

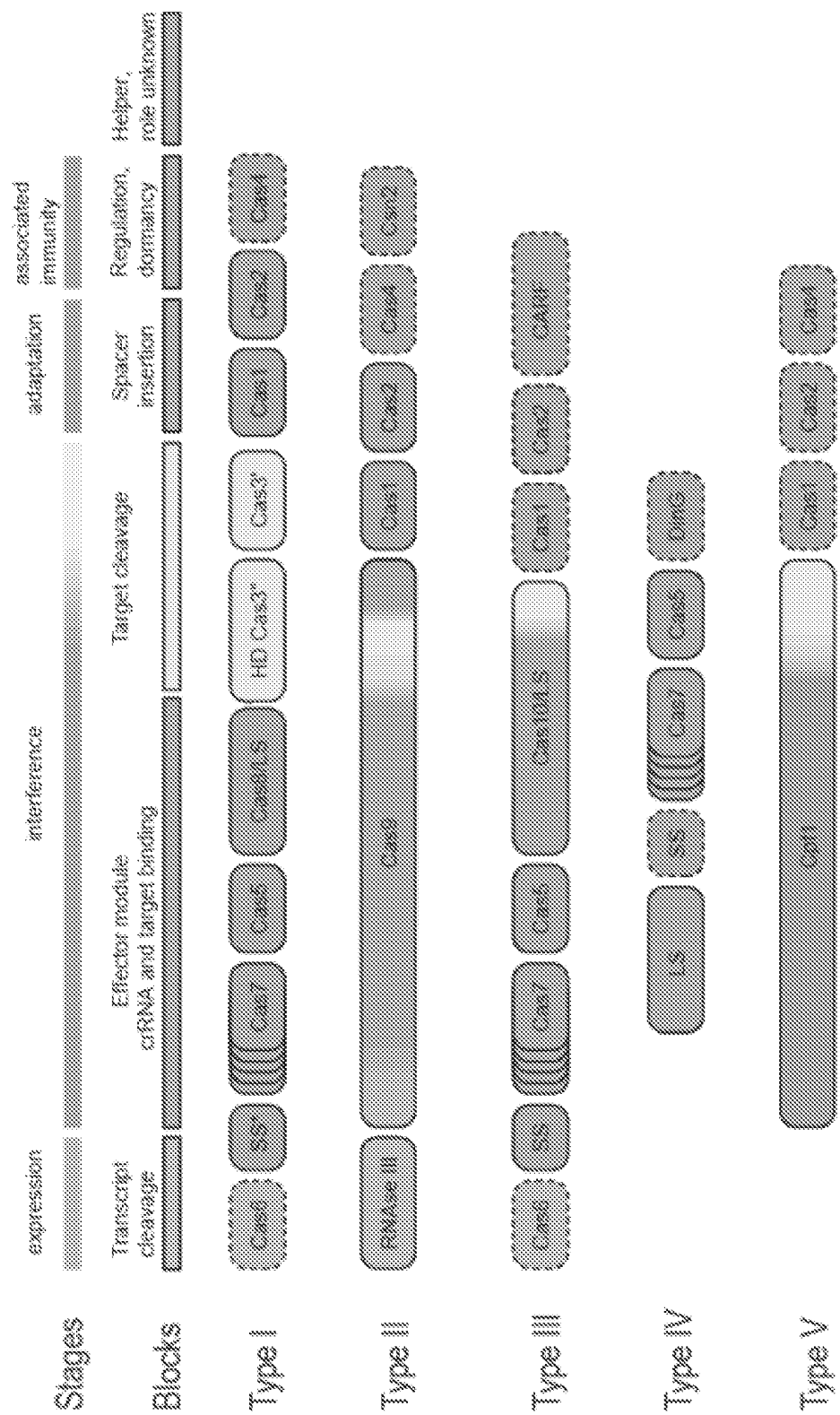


FIG. 2

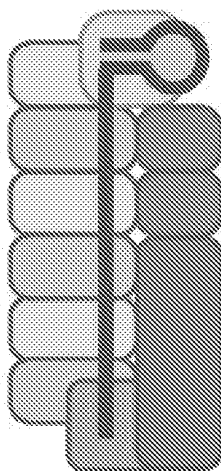


FIG. 3A

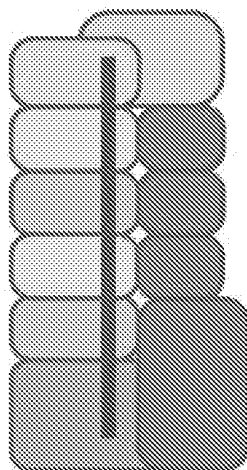


FIG. 3B

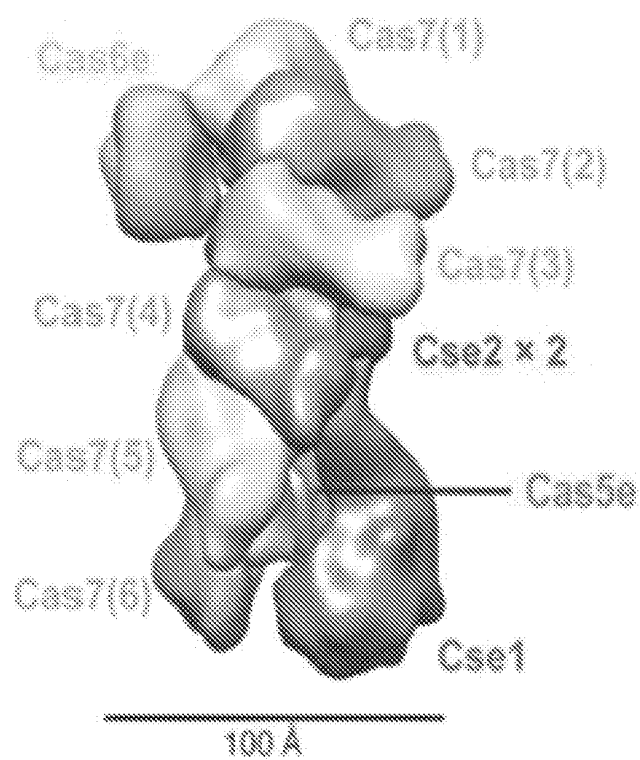


FIG. 3C

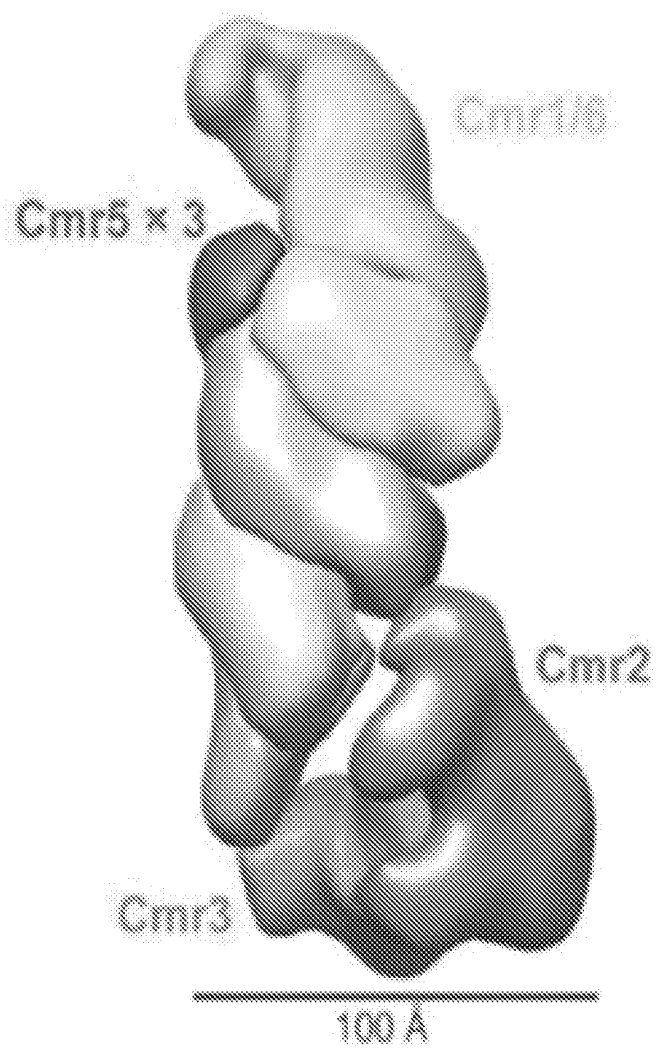
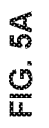
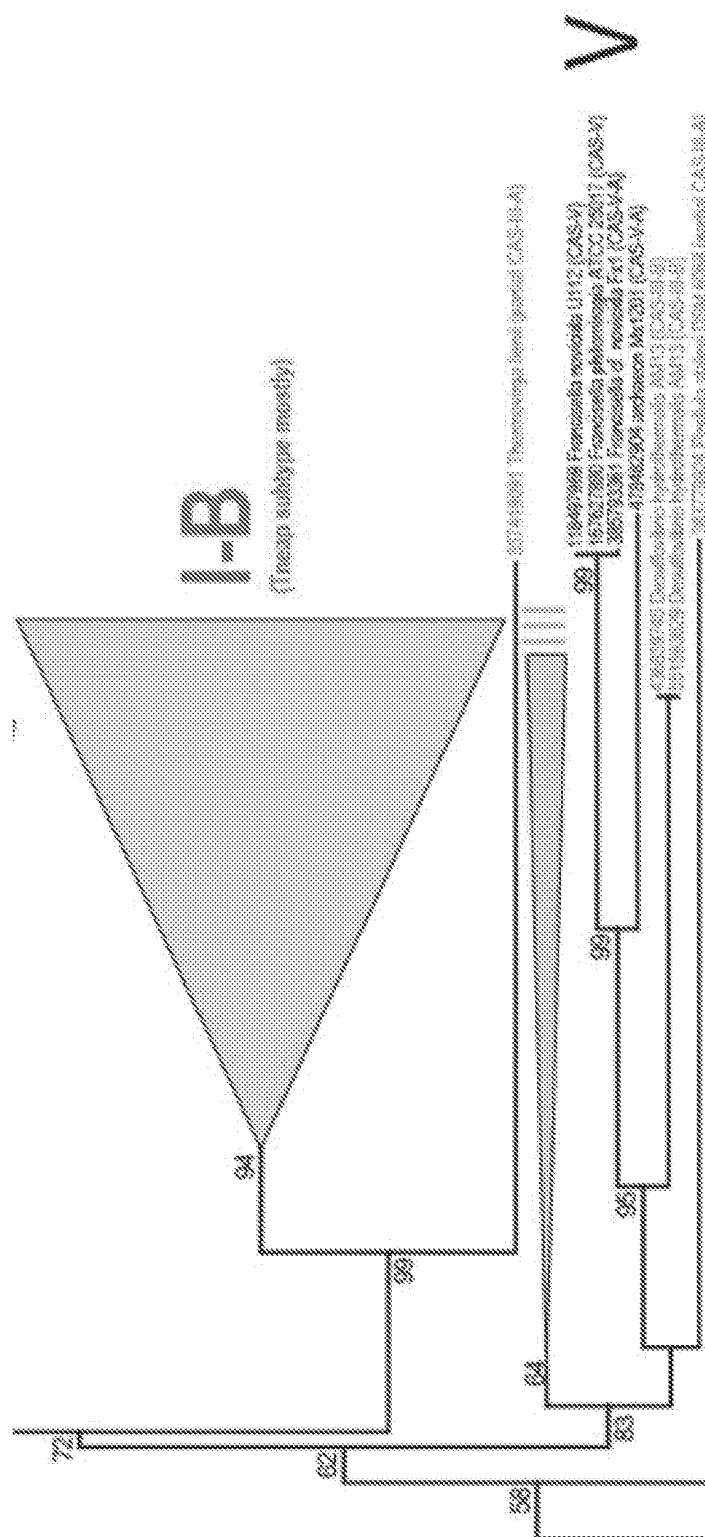


FIG. 3D





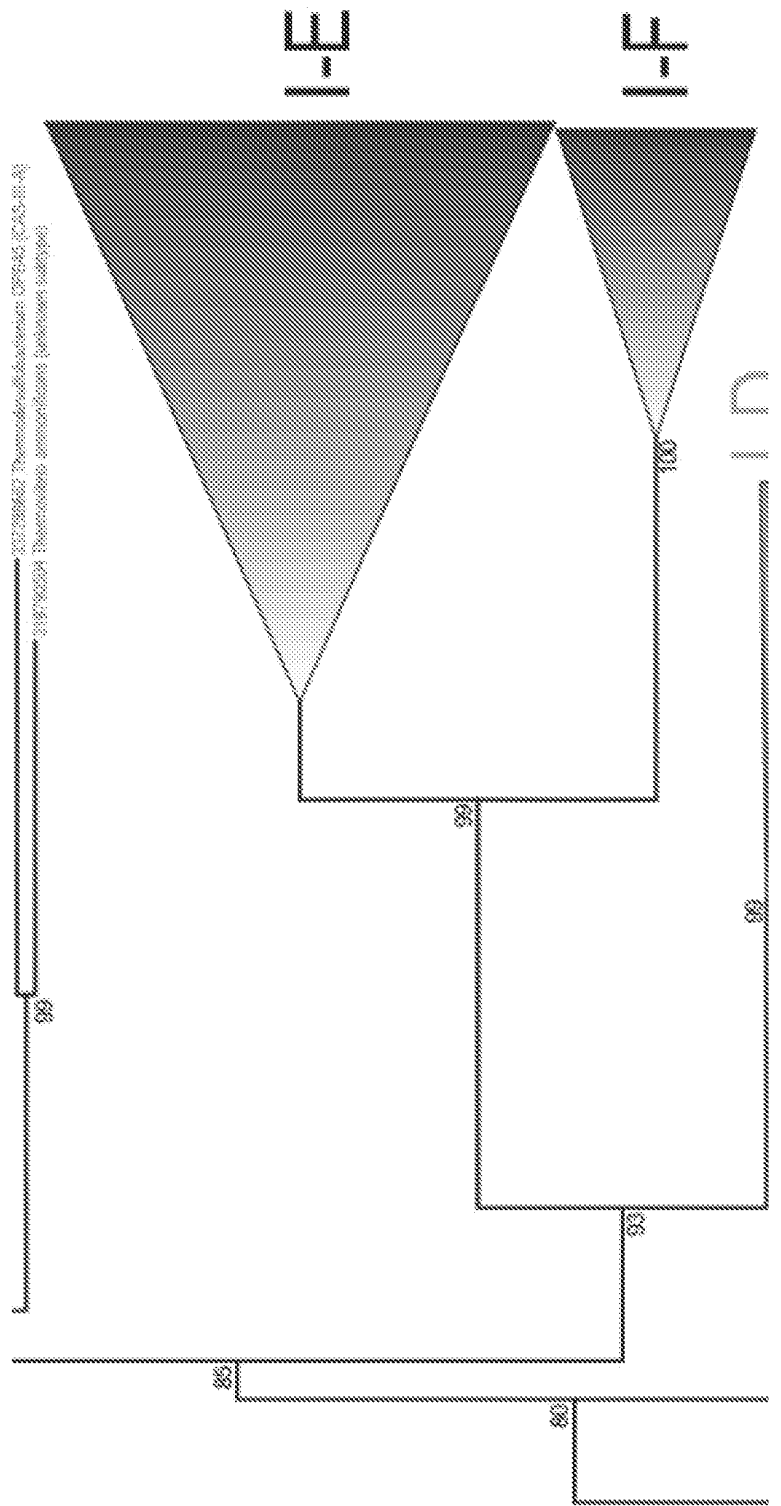


FIG. 5C

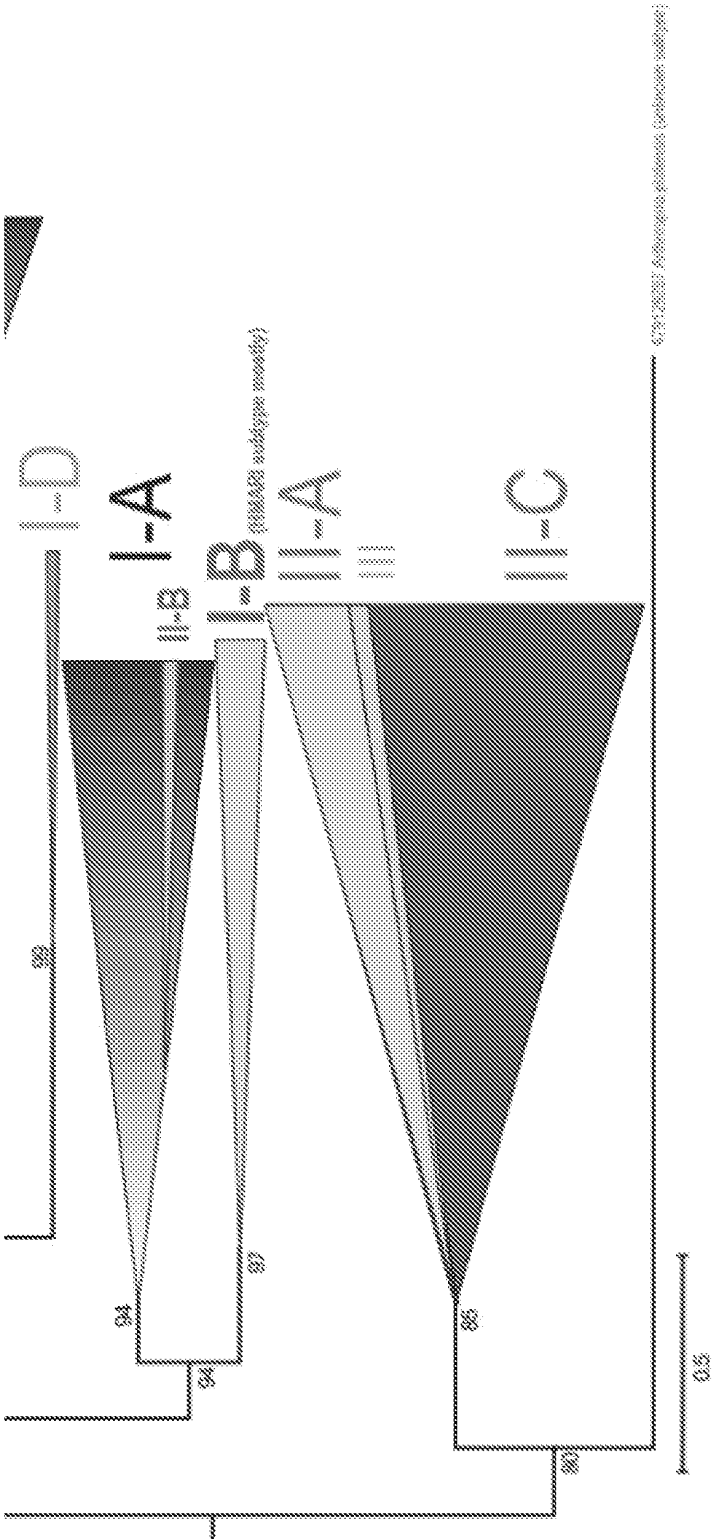
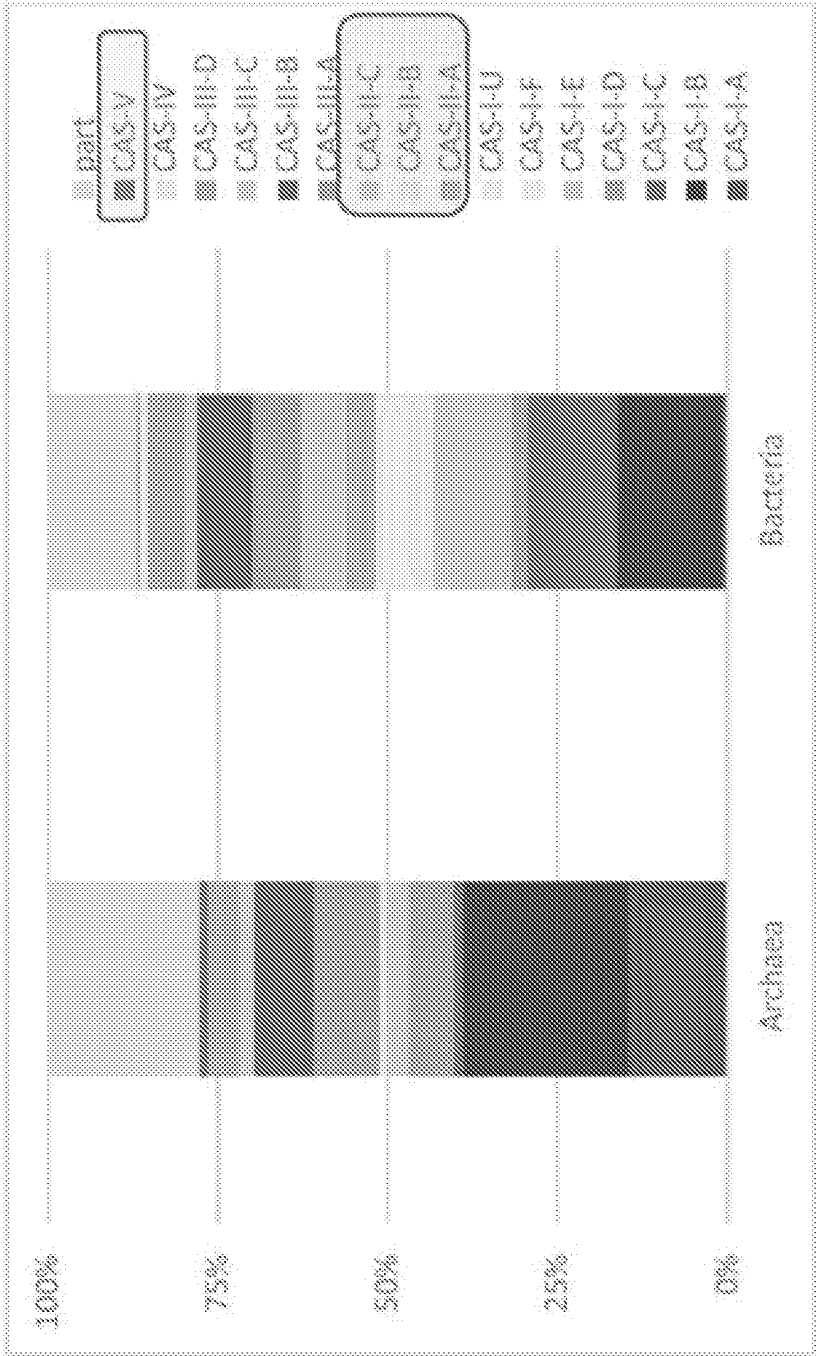


FIG. 5D



Type II: bacteria-specific, ~10% bacteria
Type V: rare – several bacteria, one archaeon

FIG. 6

Brief pipeline description

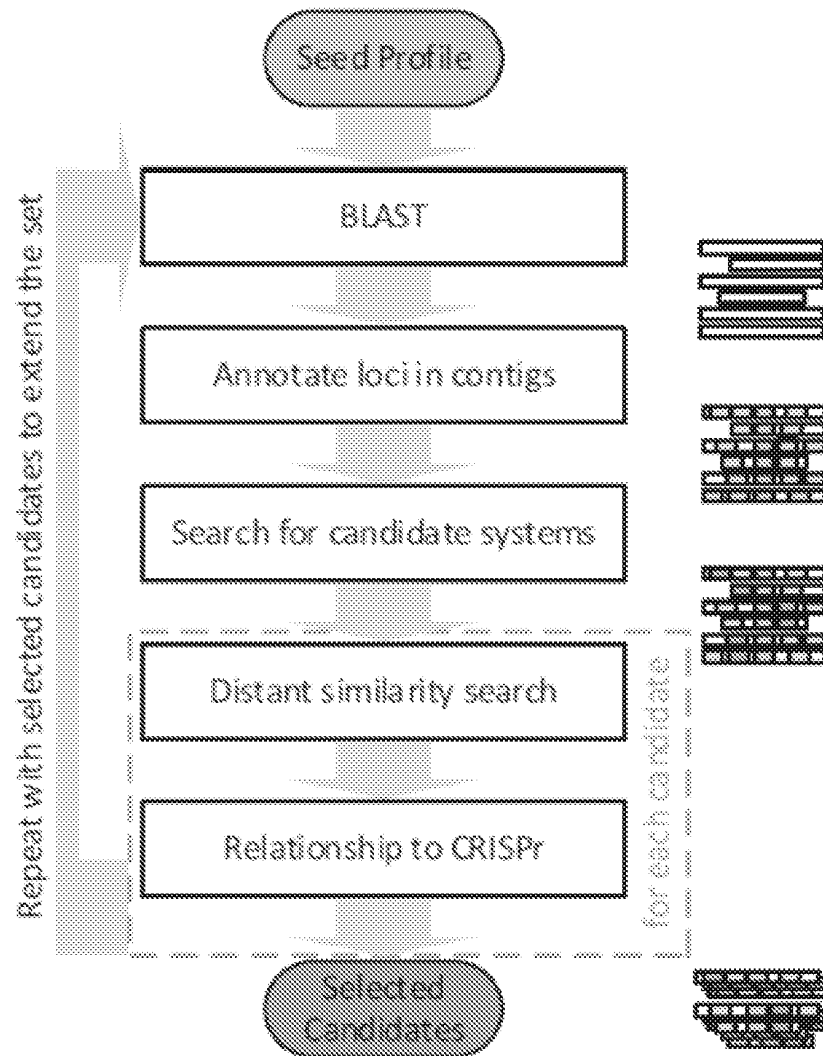


FIG. 7

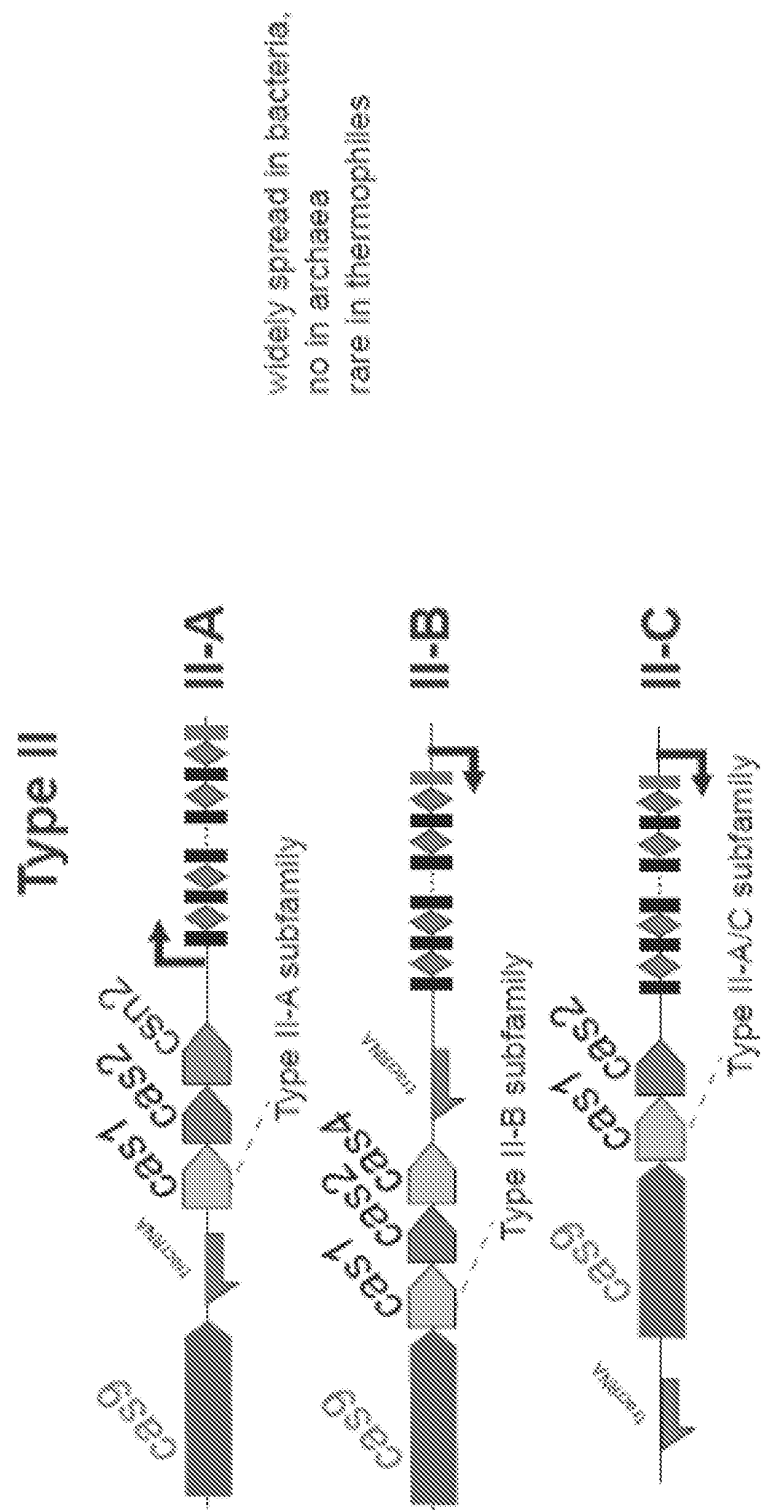
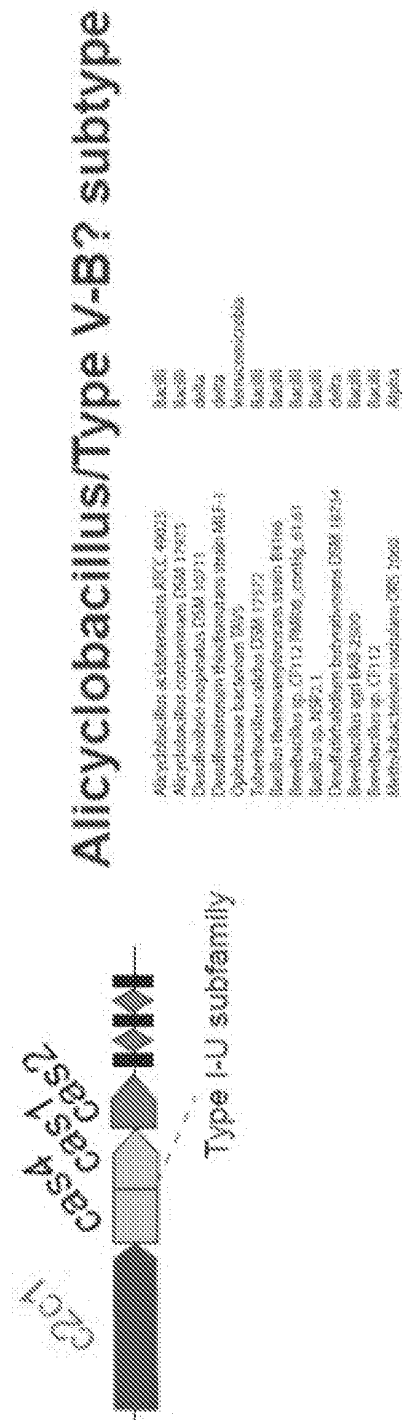
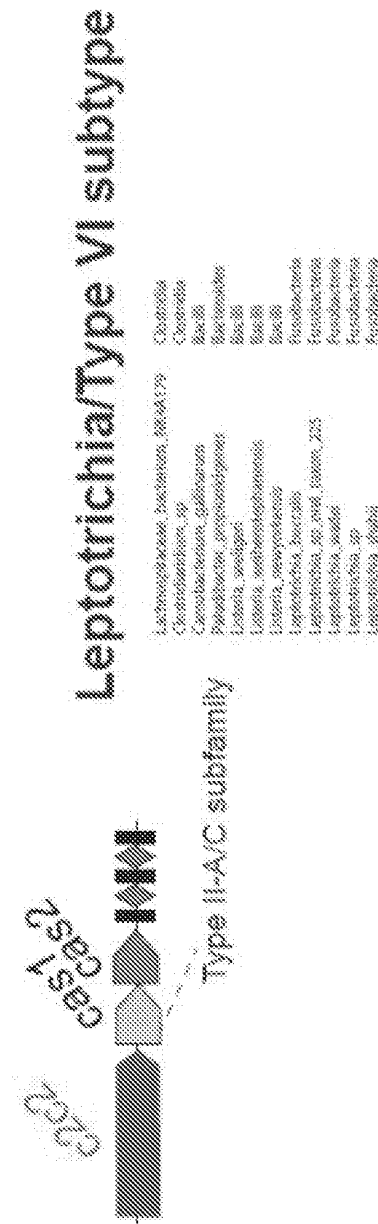


FIG. 8A



உள்ளு



၁၈၆၂

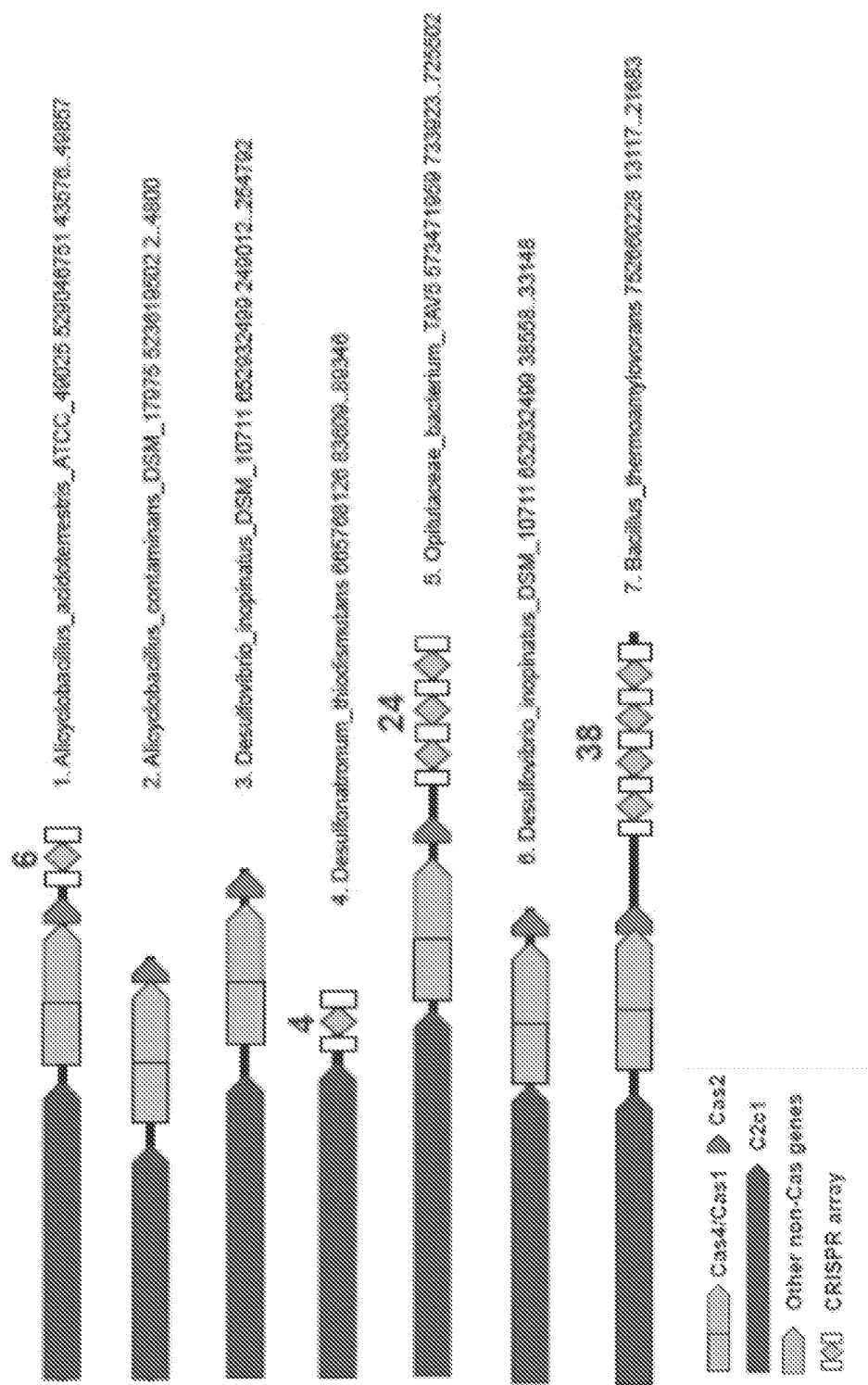


FIG. 9A

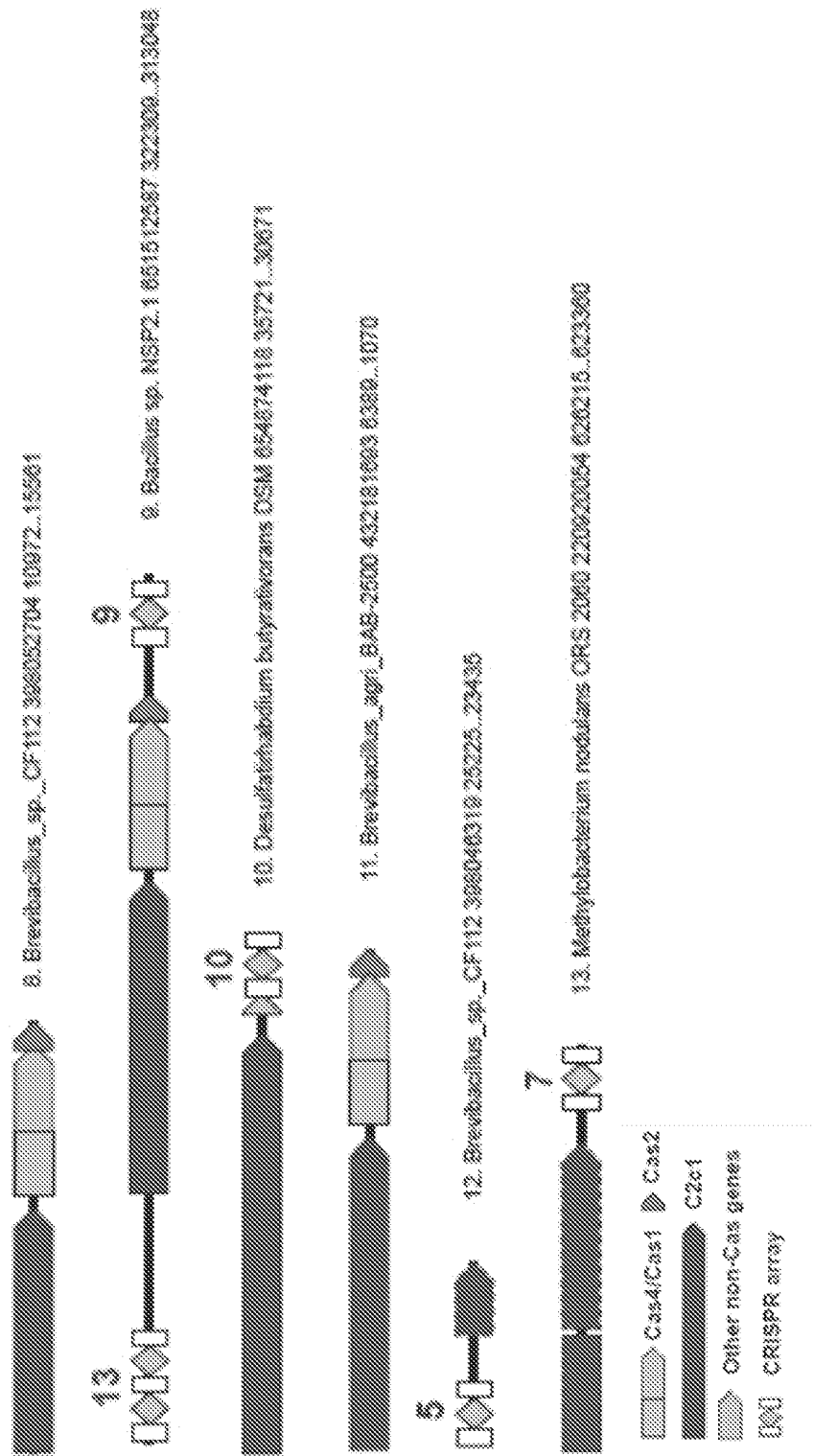


FIG. 9B

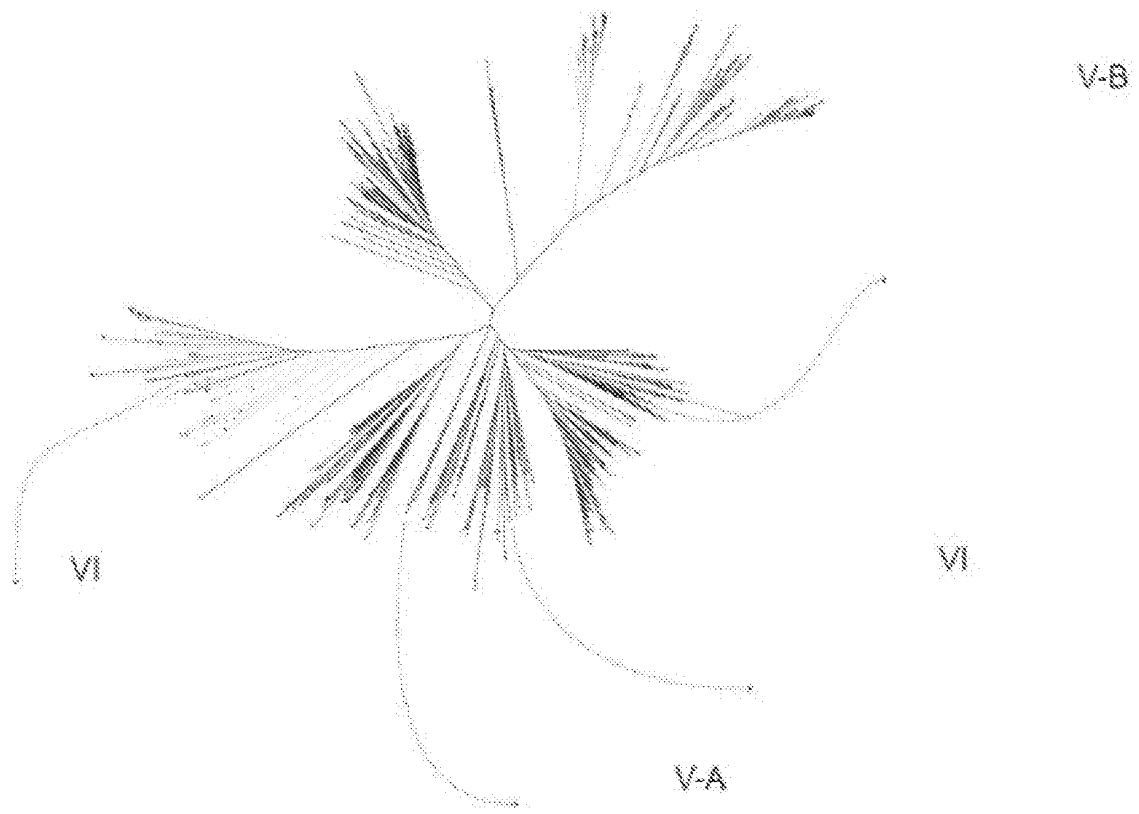
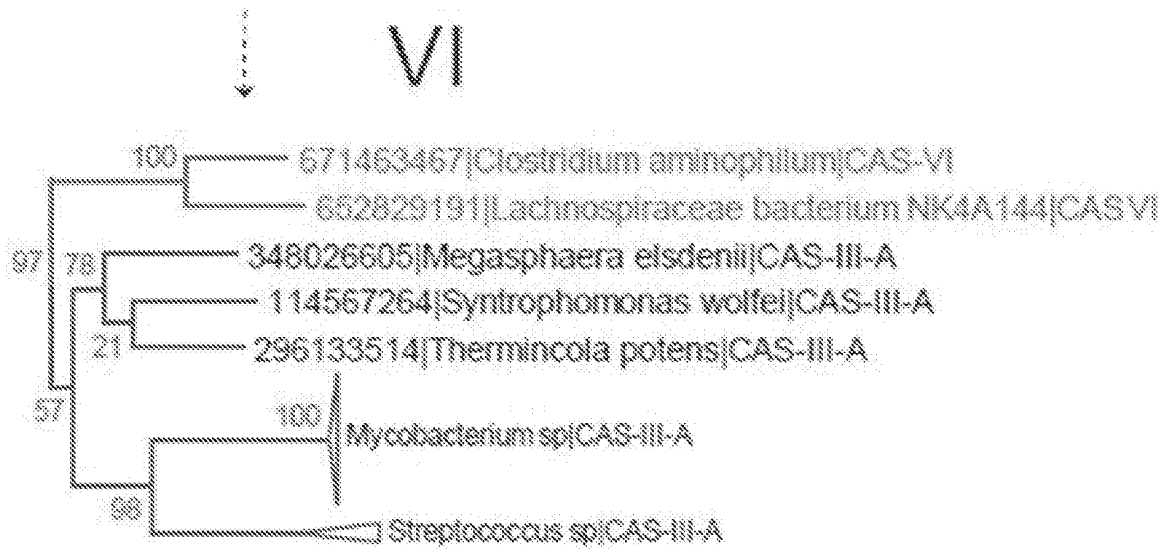


FIG. 10A



V-A

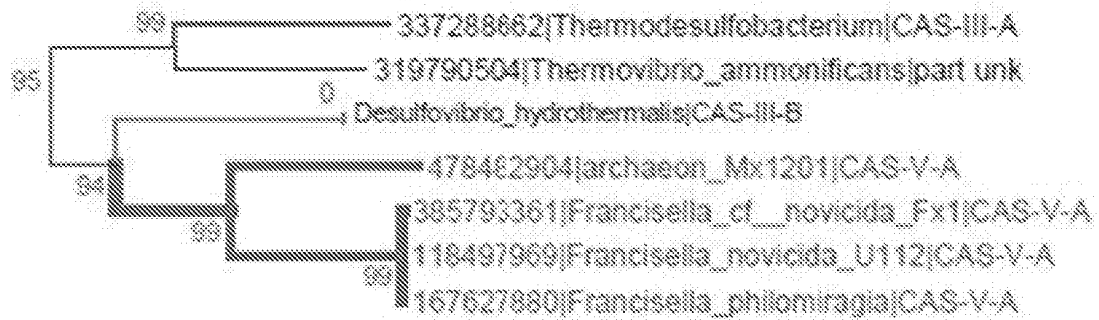
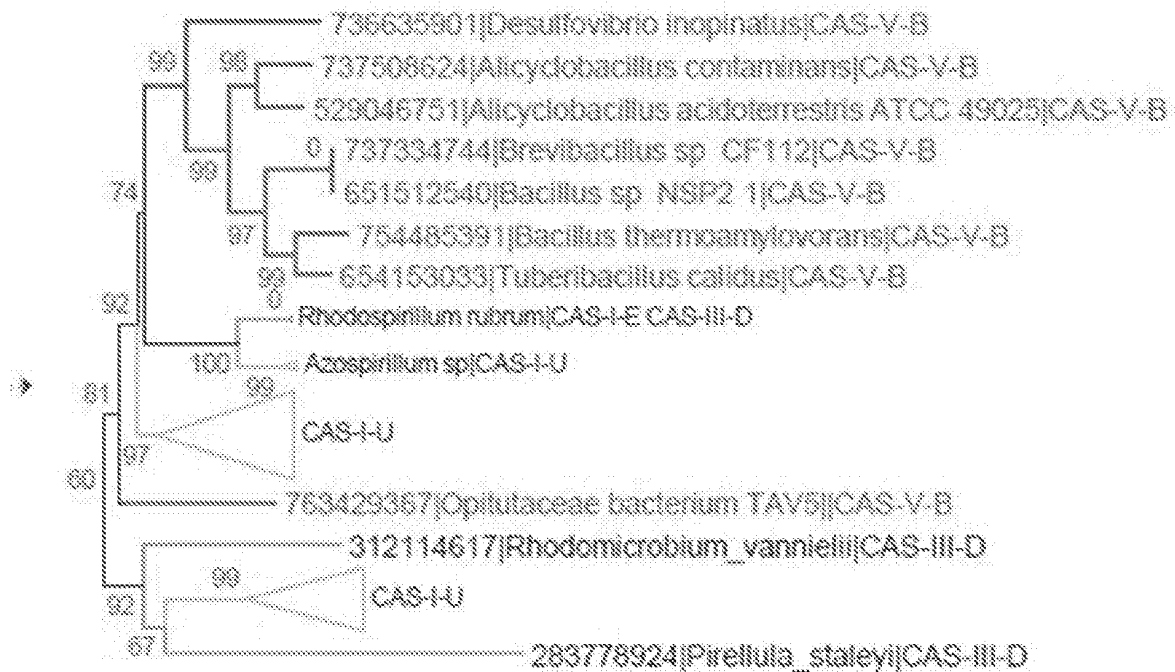


FIG. 10B

V-B



VI

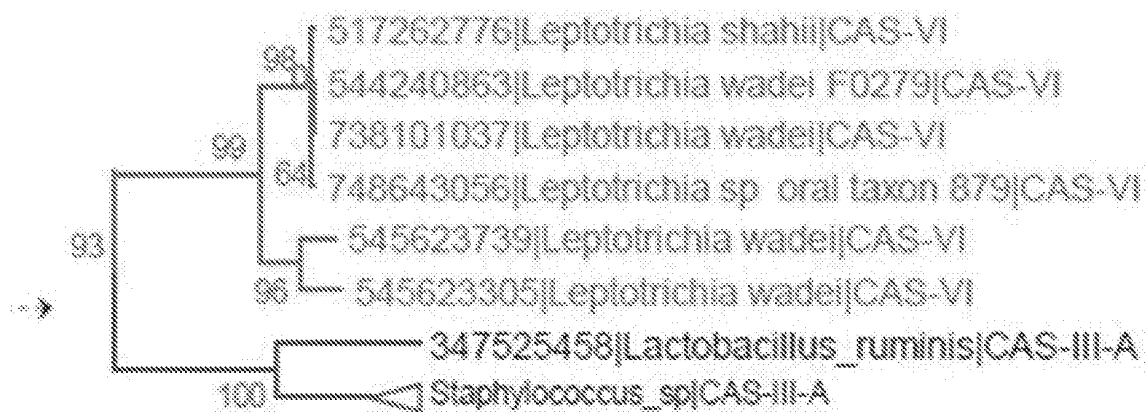


FIG. 10C

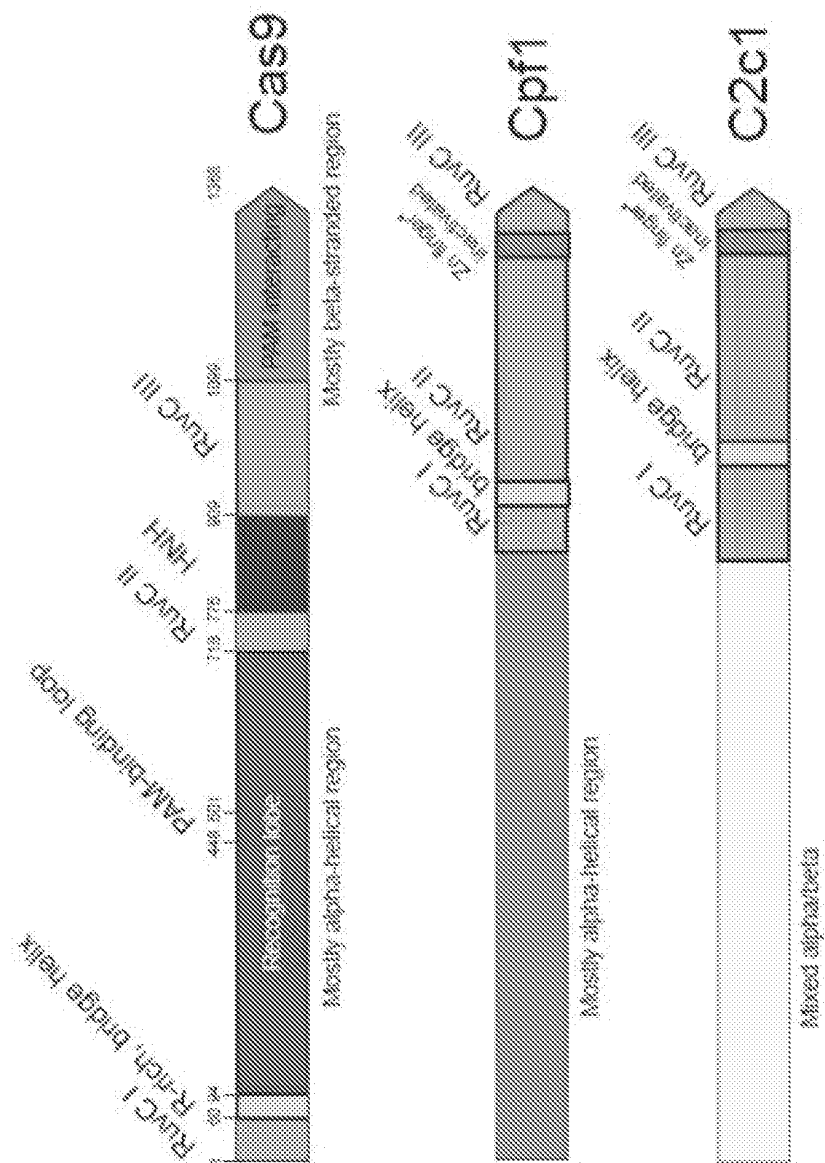


FIG. 11A

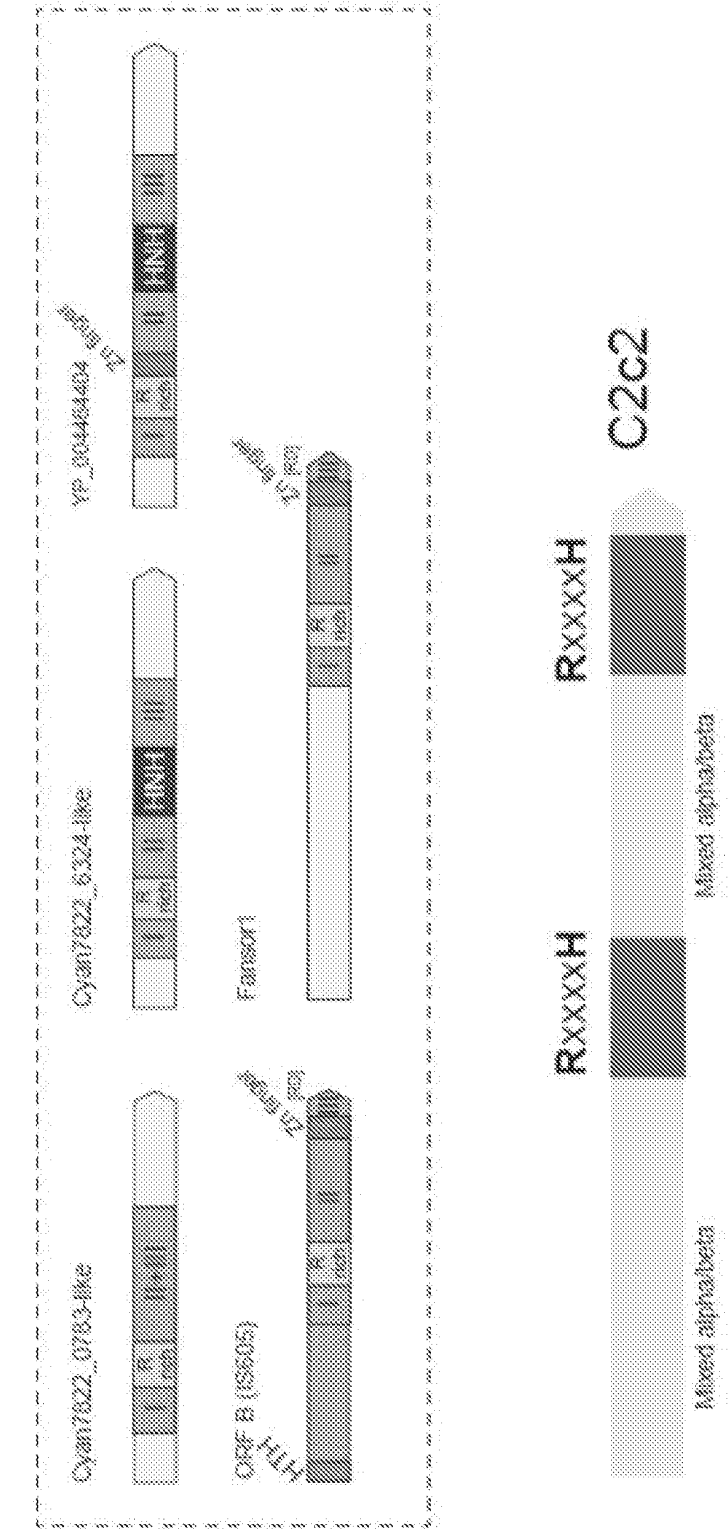


FIG. 11B

[illegible]

FIG. 12A

Case9	Case9_homologs
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

FIG. 12B

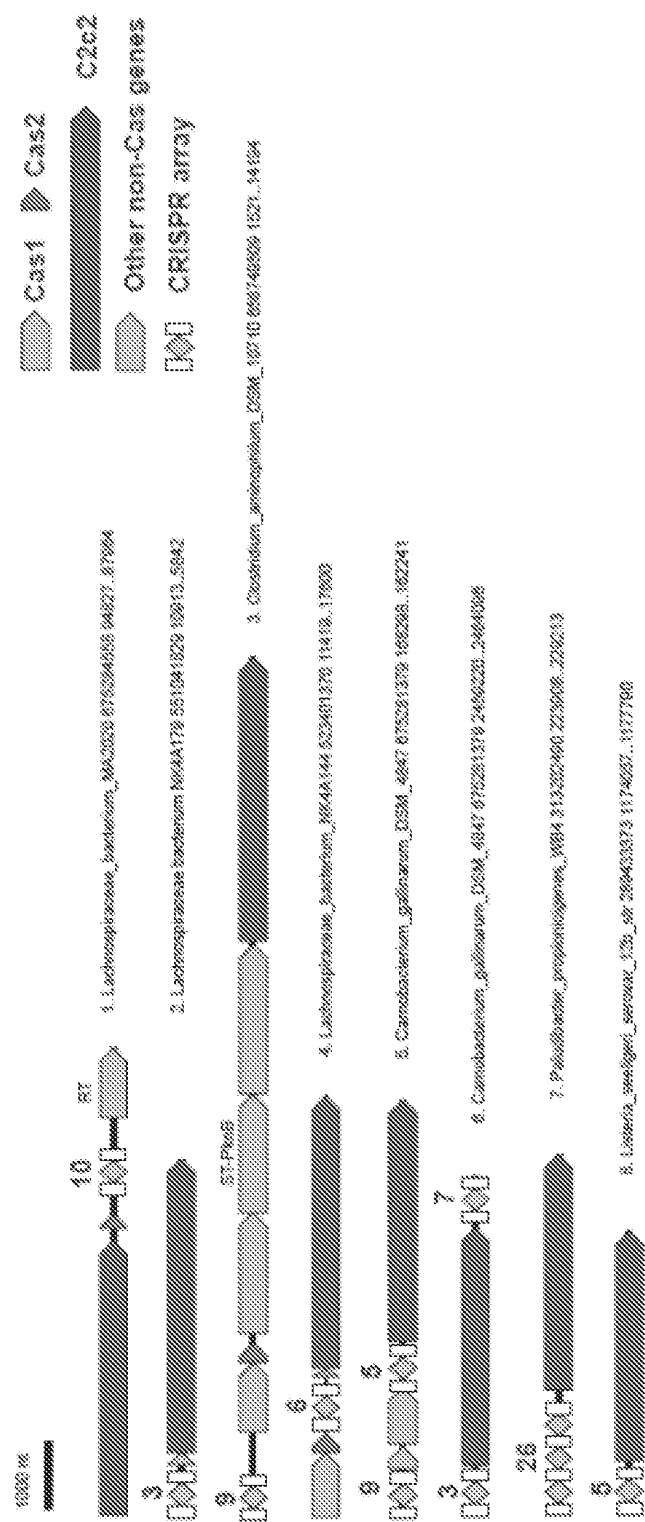


FIG. 13A

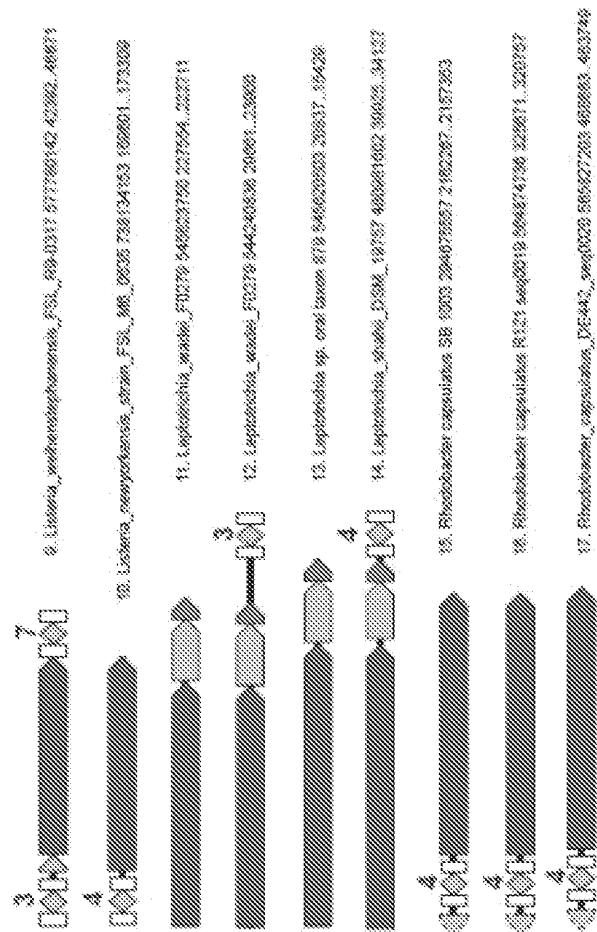


FIG. 13B

Inferred secondary structure

			XXXXXXXXXX
			<1
705395	<i>Lactococcus lactis</i> subsp. <i>lactis</i>	Ahi2/AhiF/AhiD	61 FVHHHHHH--E
190371770	<i>Lactococcus lactis</i>	Ahi2/AhiF/AhiD	60 FLAHHHHHH--E
33234700	<i>Corynebacterium diphtheriae</i> NCIC 13119	Ahi2/AhiF/AhiD	58 LLAHHHHHH--E
126178634	<i>Streptococcus lactis</i> DSM 20613	Spun_R5628	19 HHHHHHHHH--E
125424412	<i>Lactococcus lactis</i> subsp. <i>cremoris</i> DSM 1363	Spun_R5629	13 LHHHHHHHH--E
187114919	<i>Geobacter brennandii</i> DSM 1363	Spun_R5629	98 HHHHHHHHH--E
16743498	<i>Streptococcus lactis</i> str. C118	STR4199	106 FVHHHHHHHH--E
391332372	<i>Streptococcus lactis</i> str. C118	STR4199	104 FVHHHHHHHH--E
29823499	<i>Streptococcus lactis</i> str. C118	STR4199	28 FVHHHHHHHH--E
108343888	<i>Streptococcus lactis</i> str. C118	STR4199	34 FVHHHHHHHH--E
392375944	<i>Candidatus Methylobacterium</i> sp.	Sera_3948	694 GHHH--TTHHHH
146276545	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	582 TTHHHHHHHHH
345879247	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	577 GHHH--TTHHHH
210706146	<i>Escherichia coli</i> O157:H7	alk3009	209 FHHHHHHHHHH
55774553	<i>Escherichia coli</i> O157:H7	alk3009	97 FHHHHHHHHHH
17230501	<i>Escherichia coli</i> O157:H7	alk3009	97 FHHHHHHHHHH
184244822	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	9 FHHHHHHHHHH
66403013	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	9 FHHHHHHHHHH
146276545	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	9 FHHHHHHHHHH
92752344	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	764 FHHHHHHHHHH
71029722	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	584 FHHHHHHHHHH
921465571	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	13 FHHHHHHHHHH
21433643	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	21 FHHHHHHHHHH
260791233	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	18 FHHHHHHHHHH
157802267	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	582 FHHHHHHHHHH
146276545	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	110 FHHHHHHHHHH
146276545	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	775 FHHHHHHHHHH
146276545	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	11 FHHHHHHHHHH
118311131	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	17 FHHHHHHHHHH
66740878	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	22 FHHHHHHHHHH
121634053	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	6 FHHHHHHHHHH
49199884	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	10 FHHHHHHHHHH
146276545	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	9 FHHHHHHHHHH
146276545	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	1 FHHHHHHHHHH
705191249	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	1 FHHHHHHHHHH
291234929	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	1 FHHHHHHHHHH
191439999	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	13 FHHHHHHHHHH
16743498	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	10 FHHHHHHHHHH
146276545	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	14 FHHHHHHHHHH
22124749	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	32 FHHHHHHHHHH
291234929	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	94 FHHHHHHHHHH
398365293	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	9 FHHHHHHHHHH
15902826	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	17 FHHHHHHHHHH
146276545	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	16 FHHHHHHHHHH
226524784	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	16 FHHHHHHHHHH
394217453	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	208 FHHHHHHHHHH
400790360	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	208 FHHHHHHHHHH
637030412	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	126 FHHHHHHHHHH
149930787	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	110 FHHHHHHHHHH
1304567	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	130 FHHHHHHHHHH
398365293	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	262 FHHHHHHHHHH
15409855	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	218 FHHHHHHHHHH
70521919	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	214 FHHHHHHHHHH
398365293	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	317 FHHHHHHHHHH
15409855	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	381 FHHHHHHHHHH
398365293	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	339 FHHHHHHHHHH
15409855	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	295 FHHHHHHHHHH
15409855	<i>Neisseria meningitidis</i> serotype 4	Sera_3948	295 FHHHHHHHHHH

FIG. 14A

a2		a2.1		a3		a4	
33	WYLLIAMS	44	WYLLIAMS	55	WYLLIAMS	66	WYLLIAMS
34	WYLLIAMS	45	WYLLIAMS	56	WYLLIAMS	67	WYLLIAMS
35	WYLLIAMS	46	WYLLIAMS	57	WYLLIAMS	68	WYLLIAMS
36	WYLLIAMS	47	WYLLIAMS	58	WYLLIAMS	69	WYLLIAMS
37	WYLLIAMS	48	WYLLIAMS	59	WYLLIAMS	70	WYLLIAMS
38	WYLLIAMS	49	WYLLIAMS	60	WYLLIAMS	71	WYLLIAMS
39	WYLLIAMS	50	WYLLIAMS	61	WYLLIAMS	72	WYLLIAMS
40	WYLLIAMS	51	WYLLIAMS	62	WYLLIAMS	73	WYLLIAMS
41	WYLLIAMS	52	WYLLIAMS	63	WYLLIAMS	74	WYLLIAMS
42	WYLLIAMS	53	WYLLIAMS	64	WYLLIAMS	75	WYLLIAMS
43	WYLLIAMS	54	WYLLIAMS	65	WYLLIAMS	76	WYLLIAMS
44	WYLLIAMS	55	WYLLIAMS	66	WYLLIAMS	77	WYLLIAMS
45	WYLLIAMS	56	WYLLIAMS	67	WYLLIAMS	78	WYLLIAMS
46	WYLLIAMS	57	WYLLIAMS	68	WYLLIAMS	79	WYLLIAMS
47	WYLLIAMS	58	WYLLIAMS	69	WYLLIAMS	80	WYLLIAMS
48	WYLLIAMS	59	WYLLIAMS	70	WYLLIAMS	81	WYLLIAMS
49	WYLLIAMS	60	WYLLIAMS	71	WYLLIAMS	82	WYLLIAMS
50	WYLLIAMS	61	WYLLIAMS	72	WYLLIAMS	83	WYLLIAMS
51	WYLLIAMS	62	WYLLIAMS	73	WYLLIAMS	84	WYLLIAMS
52	WYLLIAMS	63	WYLLIAMS	74	WYLLIAMS	85	WYLLIAMS
53	WYLLIAMS	64	WYLLIAMS	75	WYLLIAMS	86	WYLLIAMS
54	WYLLIAMS	65	WYLLIAMS	76	WYLLIAMS	87	WYLLIAMS
55	WYLLIAMS	66	WYLLIAMS	77	WYLLIAMS	88	WYLLIAMS
56	WYLLIAMS	67	WYLLIAMS	78	WYLLIAMS	89	WYLLIAMS
57	WYLLIAMS	68	WYLLIAMS	79	WYLLIAMS	90	WYLLIAMS
58	WYLLIAMS	69	WYLLIAMS	80	WYLLIAMS	91	WYLLIAMS
59	WYLLIAMS	70	WYLLIAMS	81	WYLLIAMS	92	WYLLIAMS
60	WYLLIAMS	71	WYLLIAMS	82	WYLLIAMS	93	WYLLIAMS
61	WYLLIAMS	72	WYLLIAMS	83	WYLLIAMS	94	WYLLIAMS
62	WYLLIAMS	73	WYLLIAMS	84	WYLLIAMS	95	WYLLIAMS
63	WYLLIAMS	74	WYLLIAMS	85	WYLLIAMS	96	WYLLIAMS
64	WYLLIAMS	75	WYLLIAMS	86	WYLLIAMS	97	WYLLIAMS
65	WYLLIAMS	76	WYLLIAMS	87	WYLLIAMS	98	WYLLIAMS
66	WYLLIAMS	77	WYLLIAMS	88	WYLLIAMS	99	WYLLIAMS
67	WYLLIAMS	78	WYLLIAMS	89	WYLLIAMS	100	WYLLIAMS
68	WYLLIAMS	79	WYLLIAMS	90	WYLLIAMS	101	WYLLIAMS
69	WYLLIAMS	80	WYLLIAMS	91	WYLLIAMS	102	WYLLIAMS
70	WYLLIAMS	81	WYLLIAMS	92	WYLLIAMS	103	WYLLIAMS
71	WYLLIAMS	82	WYLLIAMS	93	WYLLIAMS	104	WYLLIAMS
72	WYLLIAMS	83	WYLLIAMS	94	WYLLIAMS	105	WYLLIAMS
73	WYLLIAMS	84	WYLLIAMS	95	WYLLIAMS	106	WYLLIAMS
74	WYLLIAMS	85	WYLLIAMS	96	WYLLIAMS	107	WYLLIAMS
75	WYLLIAMS	86	WYLLIAMS	97	WYLLIAMS	108	WYLLIAMS
76	WYLLIAMS	87	WYLLIAMS	98	WYLLIAMS	109	WYLLIAMS
77	WYLLIAMS	88	WYLLIAMS	99	WYLLIAMS	110	WYLLIAMS
78	WYLLIAMS	89	WYLLIAMS	100	WYLLIAMS	111	WYLLIAMS
79	WYLLIAMS	90	WYLLIAMS	101	WYLLIAMS	112	WYLLIAMS
80	WYLLIAMS	91	WYLLIAMS	102	WYLLIAMS	113	WYLLIAMS
81	WYLLIAMS	92	WYLLIAMS	103	WYLLIAMS	114	WYLLIAMS
82	WYLLIAMS	93	WYLLIAMS	104	WYLLIAMS	115	WYLLIAMS
83	WYLLIAMS	94	WYLLIAMS	105	WYLLIAMS	116	WYLLIAMS
84	WYLLIAMS	95	WYLLIAMS	106	WYLLIAMS	117	WYLLIAMS
85	WYLLIAMS	96	WYLLIAMS	107	WYLLIAMS	118	WYLLIAMS
86	WYLLIAMS	97	WYLLIAMS	108	WYLLIAMS	119	WYLLIAMS
87	WYLLIAMS	98	WYLLIAMS	109	WYLLIAMS	120	WYLLIAMS
88	WYLLIAMS	99	WYLLIAMS	110	WYLLIAMS	121	WYLLIAMS
89	WYLLIAMS	100	WYLLIAMS	111	WYLLIAMS	122	WYLLIAMS
90	WYLLIAMS	101	WYLLIAMS	112	WYLLIAMS	123	WYLLIAMS
91	WYLLIAMS	102	WYLLIAMS	113	WYLLIAMS	124	WYLLIAMS
92	WYLLIAMS	103	WYLLIAMS	114	WYLLIAMS	125	WYLLIAMS
93	WYLLIAMS	104	WYLLIAMS	115	WYLLIAMS	126	WYLLIAMS
94	WYLLIAMS	105	WYLLIAMS	116	WYLLIAMS	127	WYLLIAMS
95	WYLLIAMS	106	WYLLIAMS	117	WYLLIAMS	128	WYLLIAMS
96	WYLLIAMS	107	WYLLIAMS	118	WYLLIAMS	129	WYLLIAMS
97	WYLLIAMS	108	WYLLIAMS	119	WYLLIAMS	130	WYLLIAMS
98	WYLLIAMS	109	WYLLIAMS	120	WYLLIAMS	131	WYLLIAMS
99	WYLLIAMS	110	WYLLIAMS	121	WYLLIAMS	132	WYLLIAMS
100	WYLLIAMS	111	WYLLIAMS	122	WYLLIAMS	133	WYLLIAMS
101	WYLLIAMS	112	WYLLIAMS	123	WYLLIAMS	134	WYLLIAMS
102	WYLLIAMS	113	WYLLIAMS	124	WYLLIAMS	135	WYLLIAMS
103	WYLLIAMS	114	WYLLIAMS	125	WYLLIAMS	136	WYLLIAMS
104	WYLLIAMS	115	WYLLIAMS	126	WYLLIAMS	137	WYLLIAMS
105	WYLLIAMS	116	WYLLIAMS	127	WYLLIAMS	138	WYLLIAMS
106	WYLLIAMS	117	WYLLIAMS	128	WYLLIAMS	139	WYLLIAMS
107	WYLLIAMS	118	WYLLIAMS	129	WYLLIAMS	140	WYLLIAMS
108	WYLLIAMS	119	WYLLIAMS	130	WYLLIAMS	141	WYLLIAMS
109	WYLLIAMS	120	WYLLIAMS	131	WYLLIAMS	142	WYLLIAMS
110	WYLLIAMS	121	WYLLIAMS	132	WYLLIAMS	143	WYLLIAMS
111	WYLLIAMS	122	WYLLIAMS	133	WYLLIAMS	144	WYLLIAMS
112	WYLLIAMS	123	WYLLIAMS	134	WYLLIAMS	145	WYLLIAMS
113	WYLLIAMS	124	WYLLIAMS	135	WYLLIAMS	146	WYLLIAMS
114	WYLLIAMS	125	WYLLIAMS	136	WYLLIAMS	147	WYLLIAMS
115	WYLLIAMS	126	WYLLIAMS	137	WYLLIAMS	148	WYLLIAMS
116	WYLLIAMS	127	WYLLIAMS	138	WYLLIAMS	149	WYLLIAMS
117	WYLLIAMS	128	WYLLIAMS	139	WYLLIAMS	150	WYLLIAMS
118	WYLLIAMS	129	WYLLIAMS	140	WYLLIAMS	151	WYLLIAMS
119	WYLLIAMS	130	WYLLIAMS	141	WYLLIAMS	152	WYLLIAMS
120	WYLLIAMS	131	WYLLIAMS	142	WYLLIAMS	153	WYLLIAMS
121	WYLLIAMS	132	WYLLIAMS	143	WYLLIAMS	154	WYLLIAMS
122	WYLLIAMS	133	WYLLIAMS	144	WYLLIAMS	155	WYLLIAMS
123	WYLLIAMS	134	WYLLIAMS	145	WYLLIAMS	156	WYLLIAMS
124	WYLLIAMS	135	WYLLIAMS	146	WYLLIAMS	157	WYLLIAMS
125	WYLLIAMS	136	WYLLIAMS	147	WYLLIAMS	158	WYLLIAMS
126	WYLLIAMS	137	WYLLIAMS	148	WYLLIAMS	159	WYLLIAMS
127	WYLLIAMS	138	WYLLIAMS	149	WYLLIAMS	160	WYLLIAMS
128	WYLLIAMS	139	WYLLIAMS	150	WYLLIAMS	161	WYLLIAMS
129	WYLLIAMS	140	WYLLIAMS	151	WYLLIAMS	162	WYLLIAMS
130	WYLLIAMS	141	WYLLIAMS	152	WYLLIAMS	163	WYLLIAMS
131	WYLLIAMS	142	WYLLIAMS	153	WYLLIAMS	164	WYLLIAMS
132	WYLLIAMS	143	WYLLIAMS	154	WYLLIAMS	165	WYLLIAMS
133	WYLLIAMS	144	WYLLIAMS	155	WYLLIAMS	166	WYLLIAMS
134	WYLLIAMS	145	WYLLIAMS	156	WYLLIAMS	167	WYLLIAMS
135	WYLLIAMS	146	WYLLIAMS	157	WYLLIAMS	168	WYLLIAMS
136	WYLLIAMS	147	WYLLIAMS	158	WYLLIAMS	169	WYLLIAMS
137	WYLLIAMS	148	WYLLIAMS	159	WYLLIAMS	170	WYLLIAMS
138	WYLLIAMS	149	WYLLIAMS	160	WYLLIAMS	171	WYLLIAMS
139	WYLLIAMS	150	WYLLIAMS	161	WYLLIAMS	172	WYLLIAMS
140	WYLLIAMS	151	WYLLIAMS	162	WYLLIAMS	173	WYLLIAMS
141	WYLLIAMS	152	WYLLIAMS	163	WYLLIAMS	174	WYLLIAMS
142	WYLLIAMS	153	WYLLIAMS	164	WYLLIAMS	175	WYLLIAMS
143	WYLLIAMS	154	WYLLIAMS	165	WYLLIAMS	176	WYLLIAMS
144	WYLLIAMS	155	WYLLIAMS	166	WYLLIAMS	177	WYLLIAMS
145	WYLLIAMS	156	WYLLIAMS	167	WYLLIAMS	178	WYLLIAMS
146	WYLLIAMS	157	WYLLIAMS	168	WYLLIAMS	179	WYLLIAMS
147	WYLLIAMS	158	WYLLIAMS	169	WYLLIAMS	180	WYLLIAMS
148	WYLLIAMS	159	WYLLIAMS	170	WYLLIAMS	181	WYLLIAMS
149	WYLLIAMS	160	WYLLIAMS	171	WYLLIAMS	182	WYLLIAMS
150	WYLLIAMS	161	WYLLIAMS	172	WYLLIAMS	183	WYLLIAMS
151	WYLLIAMS	162	WYLLIAMS	173	WYLLIAMS	184	WYLLIAMS
152	WYLLIAMS	163	WYLLIAMS	174	WYLLIAMS	185	WYLLIAMS
153	WYLLIAMS	164	WYLLIAMS	175	WYLLIAMS	186	WYLLIAMS
154	WYLLIAMS	165	WYLLIAMS	176	WYLLIAMS	187	WYLLIAMS
155	WYLLIAMS	166	WYLLIAMS	177	WYLLIAMS	188	WYLLIAMS
156	WYLLIAMS	167	WYLLIAMS	178	WYLLIAMS	189	WYLLIAMS
157	WYLLIAMS	168	WYLLIAMS	179	WYLLIAMS	190	WYLLIAMS
158	WYLLIAMS	169	WYLLIAMS	180	WYLLIAMS	191	WYLLIAMS
159	WYLLIAMS	170	WYLLIAMS	181	WYLLIAMS	192	WYLLIAMS
160	WYLLIAMS	171	WYLLIAMS	182	WYLLIAMS	193	WYLLIAMS
161	WYLLIAMS	172	WYLLIAMS	183	WYLLIAMS	194	WYLLIAMS
162	WYLLIAMS	173	WYLLIAMS	184	WYLLIAMS	195	WYLLIAMS
163	WYLLIAMS	174	WYLLIAMS	185	WYLLIAMS	196	WYLLIAMS
164	WYLLIAMS	175	WYLLIAMS	186	WYLLIAMS	197	WYLLIAMS
165	WYLLIAMS	176	WYLLIAMS	187	WYLLIAMS	198	WYLLIAMS
166	WYLLIAMS	177	WYLLIAMS	188	WYLLIAMS	199	WYLLIAMS
167	WYLLIAMS	178	WYLLIAMS	189	WYLLIAMS	200	WYLLIAMS
168	WYLLIAMS	179	WYLLIAMS	190	WYLLIAMS	201	WYLLIAMS
169	WYLLIAMS	180	WYLLIAMS	191	WYLLIAMS	202	WYLLIAMS
170	WYLLIAMS	181	WYLLIAMS	192	WYLLIAMS	203	WYLLIAMS
171	WYLLIAMS	182	WYLLIAMS	193	WYLLIAMS	204	WYLLIAMS
172	WYLLIAMS	183	WYLLIAMS	194	WYLLIAMS	205	WYLLIAMS
173	WYLLIAMS	184	WYLLIAMS	195	WYLLIAMS	206	WYLLIAMS
174	WYLLIAMS	185	WYLLIAMS	196	WYLLIAMS	207	WYLLIAMS
175	WYLLIAMS	186	WYLLIAMS	197	WYLLIAMS	208	WYLLIAMS
176	WYLLIAMS	187	WYLLIAMS	198	WYLLIAMS	209	WYLLIAMS
177	WYLLIAMS	188	WYLLIAMS	199	WYLLIAMS	210	WYLLIAMS
178	WYLLIAMS	189	WYLLIAMS	200	WYLLIAMS	211	WYLLIAMS
179	WYLLIAMS	190	WYLLIAMS	201	WYLLIAMS	212	WYLLIAMS
180	WYLLIAMS	191	WYLLIAMS	202	WYLLIAMS	213	WYLLIAMS
181	WYLLIAMS	192	WYLLIAMS	203	WYLLIAMS	214	WYLLIAMS
182	WYLLIAMS	193	WYLLIAMS	204	WYLLIAMS	215	WYLLIAMS
183	WYLLIAMS	194	WYLLIAMS	205	WYLLIAMS	216	WYLLIAMS
184	WYLLIAMS	195	WYLLIAMS	206	WYLLIAMS	217	WYLLIAMS
185	WYLLIAMS	196	WYLLIAMS	207	WYLLIAMS	218	WYLLIAMS
186	WYLLIAMS	197	WYLLIAMS	208	WYLLIAMS	219	WYLLIAMS
187	WYLLIAMS	198	WYLLIAMS	209	WYLLIAMS	220	WYLLIAMS
188	WYLLIAMS	199	WYLLIAMS	210	WYLLIAMS	221	WYLLIAMS
189	WYLLIAMS	200	WYLLIAMS	211	WYLLIAMS	222	WYLLIAMS
190	WYLLIAMS	201	WYLLIAMS	212	WYLLIAMS	223	WYLLIAMS
191	WYLLIAMS	202	WYLLIAMS	213	WYLLIAMS	224	WYLLIAMS
192	WYLLIAMS	203	WYLLIAMS	214	WYLLIAMS	225	WYLLIAMS
193	WYLLIAMS	204	WYLLIAMS	215	WYLLIAMS	226	WYLLIAMS
194	WYLLIAMS	205	WYLLIAMS	216	WYLLIAMS	227	WYLLIAMS
195	WYLLIAMS	206	WYLLIAMS	217	WYLLIAMS	228	WYLLIAMS
196	WYLLIAMS	207	WYLLIAMS	218	WYLLIAMS	229	WYLLIAMS
197	WYLLIAMS	208	WYLLIAMS	219	WYLLIAMS	230	WYLLIAMS

FIG. 14B

305791853	<i>Oedillochloa trichoides</i> D6-6	Cas6 (Cas_Cas02718)	279	LRNEMFARAR
62360331	<i>Hom sapiens</i>	C6orf70	107	PRFPHHPRAC
103104044	<i>Enterococcus histolytica</i> HM-1-IMSE	C6orf70	88	WPHHPPHICPH
307601154	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> 80303	DUF4209	473	TDHSHHPLHPC
213441893	<i>Cyanothera</i> sp. DCC 7424	DUF4209	463	LHSHHSHHSHD
154908293	<i>Salmonella enteritidis</i> PT	HEPN-T	3	--HSHHSHH-T
103103515	<i>Streptococcus enteritidis</i> str. IMC	HEPN-T (Darep1/8)	44	--YHSHHSHL-T
154908293	<i>Hom sapiens</i>	ecovin	10	RHSHHSHHSHH
27573278	<i>Neisseria meningitidis</i>	HEPN-M	5	SHSHHSHHSHH
73555476	<i>Thermus thermophilus</i> HB8	HEPN-M	1	-----HSHHSHH
321149183	<i>Shadococcus sp.</i> 1533	HEPN-M (SAV 6187)	12	VHSHHSHHSHH
293103793	<i>Streptococcus enteritidis</i> HM-1589	HEPN-M (SAV 6187)	43	LHSHHSHHSHH
1837771	<i>Staphylococcus aureus</i>	Aminoglycoside_NT_C	137	SHSHHSHHSHH
282783117	<i>Enterococcus faecium</i>	Aminoglycoside_NT_C	146	SHSHHSHHSHH
110443294	<i>Escherichia coli</i> 536	glnD/glnE/DUF254_C	48	--SHSHHSHHSHH
2491031704	<i>Salmonella enteritidis</i> ATCC 3396	Aminoglycoside_NT_C (DUF4037)	172	SHSHHSHHSHH
183103337	<i>Streptococcus enteritidis</i> ATCC	Aminoglycoside_NT_C (DUF4037)	137	SHSHHSHHSHH
154908293	<i>Flavobacterium pyrophilum</i> JF01-84	AshV	17	SHSHHSHHSHH
103103169	<i>Lactococcus lactis</i> subsp. <i>cremoris</i> HM1563	AshV	33	SHSHHSHHSHH
170721573	<i>Streptococcus enteritidis</i> HM15	AshV	13	SHSHHSHHSHH
305791853	<i>Salmonella enteritidis</i> ATCC 3396	AshA_CTD	509	SHSHHSHHSHH
403015463	<i>Lactobacillus helveticus</i> SAG51	AshA_CTD	507	SHSHHSHHSHH
149103	<i>Lactococcus lactis</i>	AshA_CTD	494	SHSHHSHHSHH
305270339	<i>Streptococcus</i> sp. T01-81	Integron cassette HEPN	11	SHSHHSHHSHH
354090384	<i>Vibrio parahaemolyticus</i>	Integron cassette HEPN	32	SHSHHSHHSHH
157162099	<i>Escherichia coli</i> 53	BC042_2821	13	SHSHHSHHSHH
403012344	<i>Geobacter sulfurreducens</i> KM402	BC042_2821	9	SHSHHSHHSHH
213441893	<i>Streptococcus enteritidis</i> ATCC	BC042_2821	42	SHSHHSHHSHH
243103301	<i>Leptospira interrogans</i> str. 14601	LA2681	109	SHSHHSHHSHH
321149183	<i>Shigella boydii</i> Sh27	LA2681	109	SHSHHSHHSHH
257084905	<i>Enterococcus faecalis</i> F101	LA2681	109	SHSHHSHHSHH
103103169	<i>Clostridium thermocellum</i> ATCC 27405	CthA_2314	42	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	CthA_2314	41	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	Rec_C0808	15	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	Rec_C0808	15	SHSHHSHHSHH
143103169	<i>Micrococcus aeruginosus</i> H123-143	Yah	112	SHSHHSHHSHH
305270339	<i>Vibrio parahaemolyticus</i> HM1563	Yah	139	SHSHHSHHSHH
125623645	<i>Lactococcus lactis</i> subsp. <i>cremoris</i> HM1563	Yah	134	SHSHHSHHSHH
157162099	<i>Neisseria meningitidis</i>	DZ1P3/HM1138	39	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	DZ1P3/HM1138	31	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	DZ1P3/HM1138	418	SHSHHSHHSHH
157162099	<i>Neisseria meningitidis</i>	HNAse_L3	214	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	HNAse_L3	214	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	HNAse_L3	214	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	ApeA/HNAse11217	81	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	ApeA/HNAse11217	180	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	ApeA/HNAse11217	85	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	Ash02	28	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	Ash02	98	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	DUF2526	1	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	DUF2526	1	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	REN (RnnseL/Irel)	505	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	REN (RnnseL/Irel)	703	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	FrrC/BioC/AFPCO1_4465	443	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	FrrC/BioC/AFPCO1_4465	377	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	FrrC/BioC/AFPCO1_4465	395	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	ERG6_01251	392	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	ERG6_01251	391	SHSHHSHHSHH
305270339	<i>Streptococcus enteritidis</i> ATCC 3396	ERG6_01251	392	SHSHHSHHSHH

consensus/60%

...hh,hh...h.

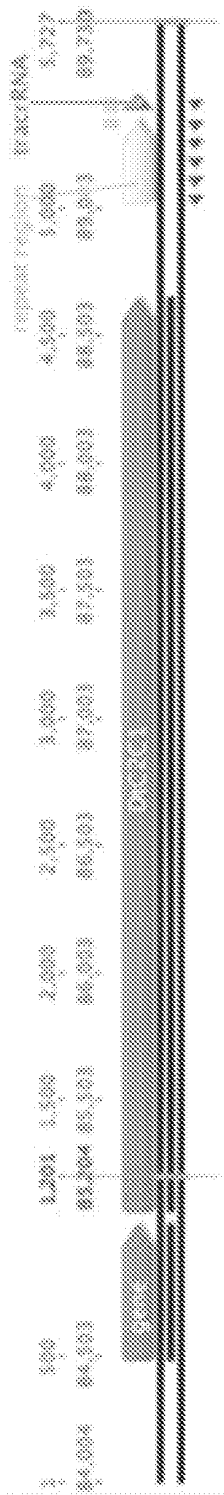
FIG. 14C

4	WAGNER	40	WAGNER	11	WAGNER	8	WAGNER	409
5	WAGNER	18	WAGNER	11	WAGNER	10	WAGNER	215
11	WAGNER	18	WAGNER	11	WAGNER	10	WAGNER	220
6	WAGNER	26	WAGNER	11	WAGNER	10	WAGNER	587
7	WAGNER	24	WAGNER	11	WAGNER	10	WAGNER	577
19	WAGNER	27	WAGNER	4	WAGNER	10	WAGNER	119
15	WAGNER	19	WAGNER	5	WAGNER	10	WAGNER	158
12	WAGNER	20	WAGNER	9	WAGNER	12	WAGNER	134
31	WAGNER	21	WAGNER	4	WAGNER	12	WAGNER	139
20	WAGNER	17	WAGNER	8	WAGNER	8	WAGNER	131
12	WAGNER	19	WAGNER	9	WAGNER	8	WAGNER	118
13	WAGNER	19	WAGNER	9	WAGNER	5	WAGNER	140
19	WAGNER	14	WAGNER	8	WAGNER	3	WAGNER	232
15	WAGNER	18	WAGNER	8	WAGNER	1	WAGNER	356
43	WAGNER	15	WAGNER	9	WAGNER	10	WAGNER	210
14	WAGNER	13	WAGNER	4	WAGNER	3	WAGNER	271
14	WAGNER	13	WAGNER	4	WAGNER	8	WAGNER	255
15	WAGNER	30	WAGNER	49	WAGNER	14	WAGNER	188
13	WAGNER	27	WAGNER	20	WAGNER	15	WAGNER	170
15	WAGNER	23	WAGNER	20	WAGNER	15	WAGNER	155
8	WAGNER	20	WAGNER	5	WAGNER	17	WAGNER	622
8	WAGNER	19	WAGNER	7	WAGNER	18	WAGNER	640
9	WAGNER	23	WAGNER	9	WAGNER	18	WAGNER	620
39	WAGNER	13	WAGNER	0	WAGNER	26	WAGNER	147
40	WAGNER	13	WAGNER	0	WAGNER	23	WAGNER	147
40	WAGNER	39	WAGNER	7	WAGNER	4	WAGNER	174
18	WAGNER	38	WAGNER	7	WAGNER	10	WAGNER	145
25	WAGNER	14	WAGNER	7	WAGNER	4	WAGNER	178
45	WAGNER	10	WAGNER	37	WAGNER	22	WAGNER	479
46	WAGNER	13	WAGNER	38	WAGNER	31	WAGNER	486
46	WAGNER	7	WAGNER	50	WAGNER	31	WAGNER	485
53	WAGNER	7	WAGNER	25	WAGNER	26	WAGNER	235
42	WAGNER	8	WAGNER	15	WAGNER	23	WAGNER	224
24	WAGNER	56	WAGNER	18	WAGNER	7	WAGNER	155
24	WAGNER	27	WAGNER	17	WAGNER	8	WAGNER	156
19	WAGNER	12	WAGNER	21	WAGNER	31	WAGNER	235
19	WAGNER	7	WAGNER	27	WAGNER	14	WAGNER	242
19	WAGNER	7	WAGNER	29	WAGNER	14	WAGNER	259
30	WAGNER	13	WAGNER	6	WAGNER	13	WAGNER	141
41	WAGNER	4	WAGNER	3	WAGNER	13	WAGNER	172
44	WAGNER	9	WAGNER	4	WAGNER	13	WAGNER	542
15	WAGNER	23	WAGNER	13	WAGNER	9	WAGNER	345
13	WAGNER	23	WAGNER	14	WAGNER	9	WAGNER	344
15	WAGNER	27	WAGNER	18	WAGNER	9	WAGNER	344
10	WAGNER	42	WAGNER	53	WAGNER	10	WAGNER	246
9	WAGNER	38	WAGNER	15	WAGNER	12	WAGNER	422
10	WAGNER	74	WAGNER	15	WAGNER	9	WAGNER	344
19	WAGNER	22	WAGNER	43	WAGNER	51	WAGNER	324
15	WAGNER	27	WAGNER	47	WAGNER	43	WAGNER	310
14	WAGNER	5	WAGNER	2	WAGNER	8	WAGNER	75
19	WAGNER	6	WAGNER	2	WAGNER	11	WAGNER	79
8	WAGNER	10	WAGNER	31	WAGNER	12	WAGNER	477
9	WAGNER	12	WAGNER	15	WAGNER	14	WAGNER	411
10	WAGNER	13	WAGNER	16	WAGNER	13	WAGNER	745
5	WAGNER	10	WAGNER	19	WAGNER	14	WAGNER	473
9	WAGNER	10	WAGNER	15	WAGNER	14	WAGNER	394
15	WAGNER	14	WAGNER	5	WAGNER	9	WAGNER	506

FIG. 14D

[illegible][illegible]

Journal of Management Inquiry 18(6)



DR: GUGCCAAUACCCACACUGACCAAGCUUGCCGAGAC

tracrRNA:
CUUGGGGAAAGCUAGGCAAGUUUUGGAUGAUAGAAUAUCAUGUCACAGGAGGA
GUUUUU

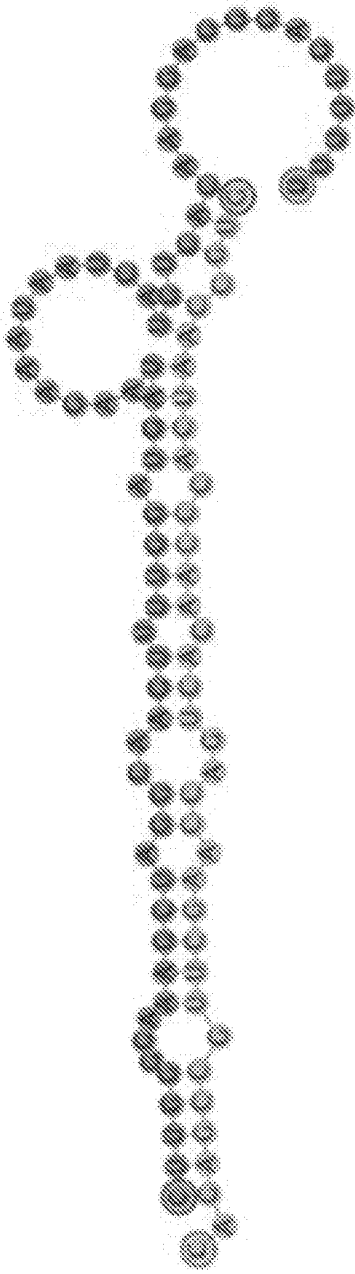


FIG. 16

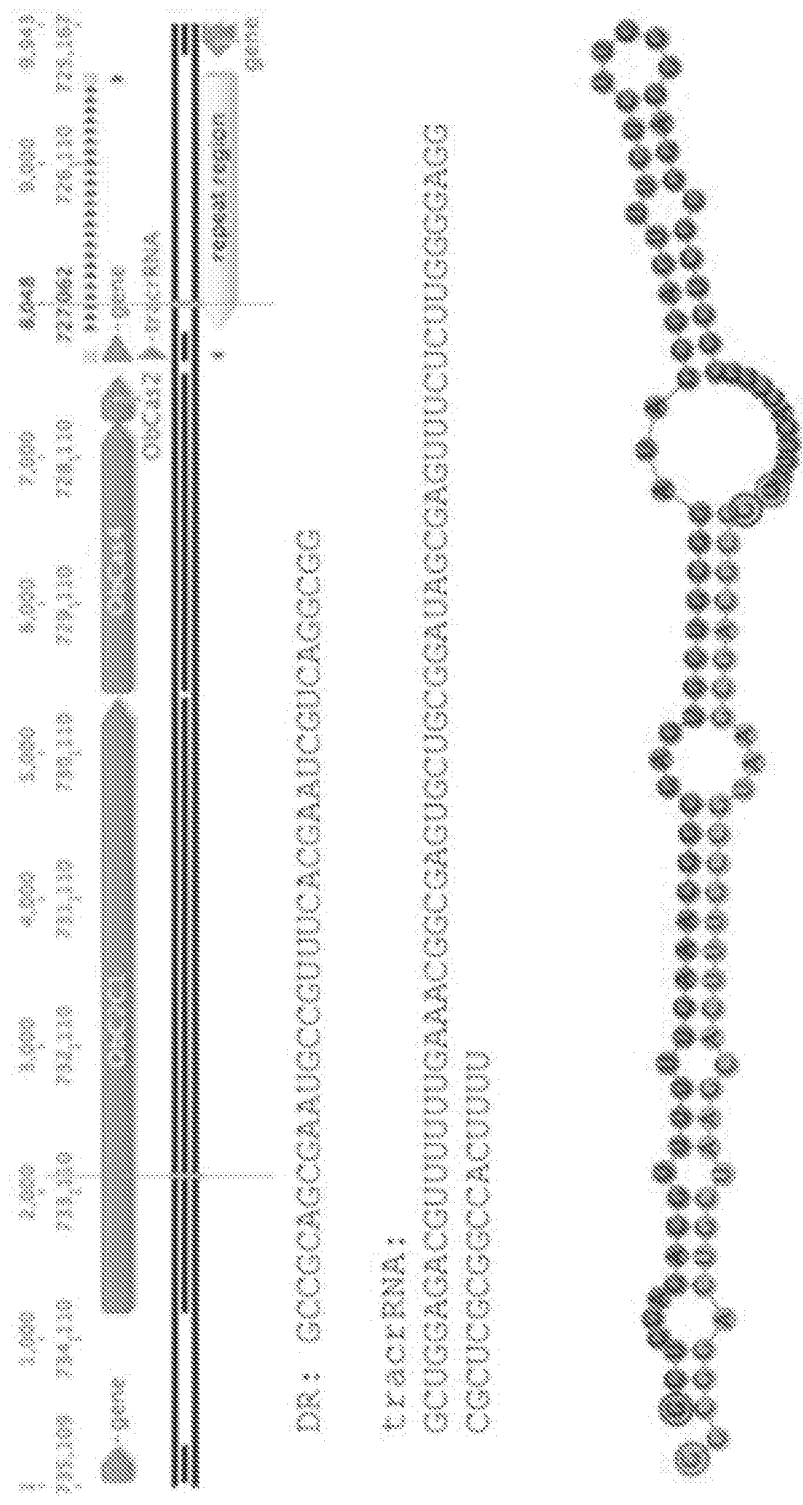
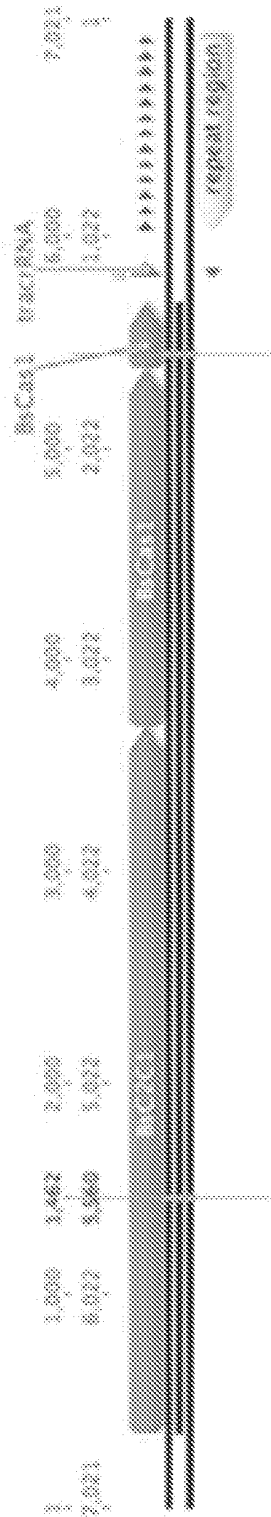


FIG. 17



DR: GUUCGAAAGCCUUAGUGGAAAGCCUUCGUGGUUAGCAC

trace RNA:

CACCGAUAUAUCACGACUUUCCACUAAGCUUUCGAAUUUAUGAUGCGAGCAUCCUCUC
AGGUCAAAAAA

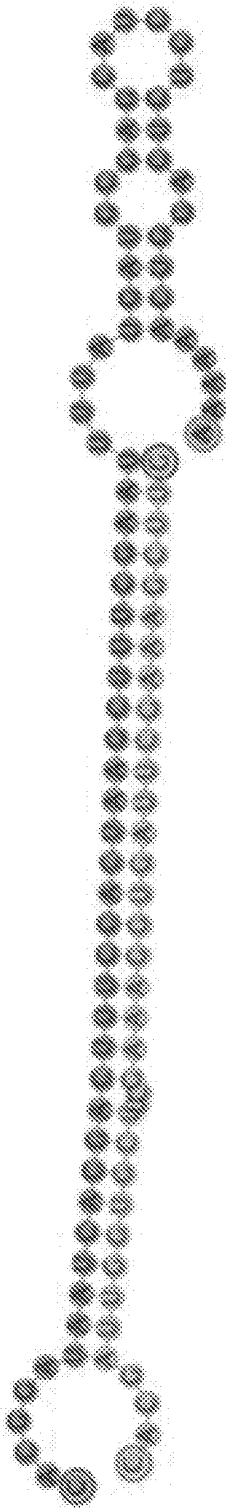
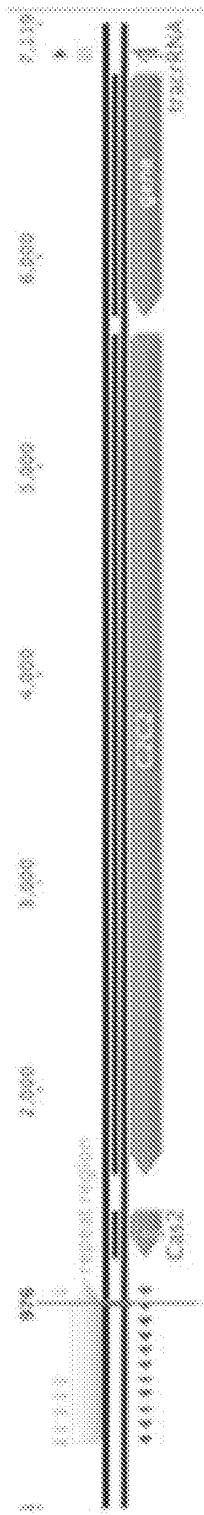
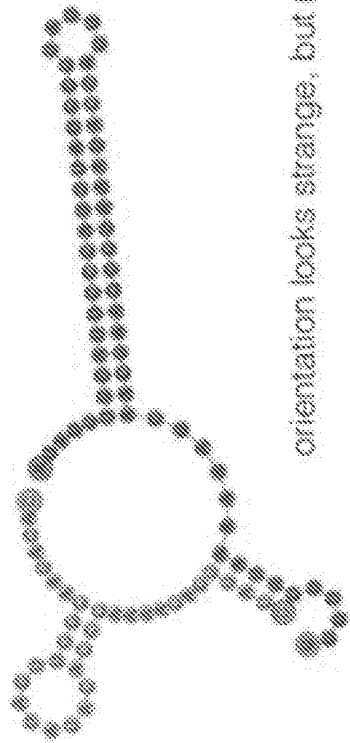


FIG. 19



DR: GUAUUGAGAAAAGCCAGAUUAUAGUUGGCAAUAGAC
tracrRNA:
AUAUUUGAUUCCCAUUUAUGGUUAUUUACCAUAAAUGGAAUCAAACUAAAAAUUU
UUUU



orientation looks strange, but maybe system is different from type II

FIG. 20

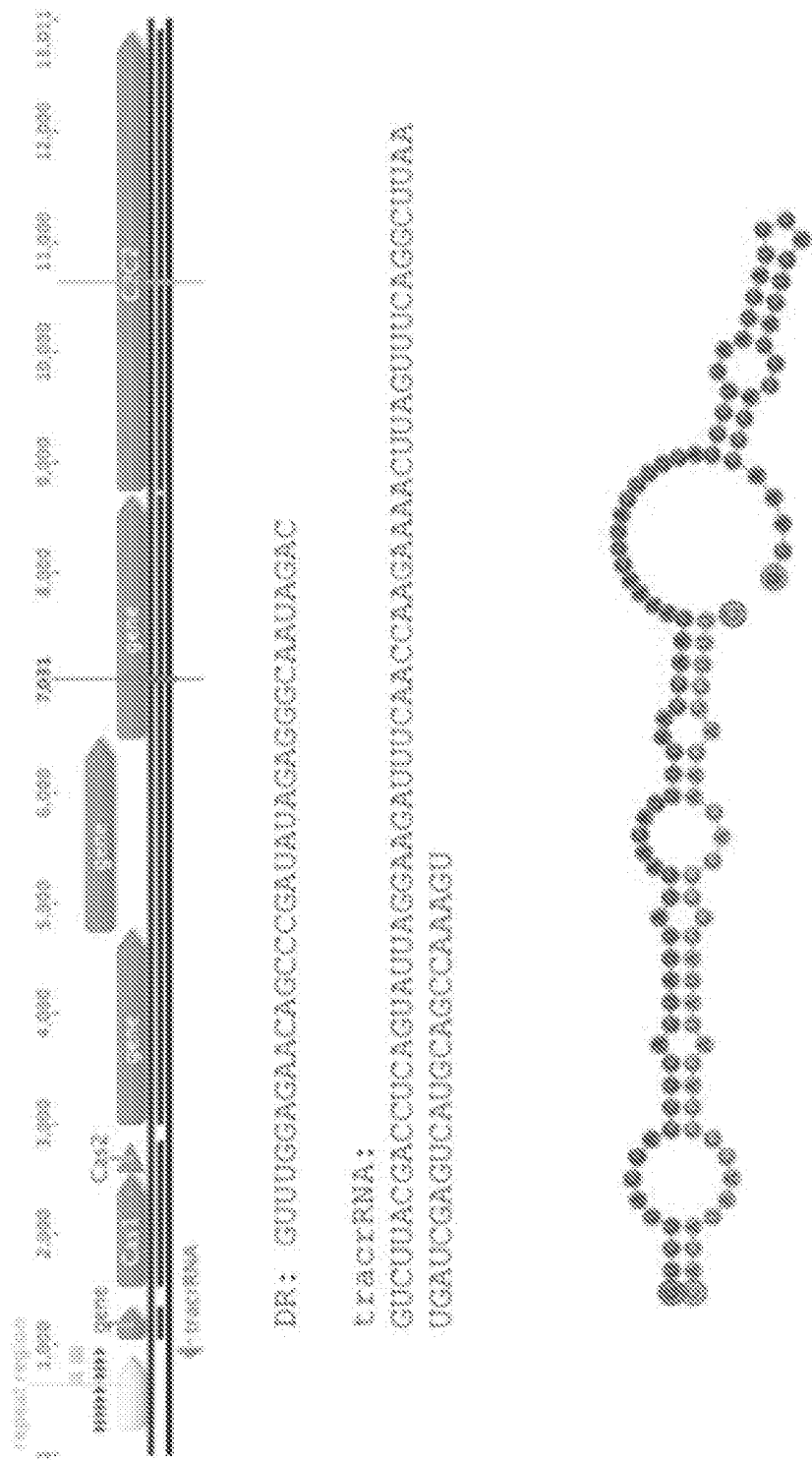


FIG. 22

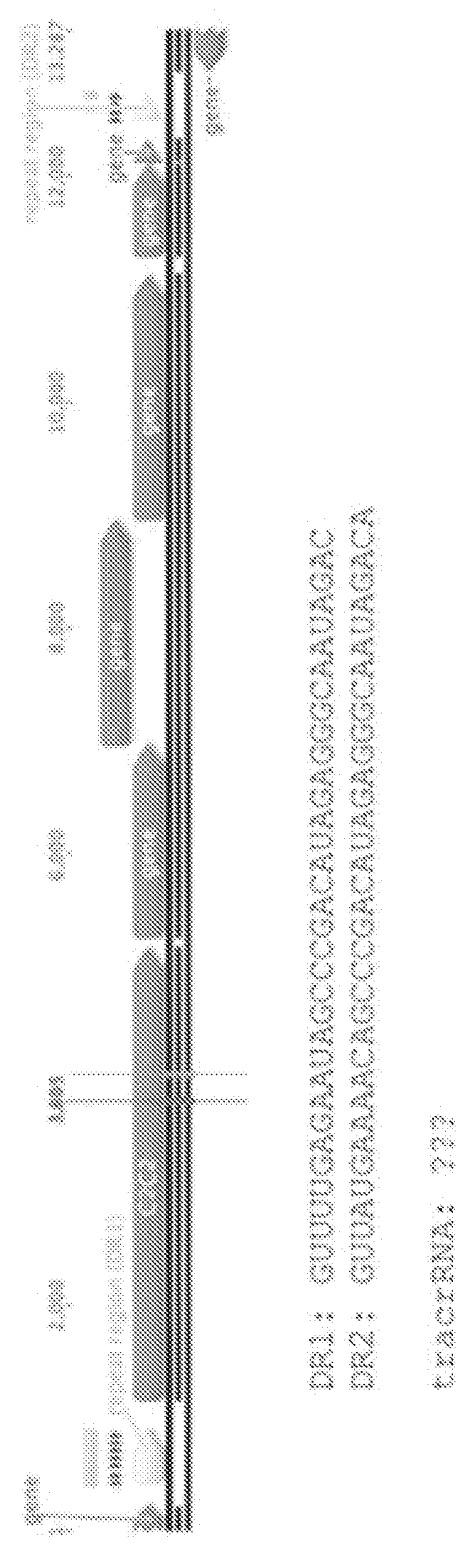
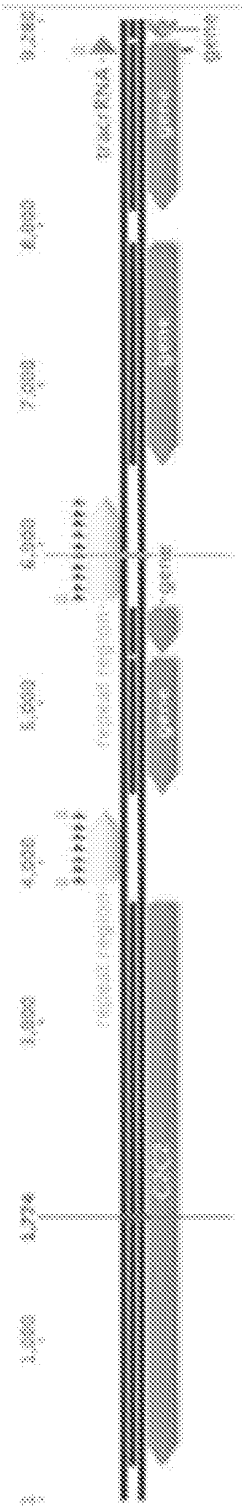


FIG. 23



DR: GUUAAGUCCUCUUAACAUUUAGAGGUAGUCUUUAAU

tracrRNA:

UCUUAAGAACUUCUCUACCUAGAGUUGGAUUUAAUAAUGACUCUUGCCUCUCAUAGAUAU
CCUCCUUUGAAAUUAUACACUGCCGAUUAUUUACCGUUUU

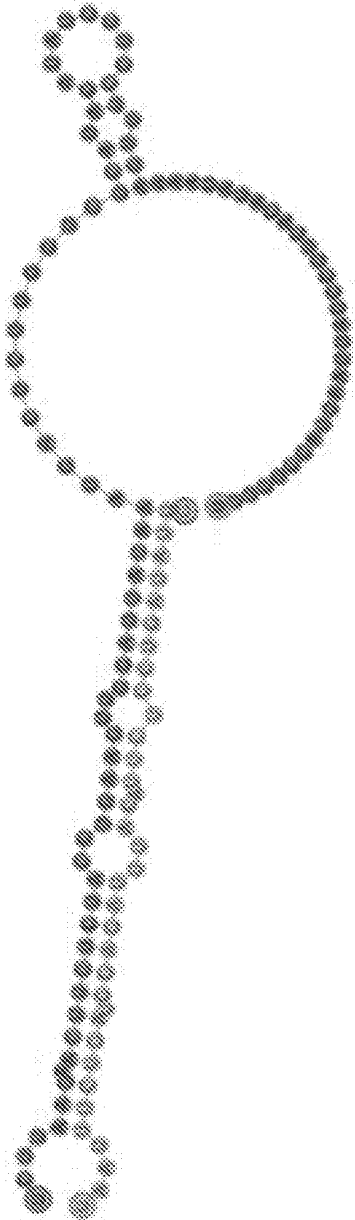
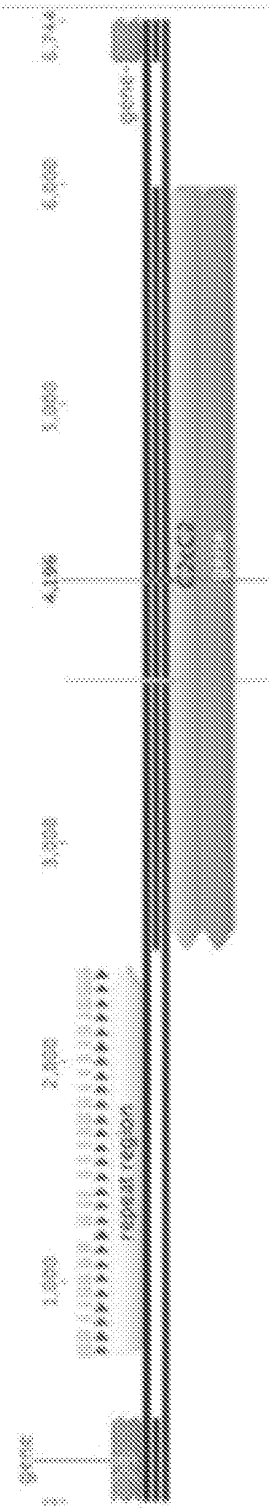


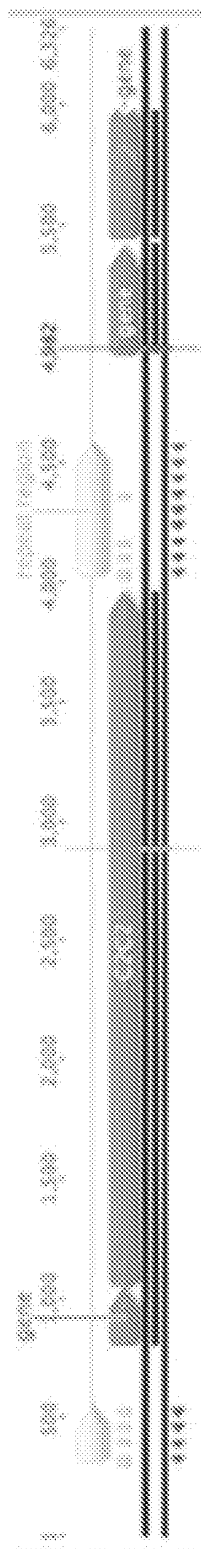
FIG. 24



DR: GUUGAGUUGCCUUCAAUUUUUGGGAUAAUCCACAAG

crRNA: ???

FIG. 26



DR: GUUUAGACCUCUUCUAUUUUUGAGGUACUCUAAAC

tracrRNA: ???

FIG. 28

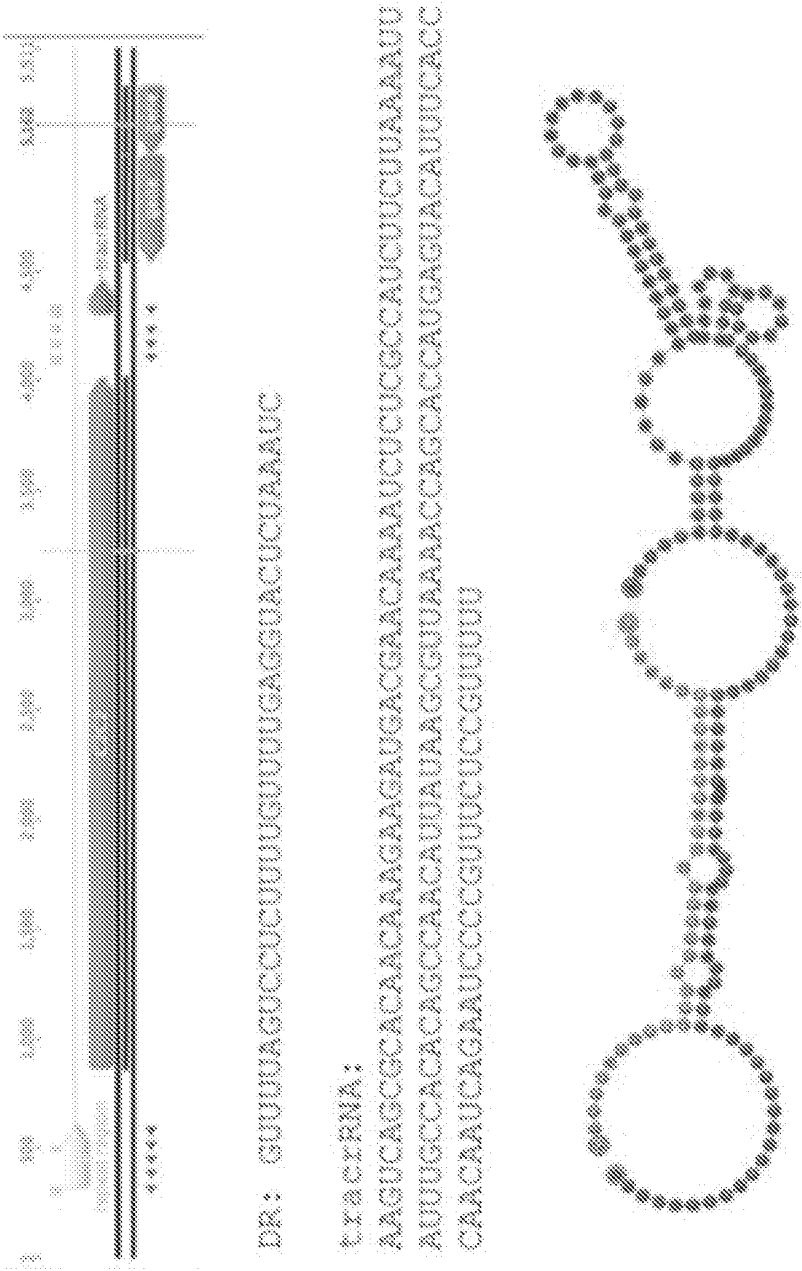
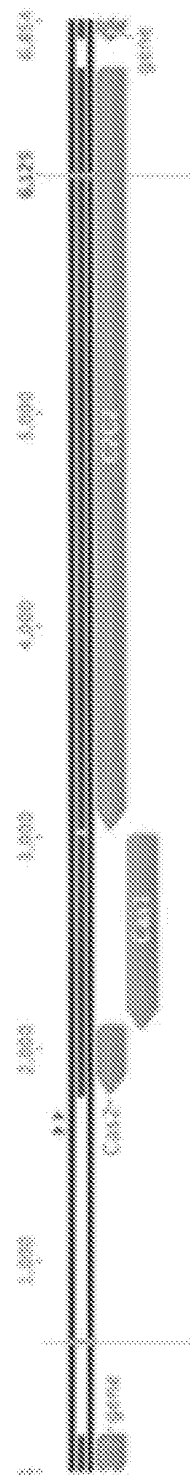


FIG. 29



DR.: GUUUAGAUCUUCGUUUUGGGUUAUCUAUAC

2003-2004

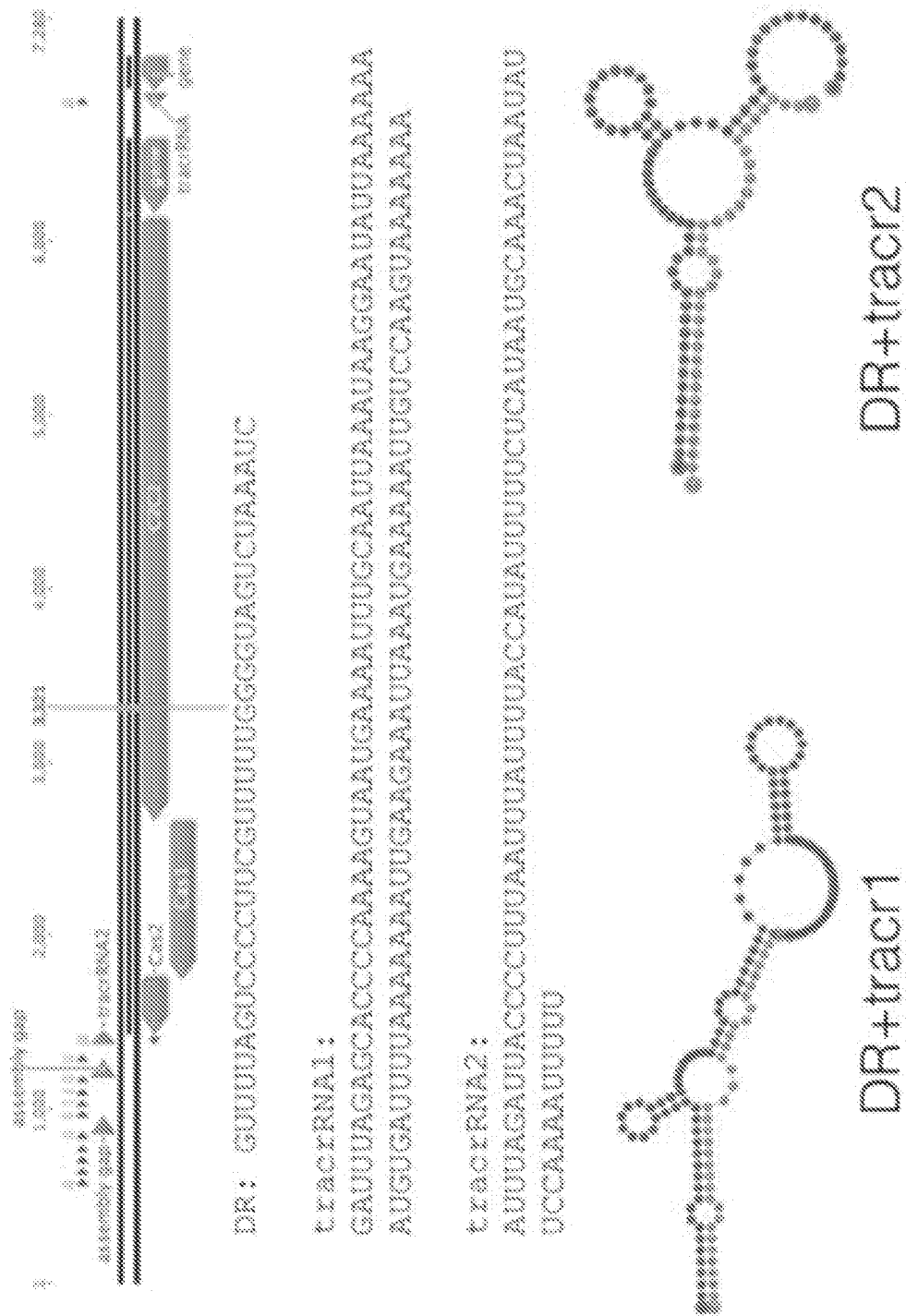


FIG. 31

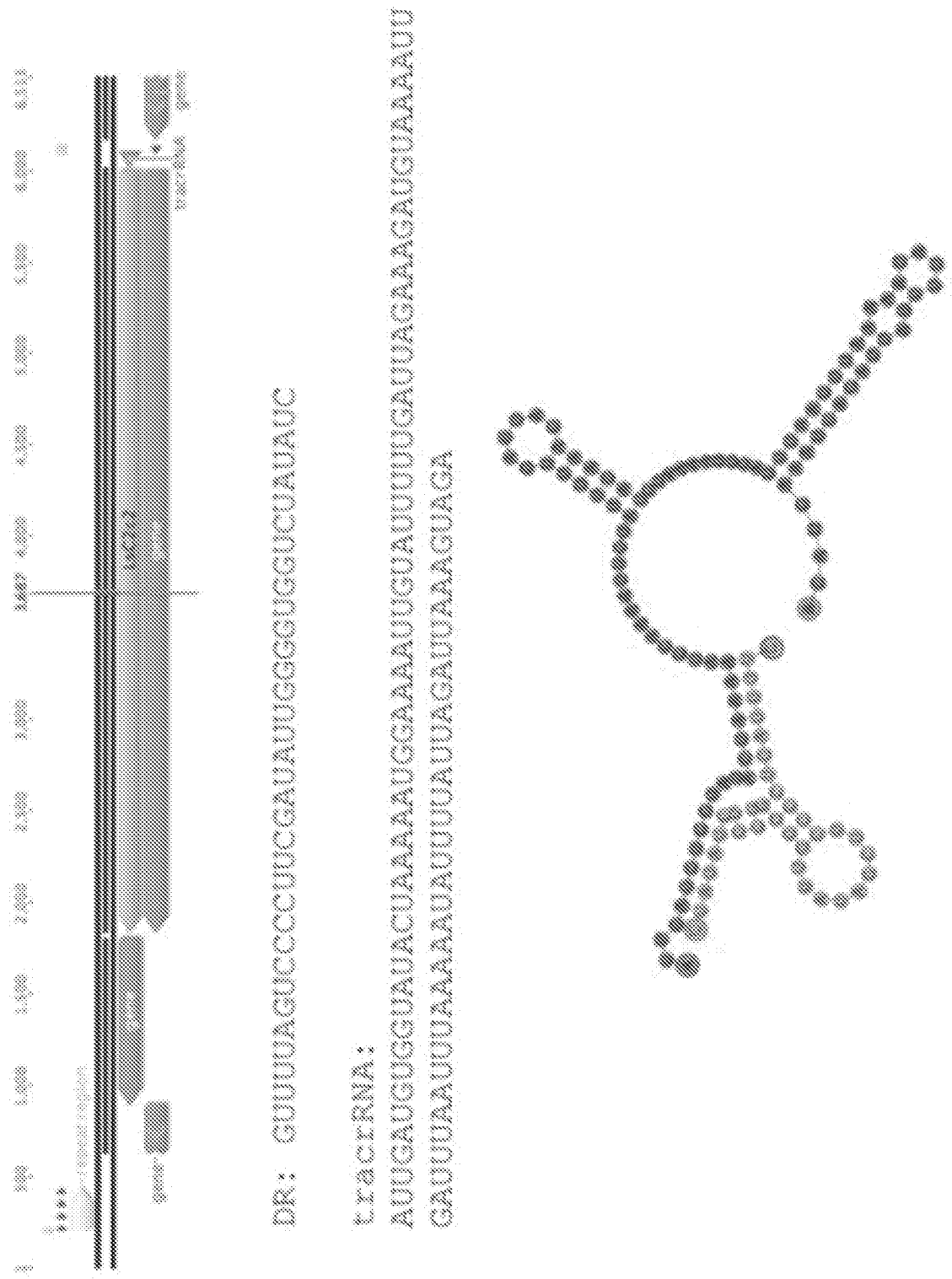
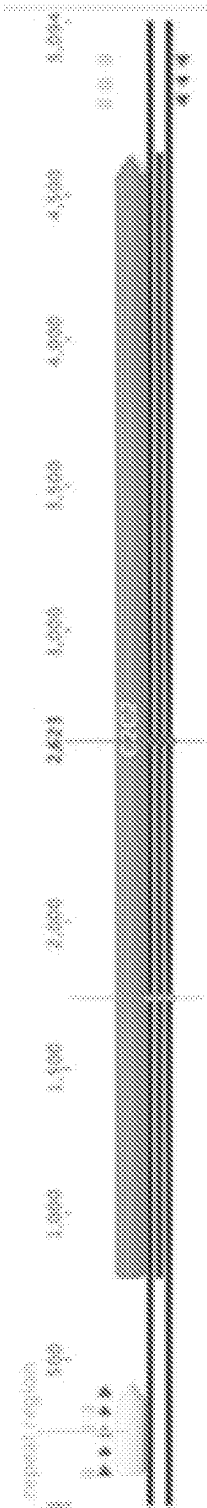


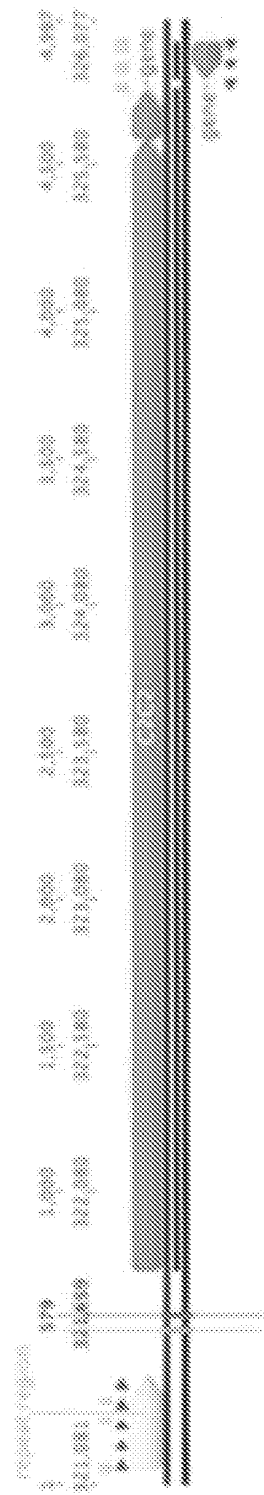
FIG. 32



DR1: GGUACAGUCCGCCGUCGUCUUGGCCGGUGAUGAGGC
DR2: GGUACAGUCCGCCGUCAUUUUGGCCGGUGAUGGCUCC

tracrRNA: ???

FIG. 33



DR1: GUUCAGUCCGCCGUCGUCUUGCGCGGUGAUGUGAGGC
DR2: GUUCAGUCCGCCGUCAUUUUGCGCGGUGAUGUGCUCC
tracrRNA: ???

FIG. 34

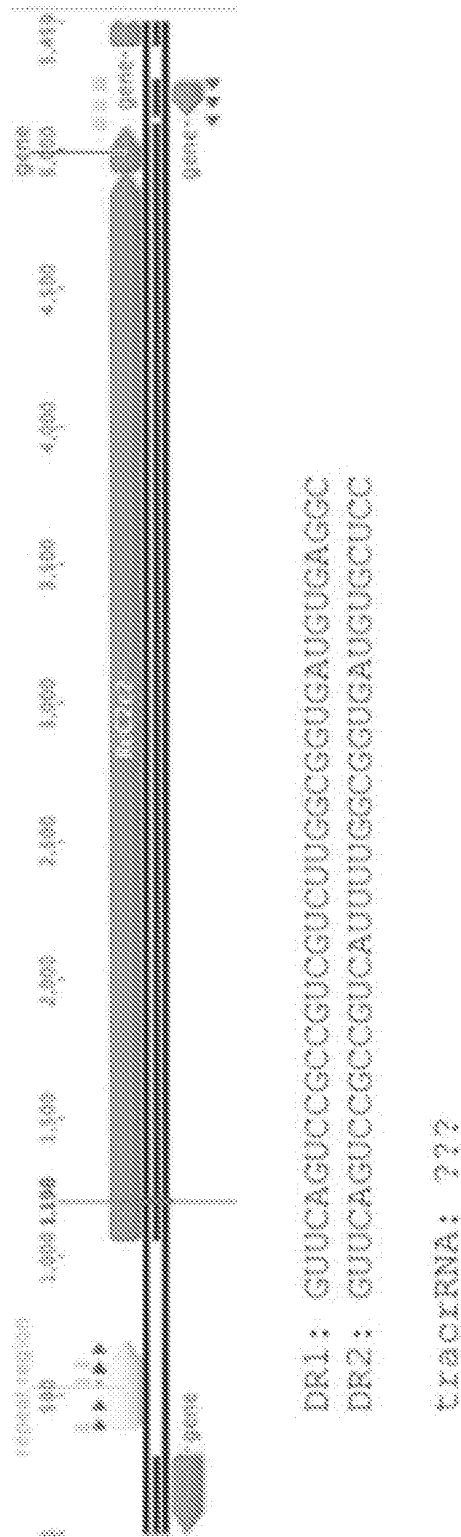


FIG. 35

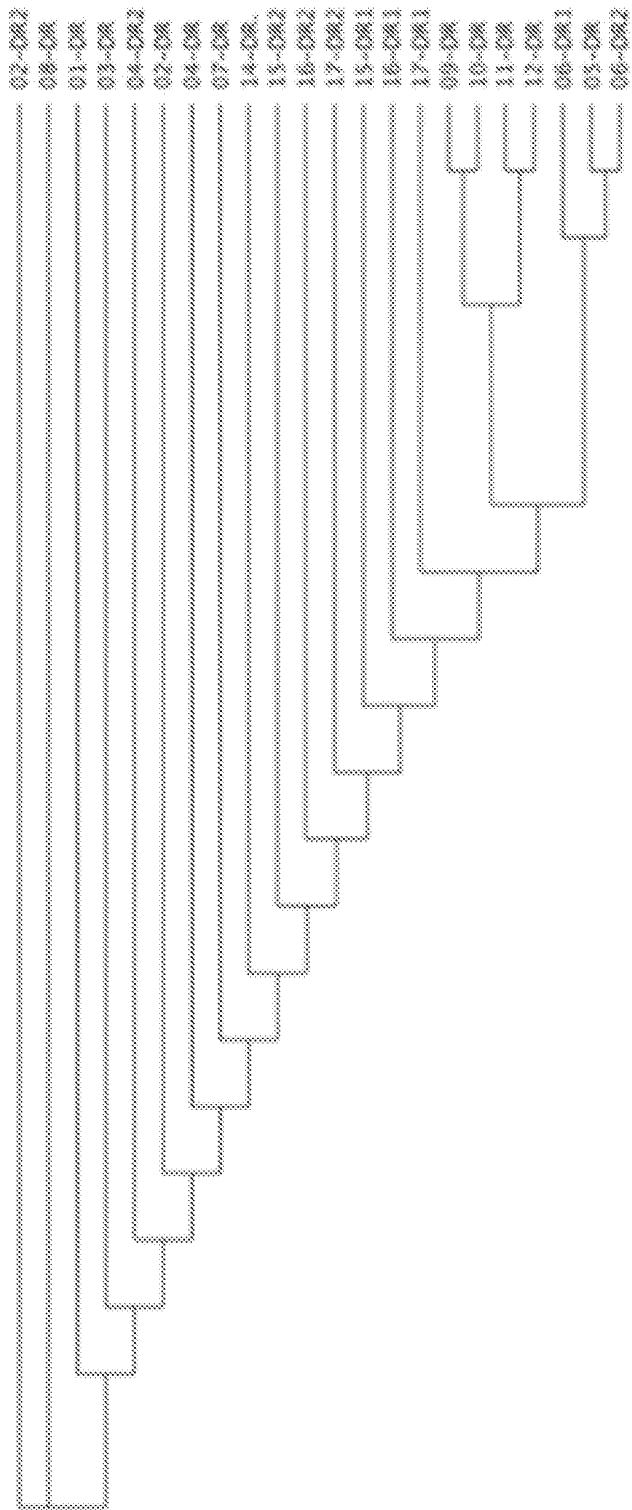


FIG. 36

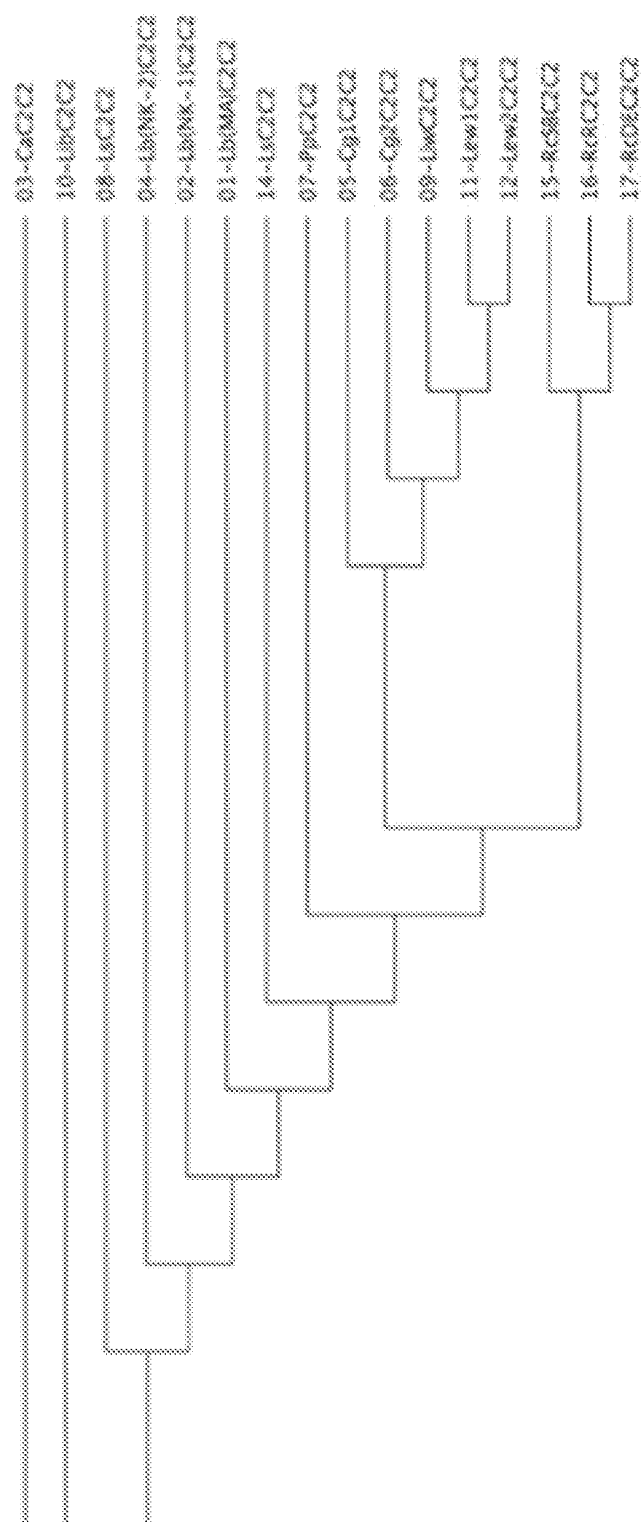


FIG. 37

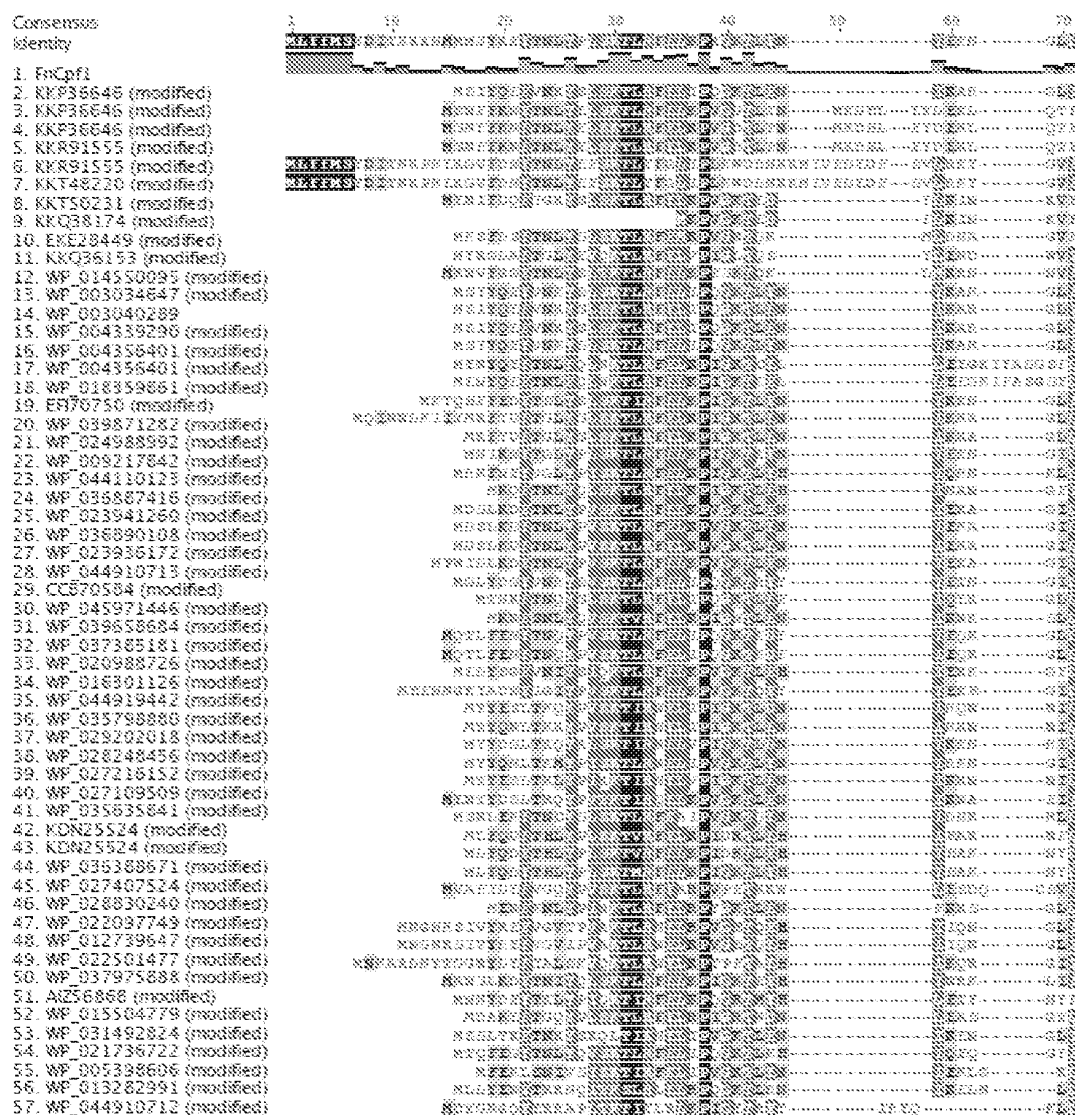


FIG. 38A

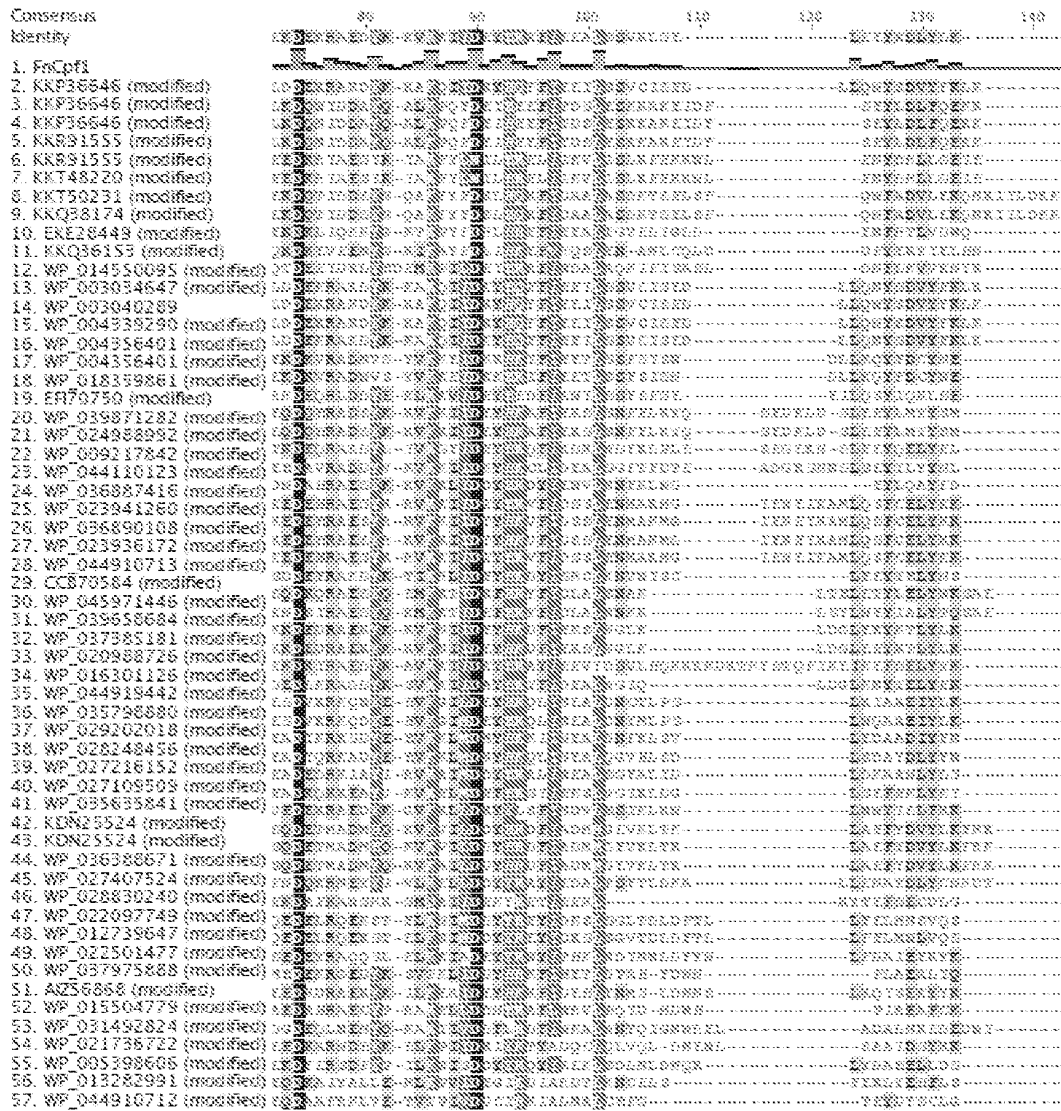


FIG. 38B

Consensus	150	160	170	180	190	200	210
Identity	-----	-----	-----	-----	-----	-----	-----
1. PnQp1	-----	-----	-----	-----	-----	-----	-----
2. KKP35646 (modified)	-----	-----	-----	-----	-----	-----	-----
3. KKP35646 (modified)	-----	-----	-----	-----	-----	-----	-----
4. KKP35646 (modified)	-----	-----	-----	-----	-----	-----	-----
5. KKR91555 (modified)	-----	-----	-----	-----	-----	-----	-----
6. KKR91555 (modified)	-----	-----	-----	-----	-----	-----	-----
7. KKT48220 (modified)	-----	-----	-----	-----	-----	-----	-----
8. KKT50231 (modified)	-----	-----	-----	-----	-----	-----	-----
9. KKT50231 (modified)	-----	-----	-----	-----	-----	-----	-----
10. KKT50231 (modified)	-----	-----	-----	-----	-----	-----	-----
11. KKT50231 (modified)	-----	-----	-----	-----	-----	-----	-----
12. WP_014550095 (modified)	-----	-----	-----	-----	-----	-----	-----
13. WP_003034647 (modified)	-----	-----	-----	-----	-----	-----	-----
14. WP_003034647 (modified)	-----	-----	-----	-----	-----	-----	-----
15. WP_003034647 (modified)	-----	-----	-----	-----	-----	-----	-----
16. WP_003034647 (modified)	-----	-----	-----	-----	-----	-----	-----
17. WP_003034647 (modified)	-----	-----	-----	-----	-----	-----	-----
18. WP_018359561 (modified)	-----	-----	-----	-----	-----	-----	-----
19. EPI70750 (modified)	-----	-----	-----	-----	-----	-----	-----
20. WP_039871282 (modified)	-----	-----	-----	-----	-----	-----	-----
21. WP_024988992 (modified)	-----	-----	-----	-----	-----	-----	-----
22. WP_008217542 (modified)	-----	-----	-----	-----	-----	-----	-----
23. WP_044110123 (modified)	-----	-----	-----	-----	-----	-----	-----
24. WP_036887416 (modified)	-----	-----	-----	-----	-----	-----	-----
25. WP_023941260 (modified)	-----	-----	-----	-----	-----	-----	-----
26. WP_036889108 (modified)	-----	-----	-----	-----	-----	-----	-----
27. WP_023936172 (modified)	-----	-----	-----	-----	-----	-----	-----
28. WP_044910713 (modified)	-----	-----	-----	-----	-----	-----	-----
29. CC70584 (modified)	-----	-----	-----	-----	-----	-----	-----
30. WP_045971446 (modified)	-----	-----	-----	-----	-----	-----	-----
31. WP_039858684 (modified)	-----	-----	-----	-----	-----	-----	-----
32. WP_037385181 (modified)	-----	-----	-----	-----	-----	-----	-----
33. WP_020988726 (modified)	-----	-----	-----	-----	-----	-----	-----
34. WP_016901126 (modified)	-----	-----	-----	-----	-----	-----	-----
35. WP_044919442 (modified)	-----	-----	-----	-----	-----	-----	-----
36. WP_035798880 (modified)	-----	-----	-----	-----	-----	-----	-----
37. WP_029202018 (modified)	-----	-----	-----	-----	-----	-----	-----
38. WP_026248456 (modified)	-----	-----	-----	-----	-----	-----	-----
39. WP_027216152 (modified)	-----	-----	-----	-----	-----	-----	-----
40. WP_027109509 (modified)	-----	-----	-----	-----	-----	-----	-----
41. WP_035635841 (modified)	-----	-----	-----	-----	-----	-----	-----
42. KDN25524 (modified)	-----	-----	-----	-----	-----	-----	-----
43. KDN25524 (modified)	-----	-----	-----	-----	-----	-----	-----
44. WP_035388671 (modified)	-----	-----	-----	-----	-----	-----	-----
45. WP_027407524 (modified)	-----	-----	-----	-----	-----	-----	-----
46. WP_026830240 (modified)	-----	-----	-----	-----	-----	-----	-----
47. WP_022097749 (modified)	-----	-----	-----	-----	-----	-----	-----
48. WP_012759647 (modified)	-----	-----	-----	-----	-----	-----	-----
49. WP_022501477 (modified)	-----	-----	-----	-----	-----	-----	-----
50. WP_037975888 (modified)	-----	-----	-----	-----	-----	-----	-----
51. AIZ56866 (modified)	-----	-----	-----	-----	-----	-----	-----
52. WP_015504779 (modified)	-----	-----	-----	-----	-----	-----	-----
53. WP_031492824 (modified)	-----	-----	-----	-----	-----	-----	-----
54. WP_021736722 (modified)	-----	-----	-----	-----	-----	-----	-----
55. WP_005398606 (modified)	-----	-----	-----	-----	-----	-----	-----
56. WP_013262991 (modified)	-----	-----	-----	-----	-----	-----	-----
57. WP_044910712 (modified)	-----	-----	-----	-----	-----	-----	-----

FIG. 38C

Consensus	478	488	498	508	518	528
Identity						
1. PnCP1						
2. KKP36646 (modified)						
3. KKP36646 (modified)						
4. KKP36646 (modified)						
5. KKP36646 (modified)						
6. KKP36646 (modified)						
7. KKT48220 (modified)						
8. KKT50231 (modified)						
9. KKT50231 (modified)						
10. KKT50231 (modified)						
11. KKT50231 (modified)						
12. WP_014550095 (modified)						
13. WP_003034647 (modified)						
14. WP_003040289						
15. WP_004339290 (modified)						
16. WP_004339290 (modified)						
17. WP_004339290 (modified)						
18. WP_018359861 (modified)						
19. ERT07050 (modified)						
20. WP_009871282 (modified)						
21. WP_024988992 (modified)						
22. WP_009217842 (modified)						
23. WP_044110123 (modified)						
24. WP_00887416 (modified)						
25. WP_023941260 (modified)						
26. WP_006890108 (modified)						
27. WP_023936172 (modified)						
28. WP_044910713 (modified)						
29. CC870584 (modified)						
30. WP_045971448 (modified)						
31. WP_009858884 (modified)						
32. WP_007385181 (modified)						
33. WP_020988726 (modified)						
34. WP_018901126 (modified)						
35. WP_044919442 (modified)						
36. WP_005758880 (modified)						
37. WP_029202018 (modified)						
38. WP_028248456 (modified)						
39. WP_027216152 (modified)						
40. WP_027109509 (modified)						
41. WP_035635841 (modified)						
42. KDN25524 (modified)						
43. KDN25524 (modified)						
44. WP_006368671 (modified)						
45. WP_027407524 (modified)						
46. WP_028830240 (modified)						
47. WP_022097749 (modified)						
48. WP_012739647 (modified)						
49. WP_022501477 (modified)						
50. WP_037975880 (modified)						
51. AIZ58868 (modified)						
52. WP_015504779 (modified)						
53. WP_031492824 (modified)						
54. WP_021736722 (modified)						
55. WP_005398506 (modified)						
56. WP_013282991 (modified)						
57. WP_044910712 (modified)						

FIG. 38D

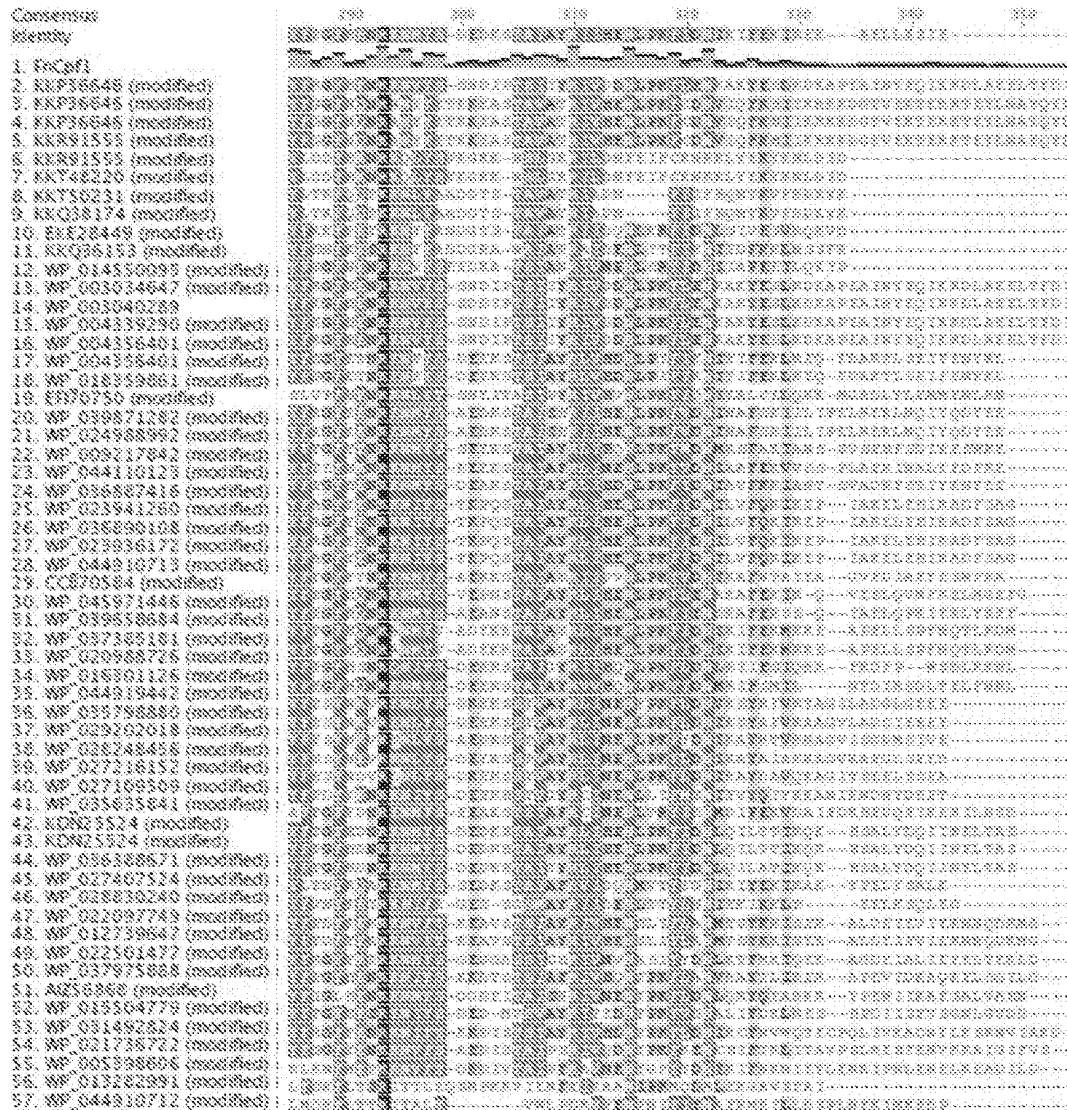


FIG. 38E

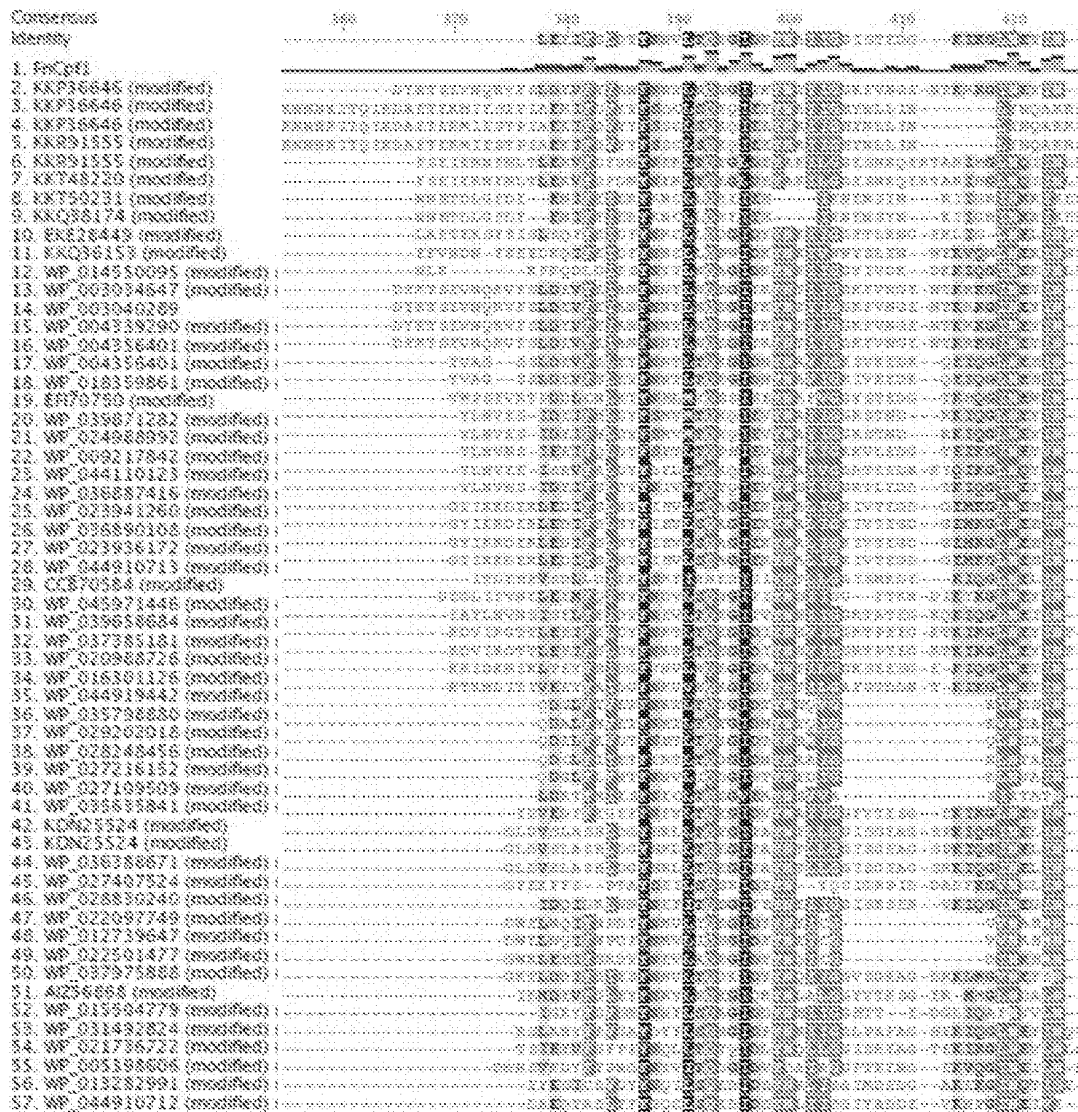


FIG. 38F

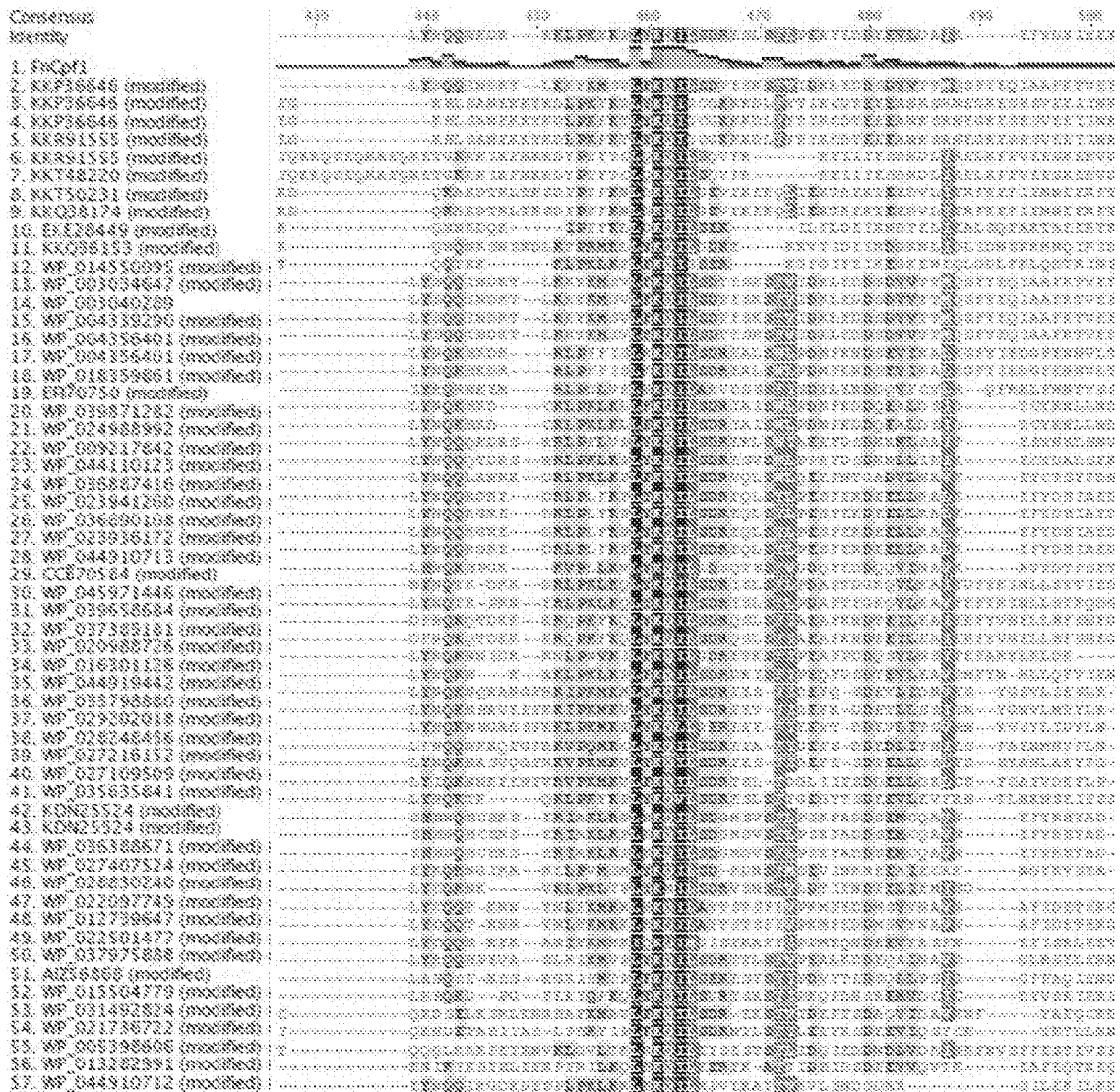


FIG. 38G

Consensus	100	100	100	100	100	100
Identity	100	100	100	100	100	100
1. EnCpf1						
2. KKP36646 (modified)						
3. KKP36646 (modified)						
4. KKP36646 (modified)						
5. KKP36646 (modified)						
6. KKP36646 (modified)						
7. KKP36646 (modified)						
8. KKP36646 (modified)						
9. KKP36646 (modified)						
10. KKP36646 (modified)						
11. KKP36646 (modified)						
12. WP_014550095 (modified)						
13. WP_003034847 (modified)						
14. WP_003040289 (modified)						
15. WP_004339290 (modified)						
16. WP_004339401 (modified)						
17. WP_004339401 (modified)						
18. WP_018359861 (modified)						
19. ERI70790 (modified)						
20. WP_035871282 (modified)						
21. WP_034808892 (modified)						
22. WP_008217843 (modified)						
23. WP_044130123 (modified)						
24. WP_036887418 (modified)						
25. WP_03941260 (modified)						
26. WP_036887418 (modified)						
27. WP_021936172 (modified)						
28. WP_044910713 (modified)						
29. CC870384 (modified)						
30. WP_045971446 (modified)						
31. WP_039656884 (modified)						
32. WP_037385181 (modified)						
33. WP_020988726 (modified)						
34. WP_018301126 (modified)						
35. WP_044919442 (modified)						
36. WP_035791888 (modified)						
37. WP_009203916 (modified)						
38. WP_028248416 (modified)						
39. WP_027216152 (modified)						
40. WP_027109509 (modified)						
41. WP_035635841 (modified)						
42. KDN25524 (modified)						
43. KDN25524 (modified)						
44. WP_036388671 (modified)						
45. WP_027407524 (modified)						
46. WP_028830240 (modified)						
47. WP_022097749 (modified)						
48. WP_012739647 (modified)						
49. WP_022501477 (modified)						
50. WP_037975888 (modified)						
51. AG26868 (modified)						
52. WP_015504779 (modified)						
53. WP_031492824 (modified)						
54. WP_021736722 (modified)						
55. WP_005396806 (modified)						
56. WP_013282993 (modified)						
57. WP_044910712 (modified)						

FIG. 38H

[The page contains faint, illegible markings, possibly bleed-through from the reverse side.]

Consensus Identity	540	550	560	570	580	590	600	610	790	795
1. FcCp1										
2. KKF36646 (modified)										
3. KKF36646 (modified)										
4. KKF36646 (modified)										
5. KKR31555 (modified)										
6. KKR31555 (modified)										
7. KKT48220 (modified)										
8. KKT50231 (modified)										
9. KKT58174 (modified)										
10. KKE26449 (modified)										
11. KKT38153 (modified)										
12. WP_014850695 (modified)										
13. WP_003034647 (modified)										
14. WP_003040289 (modified)										
15. WP_004339390 (modified)										
16. WP_004336401 (modified)										
17. WP_004336401 (modified)										
18. WP_018330861 (modified)										
19. EPI70730 (modified)										
20. WP_019871282 (modified)										
21. WP_024988992 (modified)										
22. WP_009217842 (modified)										
23. WP_044110125 (modified)										
24. WP_03687416 (modified)										
25. WP_023941260 (modified)										
26. WP_036890108 (modified)										
27. WP_023936172 (modified)										
28. WP_044910711 (modified)										
29. CKE70564 (modified)										
30. WP_045971446 (modified)										
31. WP_039610684 (modified)										
32. WP_037361181 (modified)										
33. WP_020988726 (modified)										
34. WP_016301126 (modified)										
35. WP_044919442 (modified)										
36. WP_035798880 (modified)										
37. WP_029262018 (modified)										
38. WP_028248456 (modified)										
39. WP_027216152 (modified)										
40. WP_027109509 (modified)										
41. WP_015631841 (modified)										
42. KDN25524 (modified)										
43. KDN25524 (modified)										
44. WP_038368671 (modified)										
45. WP_027407524 (modified)										
46. WP_026830340 (modified)										
47. WP_022097749 (modified)										
48. WP_012710847 (modified)										
49. WP_022501477 (modified)										
50. WP_017975888 (modified)										
51. A256886 (modified)										
52. WP_015504779 (modified)										
53. WP_011492624 (modified)										
54. WP_021736722 (modified)										
55. WP_005398608 (modified)										
56. WP_013282391 (modified)										
57. WP_044910712 (modified)										

FIG. 38J

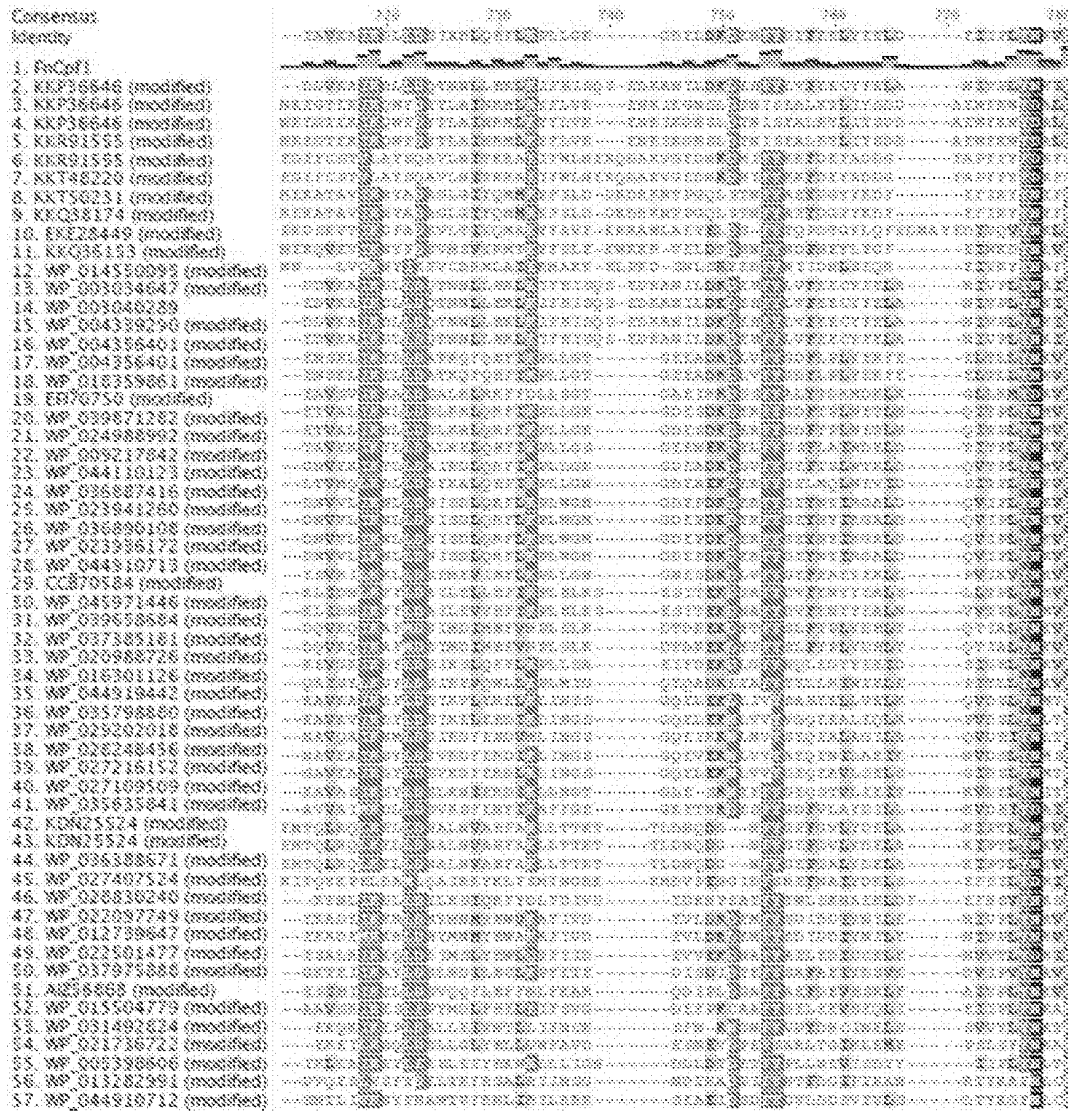


FIG. 38K

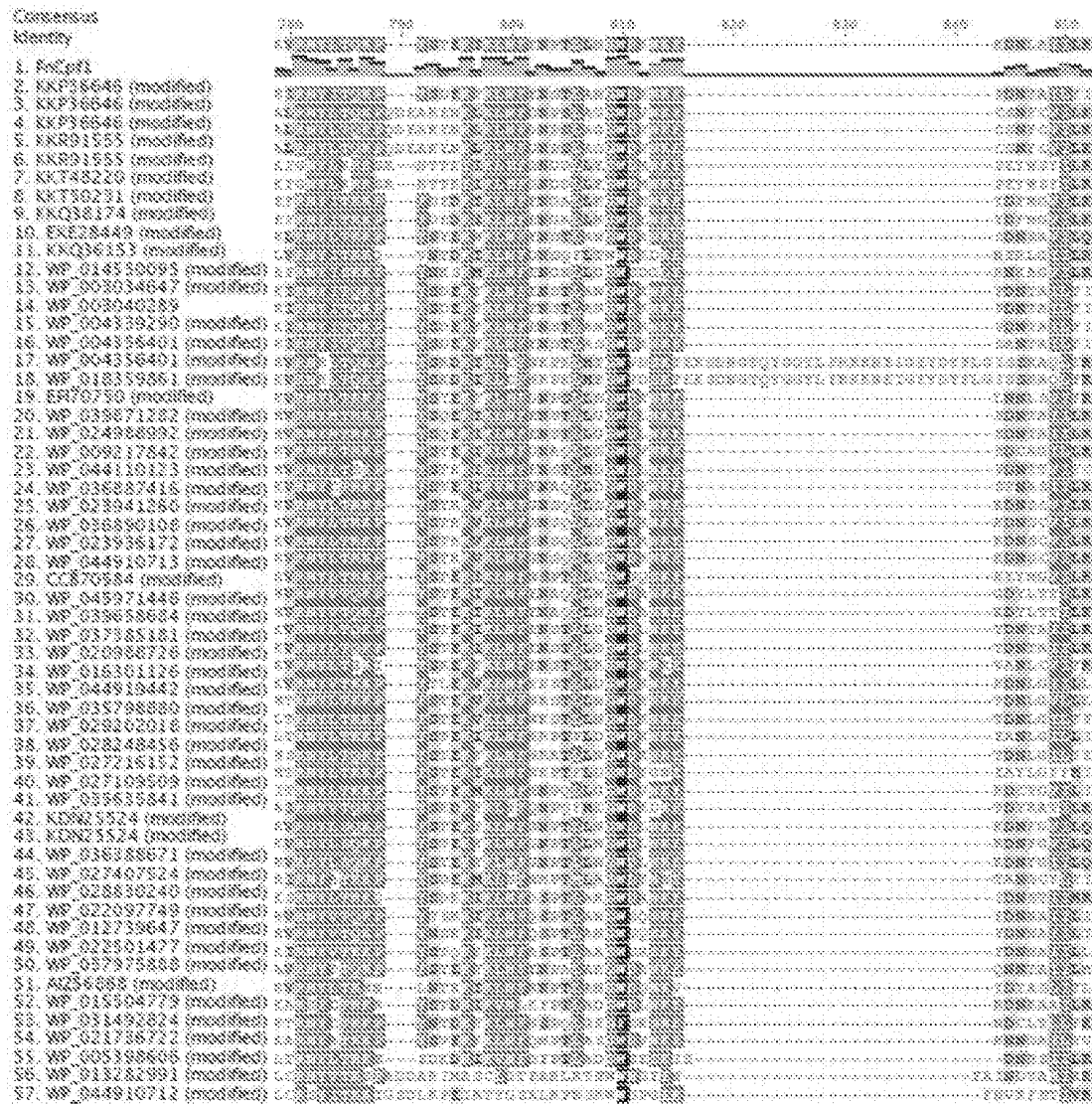


FIG. 38L

Consensus Identity	990	970	950	930	910	890
1. EnCaf1						
2. KKP36646 (modified)						
3. KKP36646 (modified)						
4. KKP36646 (modified)						
5. KKP31555 (modified)						
6. KKP31555 (modified)						
7. KKT48220 (modified)						
8. KKT50231 (modified)						
9. KKT50231 (modified)						
10. KKT50231 (modified)						
11. KKT50231 (modified)						
12. WP_014550095 (modified)						
13. WP_003034647 (modified)						
14. WP_003040289 (modified)						
15. WP_004339290 (modified)						
16. WP_004339401 (modified)						
17. WP_004339401 (modified)						
18. WP_018359861 (modified)						
19. EHT0750 (modified)						
20. WP_035871282 (modified)						
21. WP_024988992 (modified)						
22. WP_009217842 (modified)						
23. WP_044130123 (modified)						
24. WP_036887418 (modified)						
25. WP_023941280 (modified)						
26. WP_036890108 (modified)						
27. WP_023936172 (modified)						
28. WP_044910713 (modified)						
29. CC370884 (modified)						
30. WP_045971446 (modified)						
31. WP_035658684 (modified)						
32. WP_037385181 (modified)						
33. WP_020986726 (modified)						
34. WP_016301126 (modified)						
35. WP_044919442 (modified)						
36. WP_035798888 (modified)						
37. WP_029202016 (modified)						
38. WP_028248456 (modified)						
39. WP_027218152 (modified)						
40. WP_027109509 (modified)						
41. WP_035635841 (modified)						
42. KDN23524 (modified)						
43. KDN23524 (modified)						
44. WP_036386671 (modified)						
45. WP_027407524 (modified)						
46. WP_028830240 (modified)						
47. WP_022097749 (modified)						
48. WP_012739647 (modified)						
49. WP_022501477 (modified)						
50. WP_037975888 (modified)						
51. A256866 (modified)						
52. WP_015504779 (modified)						
53. WP_031492824 (modified)						
54. WP_021756722 (modified)						
55. WP_003398606 (modified)						
56. WP_013282991 (modified)						
57. WP_044910712 (modified)						

FIG. 38M

Consensus	620	670	720	770	820	870	920
Identity							
1. FcCnf1							
2. KKP36646 (modified)							
3. KKP36646 (modified)							
4. KKP36646 (modified)							
5. KKP36646 (modified)							
6. KKP36646 (modified)							
7. KKT48220 (modified)							
8. KKT50231 (modified)							
9. KKT50231 (modified)							
10. KKT50231 (modified)							
11. KKT50231 (modified)							
12. WP_014550695 (modified)							
13. WP_063034647 (modified)							
14. WP_063040269 (modified)							
15. WP_064339295 (modified)							
16. WP_064356481 (modified)							
17. WP_064356481 (modified)							
18. WP_018359861 (modified)							
19. ERI70750 (modified)							
20. WP_039871282 (modified)							
21. WP_014868892 (modified)							
22. WP_069217843 (modified)							
23. WP_044110123 (modified)							
24. WP_036887416 (modified)							
25. WP_023941260 (modified)							
26. WP_038890106 (modified)							
27. WP_021936172 (modified)							
28. WP_044910713 (modified)							
29. CCE70584 (modified)							
30. WP_045371446 (modified)							
31. WP_039858684 (modified)							
32. WP_037385181 (modified)							
33. WP_020988726 (modified)							
34. WP_018301176 (modified)							
35. WP_044919442 (modified)							
36. WP_031796880 (modified)							
37. WP_029202016 (modified)							
38. WP_028248416 (modified)							
39. WP_027216152 (modified)							
40. WP_027109509 (modified)							
41. WP_035635841 (modified)							
42. KDN29524 (modified)							
43. KDN25524 (modified)							
44. WP_036388671 (modified)							
45. WP_027407524 (modified)							
46. WP_028830240 (modified)							
47. WP_022087749 (modified)							
48. WP_012735847 (modified)							
49. WP_022501477 (modified)							
50. WP_037575888 (modified)							
51. A256868 (modified)							
52. WP_015504779 (modified)							
53. WP_031492824 (modified)							
54. WP_021796722 (modified)							
55. WP_005396506 (modified)							
56. WP_013262931 (modified)							
57. WP_044910712 (modified)							

FIG. 38N

Consensus Identity	1,000	1,000	1,000	1,000	1,000	1,000
1. FxCut3						
2. KKP36646 (modified)						
3. KKP36646 (modified)						
4. KKP36646 (modified)						
5. KKP36646 (modified)						
6. KKP36646 (modified)						
7. KKP36646 (modified)						
8. KKP36646 (modified)						
9. KKP36646 (modified)						
10. KKP36646 (modified)						
11. KKP36646 (modified)						
12. WP_014550095 (modified)						
13. WP_005034647 (modified)						
14. WP_005034647 (modified)						
15. WP_005034647 (modified)						
16. WP_005034647 (modified)						
17. WP_005034647 (modified)						
18. WP_018359861 (modified)						
19. EPI70750 (modified)						
20. WP_035871282 (modified)						
21. WP_014868993 (modified)						
22. WP_009217843 (modified)						
23. WP_044110153 (modified)						
24. WP_036874118 (modified)						
25. WP_023941260 (modified)						
26. WP_036874118 (modified)						
27. WP_023941260 (modified)						
28. WP_044910713 (modified)						
29. CC870584 (modified)						
30. WP_045871446 (modified)						
31. WP_036874118 (modified)						
32. WP_037785181 (modified)						
33. WP_020068726 (modified)						
34. WP_018359861 (modified)						
35. WP_044910713 (modified)						
36. WP_035798880 (modified)						
37. WP_020068726 (modified)						
38. WP_028248456 (modified)						
39. WP_027218152 (modified)						
40. WP_027105503 (modified)						
41. WP_035675841 (modified)						
42. KDN25524 (modified)						
43. KDN25524 (modified)						
44. WP_036874118 (modified)						
45. WP_027407524 (modified)						
46. WP_028850240 (modified)						
47. WP_023097749 (modified)						
48. WP_012719647 (modified)						
49. WP_022501477 (modified)						
50. WP_037785181 (modified)						
51. A236668 (modified)						
52. WP_015504779 (modified)						
53. WP_031491824 (modified)						
54. WP_021716723 (modified)						
55. WP_005398606 (modified)						
56. WP_013282981 (modified)						
57. WP_044910712 (modified)						

FIG. 380

Consensus Identity	1,046	1,076	1,099	1,090	1,100	1,118	1,139	1,159
1. ForCap1								
2. KKP36646 (modified)								
3. KKP36646 (modified)								
4. KKP36646 (modified)								
5. KKR31555 (modified)								
6. KKR31555 (modified)								
7. KKT48220 (modified)								
8. KKT50231 (modified)								
9. KKG38174 (modified)								
10. KKE26449 (modified)								
11. KKG36153 (modified)								
12. WP_014556095 (modified)								
13. WP_003034647 (modified)								
14. WP_003040289								
15. WP_004339290 (modified)								
16. WP_004356401 (modified)								
17. WP_004356401 (modified)								
18. WP_018359861 (modified)								
19. EKH70750 (modified)								
20. WP_019871282 (modified)								
21. WP_024960992 (modified)								
22. WP_003217642 (modified)								
23. WP_044150123								
24. WP_016887416 (modified)								
25. WP_023941260 (modified)								
26. WP_006699108 (modified)								
27. WP_023936172 (modified)								
28. WP_044910713 (modified)								
29. CCE70584 (modified)								
30. WP_045971446 (modified)								
31. WP_019656684 (modified)								
32. WP_037385181 (modified)								
33. WP_020968726 (modified)								
34. WP_016301126 (modified)								
35. WP_044919442 (modified)								
36. WP_015796580 (modified)								
37. WP_029202018 (modified)								
38. WP_025248456 (modified)								
39. WP_027216152 (modified)								
40. WP_027109509 (modified)								
41. WP_035635841 (modified)								
42. KDN25524 (modified)								
43. KDN25524 (modified)								
44. WP_006588671 (modified)								
45. WP_027407524 (modified)								
46. WP_028830240 (modified)								
47. WP_022097749 (modified)								
48. WP_012759647 (modified)								
49. WP_022501477 (modified)								
50. WP_037975888 (modified)								
51. AG56868 (modified)								
52. WP_013504779 (modified)								
53. WP_031492524 (modified)								
54. WP_021736722 (modified)								
55. WP_005398606 (modified)								
56. WP_013282991 (modified)								
57. WP_044910712 (modified)								

FIG. 38P

FIG. 38Q

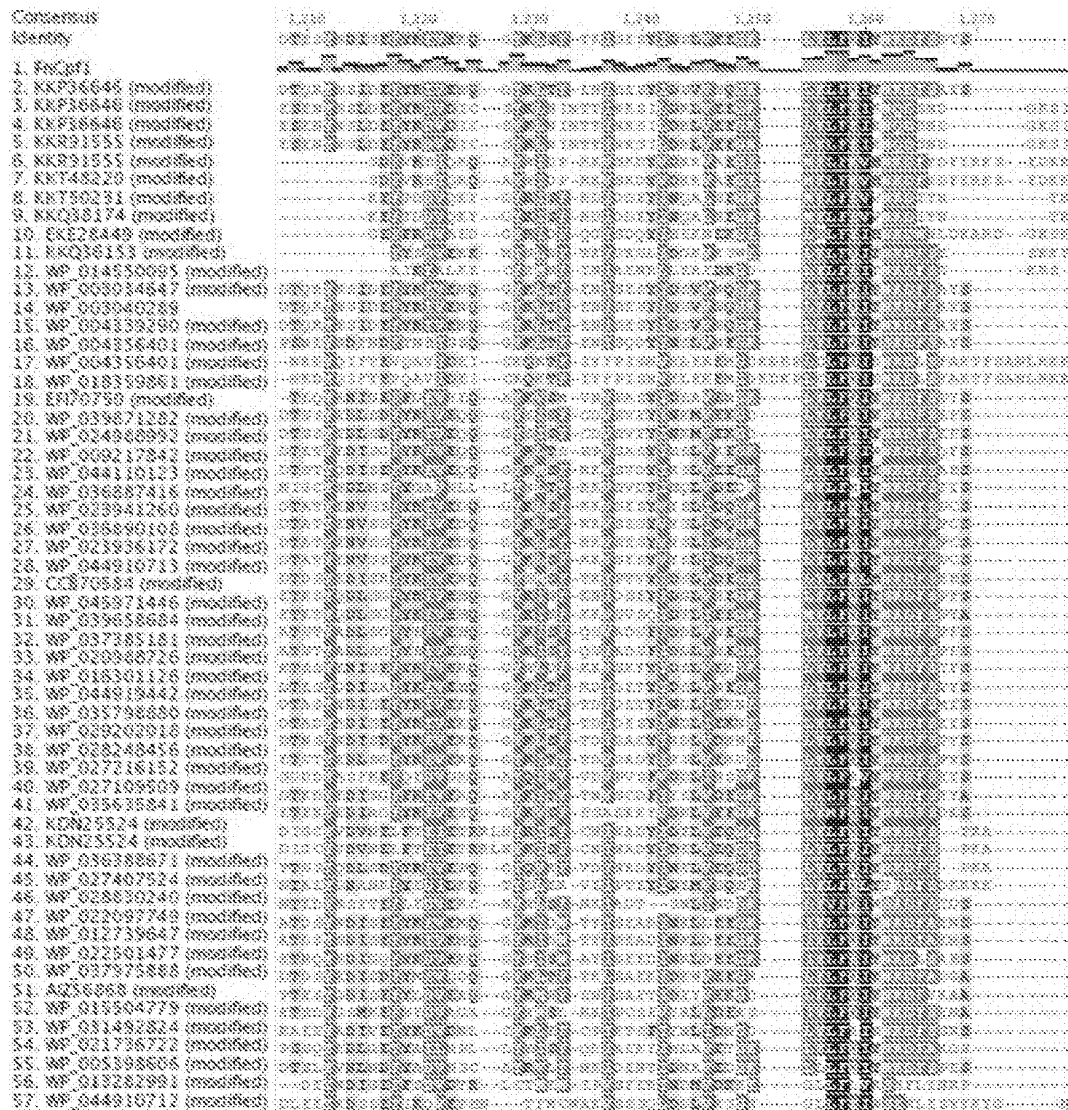


FIG. 38R

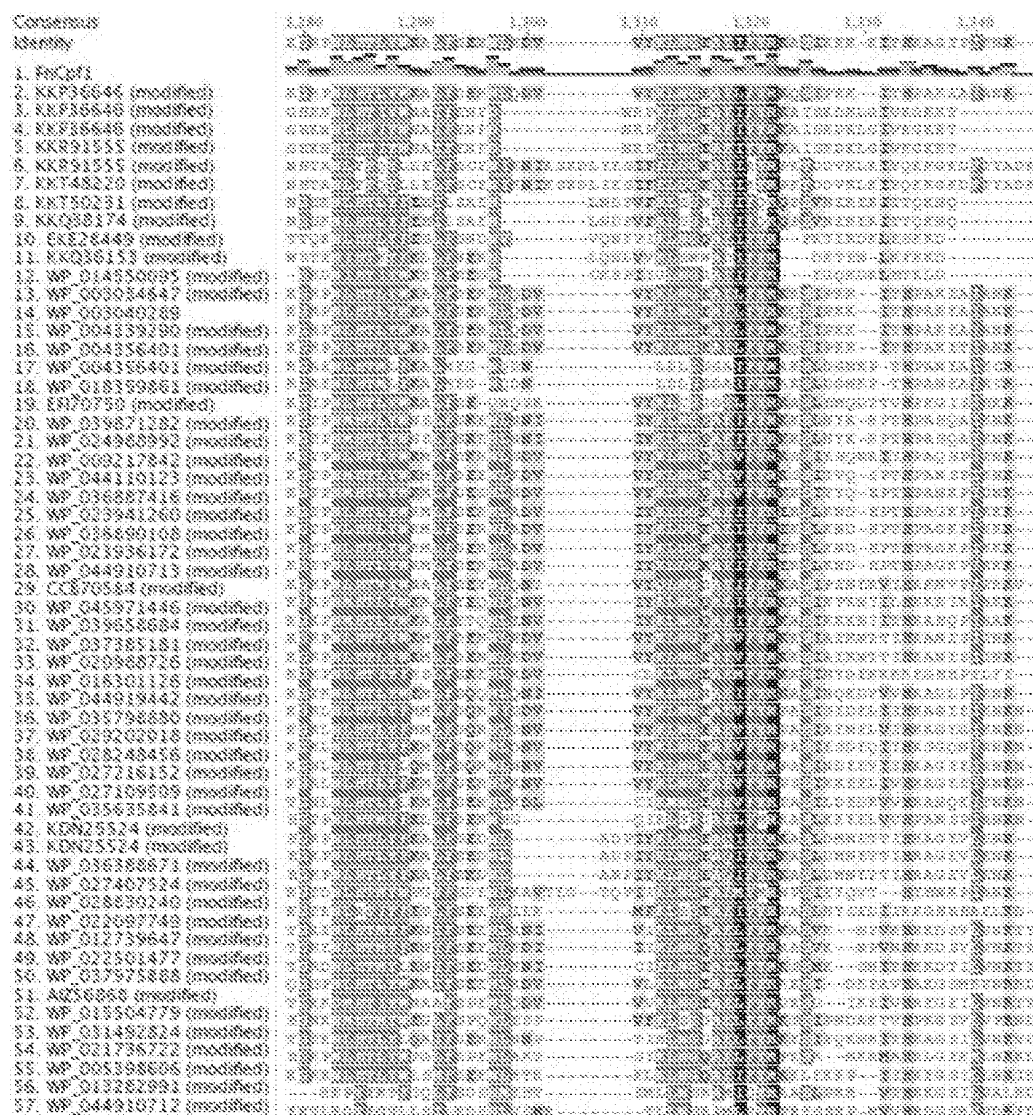


FIG. 38S

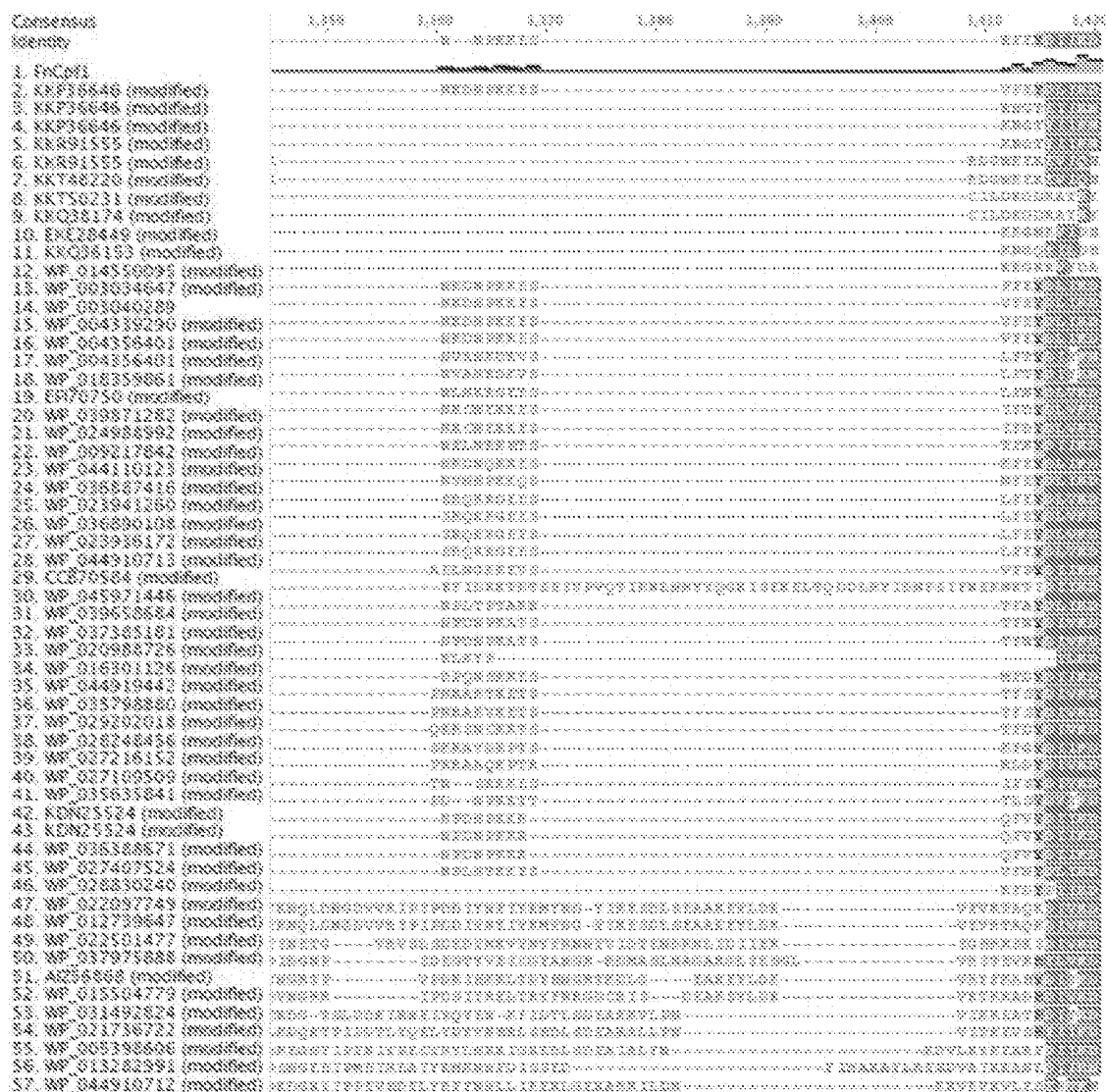


FIG. 38T

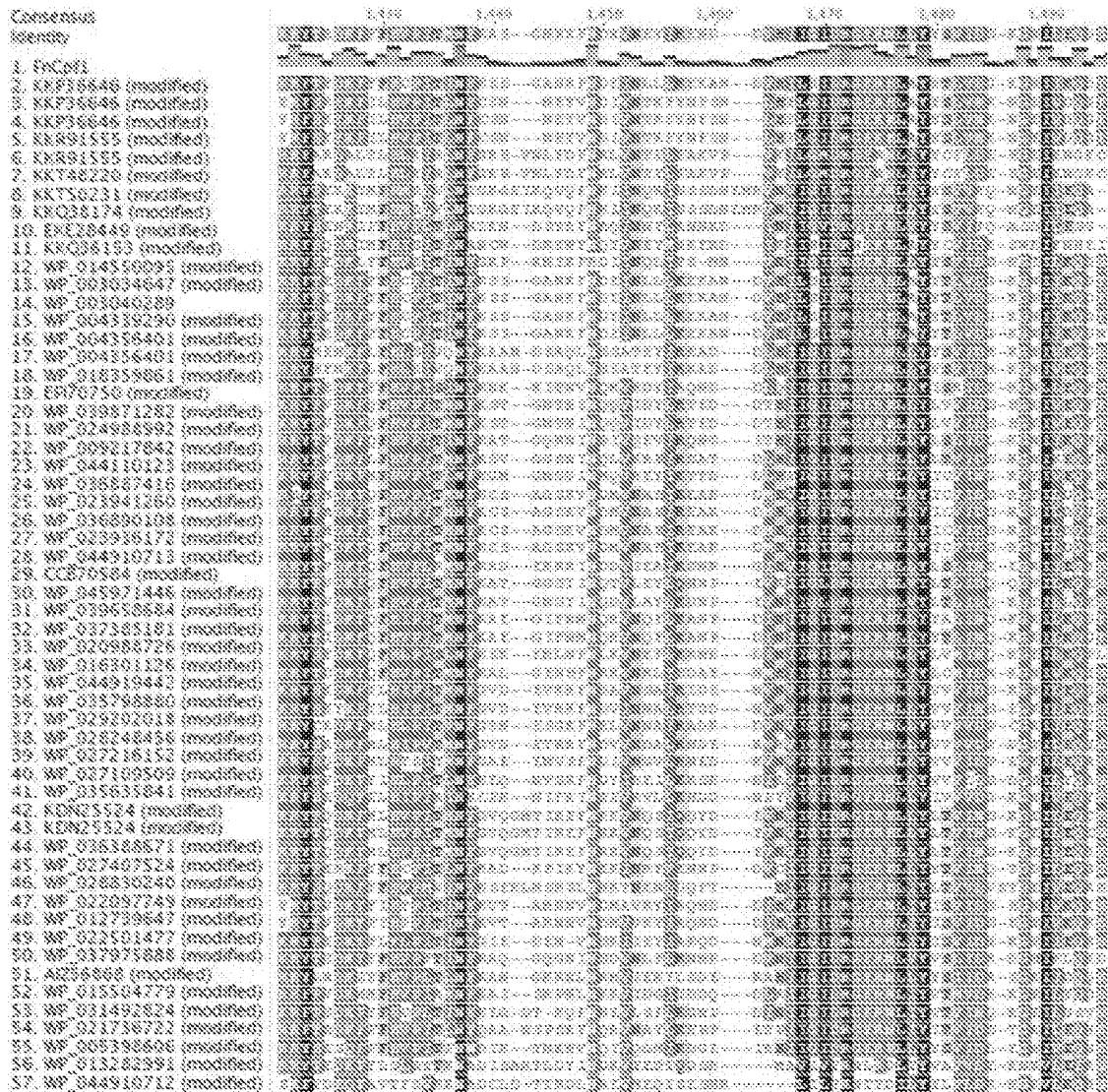


FIG. 38U

Consensus	1,394	1,414	1,434	1,454	1,474	1,494	1,514
Identity	97.8	97.8	97.8	97.8	97.8	97.8	97.8
1. EnCef1							
2. KKP36646 (modified)							
3. KKP36646 (modified)							
4. KKP36646 (modified)							
5. KKR91555 (modified)							
6. KKR91555 (modified)							
7. KKT48220 (modified)							
8. KKT50231 (modified)							
9. KKG38174 (modified)							
10. EKE26449 (modified)							
11. KKG36153 (modified)							
12. WP_014556095 (modified)							
13. WP_003034647 (modified)							
14. WP_003040289 (modified)							
15. WP_004339290 (modified)							
16. WP_004356401 (modified)							
17. WP_004356401 (modified)							
18. WP_018359661 (modified)							
19. EKE70750 (modified)							
20. WP_019871282 (modified)							
21. WP_024868992 (modified)							
22. WP_003217642 (modified)							
23. WP_044110123 (modified)							
24. WP_016887416 (modified)							
25. WP_023941260 (modified)							
26. WP_036690108 (modified)							
27. WP_023936172 (modified)							
28. WP_044910713 (modified)							
29. CC870584 (modified)							
30. WP_045971446 (modified)							
31. WP_019656684 (modified)							
32. WP_017365181 (modified)							
33. WP_020887726 (modified)							
34. WP_016301138 (modified)							
35. WP_044919442 (modified)							
36. WP_016796880 (modified)							
37. WP_029202018 (modified)							
38. WP_028248416 (modified)							
39. WP_027216152 (modified)							
40. WP_027109509 (modified)							
41. WP_035635841 (modified)							
42. KEN25524 (modified)							
43. KEN25524 (modified)							
44. WP_036588571 (modified)							
45. WP_027407524 (modified)							
46. WP_028630240 (modified)							
47. WP_022087740 (modified)							
48. WP_012739647 (modified)							
49. WP_022501477 (modified)							
50. WP_037975888 (modified)							
51. A056860 (modified)							
52. WP_015504779 (modified)							
53. WP_011492824 (modified)							
54. WP_021736722 (modified)							
55. WP_005398606 (modified)							
56. WP_013282991 (modified)							
57. WP_044910712 (modified)							

FIG. 38V

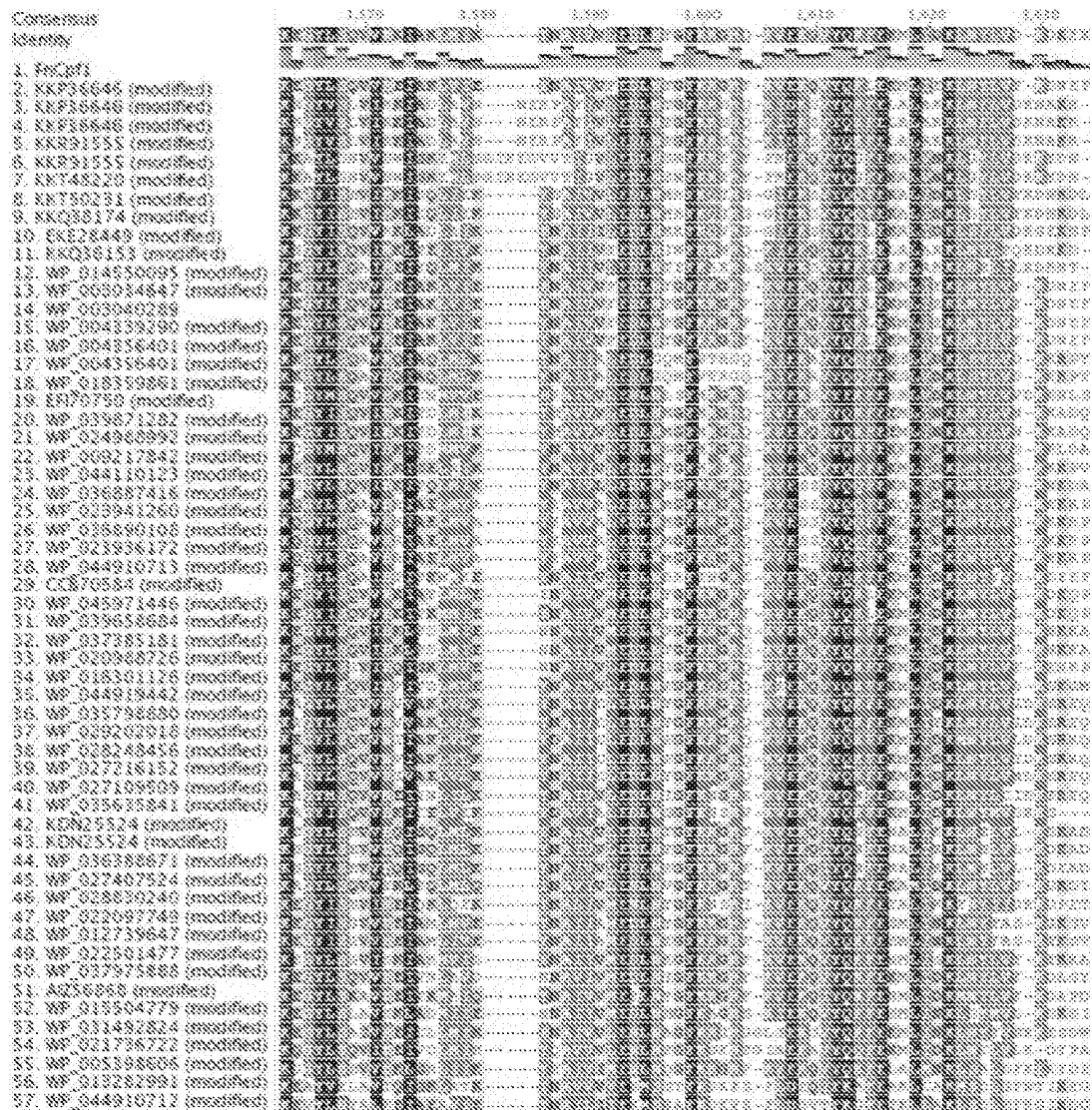


FIG. 38W

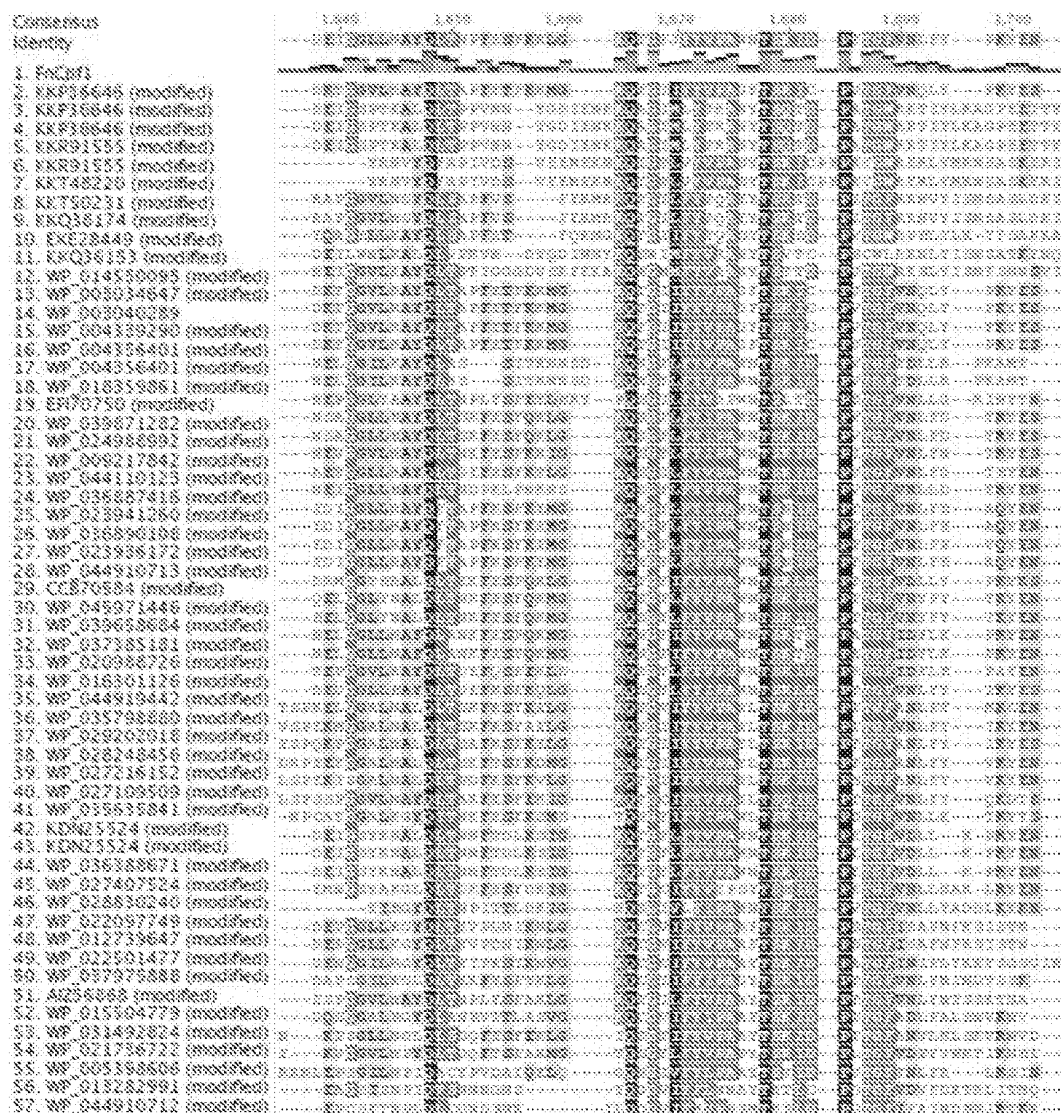


FIG. 38X

Consensus Identity	1,738	1,739	1,740	1,741	1,742	1,743	1,744	1,745
1. FncCpf1								
2. KKFP36846 (modified)								
3. KKFP36846 (modified)								
4. KKFP36846 (modified)								
5. KK9915153 (modified)								
6. KR881515 (modified)								
7. HKT488220 (modified)								
8. HKT50231 (modified)								
9. KKC38174 (modified)								
10. EIE28449 (modified)								
11. KKG38153 (modified)								
12. WP_014556095 (modified)								
13. WP_003034647 (modified)								
14. WP_003040289								
15. WP_004338290 (modified)								
16. WP_004338401 (modified)								
17. WP_004338401 (modified)								
18. WP_018359861 (modified)								
19. EBI070750 (modified)								
20. WP_039871282 (modified)								
21. WP_024988992 (modified)								
22. WP_009213642 (modified)								
23. WP_044110123 (modified)								
24. WP_026887416 (modified)								
25. WP_023941260 (modified)								
26. WP_036690108 (modified)								
27. WP_028938172 (modified)								
28. WP_044910713 (modified)								
29. CC870584 (modified)								
30. WP_045971448 (modified)								
31. WP_039858884 (modified)								
32. WP_037385181 (modified)								
33. WP_020988728 (modified)								
34. WP_016301128 (modified)								
35. WP_044919442 (modified)								
36. WP_032998880 (modified)								
37. WP_029202018 (modified)								
38. WP_028248438 (modified)								
39. WP_027218152 (modified)								
40. WP_027189508 (modified)								
41. WP_035635841 (modified)								
42. KD425524 (modified)								
43. KD425524 (modified)								
44. WP_036388871 (modified)								
45. WP_027407524 (modified)								
46. WP_028830240 (modified)								
47. WP_022887749 (modified)								
48. WP_012739647 (modified)								
49. WP_022501477 (modified)								
50. WP_037975888 (modified)								
51. AIZ68868 (modified)								
52. WP_015504779 (modified)								
53. WP_031492824 (modified)								
54. WP_021758732 (modified)								
55. WP_005398866 (modified)								
56. WP_013282991 (modified)								
57. WP_044910712 (modified)								

FIG. 38Y

Consensus	1,790	1,795	1,800	1,810	1,820	1,830	1,840
Identity	-----						
1. EnCpf1							
2. KKP36846 (modified)							
3. KKP36846 (modified)							
4. KKP36846 (modified)							
5. KKR91553 (modified)							
6. KKR91553 (modified)							
7. KKT48020 (modified)							
8. KKT50231 (modified)							
9. KKT50231 (modified)							
10. EKE26449 (modified)							
11. KKT50231 (modified)							
12. WP_014550095 (modified)							
13. WP_003034647 (modified)							
14. WP_003040269							
15. WP_004035290 (modified)							
16. WP_004035401 (modified)							
17. WP_004155401 (modified)							
18. WP_018359866 (modified)							
19. ERI70750 (modified)							
20. WP_019871282 (modified)							
21. WP_024988992 (modified)							
22. WP_009217842 (modified)							
23. WP_044110125 (modified)							
24. WP_016867416 (modified)							
25. WP_023941260 (modified)							
26. WP_036880108 (modified)							
27. WP_023036172 (modified)							
28. WP_044910713 (modified)							
29. CCE70584 (modified)							
30. WP_045971448 (modified)							
31. WP_039658584 (modified)							
32. WP_037365181 (modified)							
33. WP_020988726 (modified)							
34. WP_016301126 (modified)							
35. WP_044919442 (modified)							
36. WP_035798880 (modified)							
37. WP_029202019 (modified)							
38. WP_028248416 (modified)							
39. WP_027216152 (modified)							
40. WP_027109509 (modified)							
41. WP_035635841 (modified)							
42. KDN25524 (modified)							
43. KDN25524 (modified)							
44. WP_036388571 (modified)							
45. WP_027407524 (modified)							
46. WP_028830240 (modified)							
47. WP_022087749 (modified)							
48. WP_012739647 (modified)							
49. WP_023501477 (modified)							
50. WP_037975888 (modified)							
51. A256860 (modified)							
52. WP_015504779 (modified)							
53. WP_031492824 (modified)							
54. WP_021736722 (modified)							
55. WP_005398696 (modified)							
56. WP_013282991 (modified)							
57. WP_044910712 (modified)							

FIG. 38Z

Consensus Identity	1,879	1,880	1,879	1,880	1,881	1,880	1,879
1. FrCpf1							
2. KKP36646 (modified)							
3. KKP36646 (modified)							
4. KKP36646 (modified)							
5. KKP36646 (modified)							
6. KKP36646 (modified)							
7. KKP36646 (modified)							
8. KKP36646 (modified)							
9. KKP36646 (modified)							
10. KKP36646 (modified)							
11. KKP36646 (modified)							
12. WP_014650095 (modified)							
13. WP_003014847 (modified)							
14. WP_003040288							
15. WP_004139190 (modified)							
16. WP_004356401 (modified)							
17. WP_004356401 (modified)							
18. WP_018319861 (modified)							
19. EPI70750 (modified)							
20. WP_039871282 (modified)							
21. WP_024968892 (modified)							
22. WP_009217842 (modified)							
23. WP_044110123 (modified)							
24. WP_036887418 (modified)							
25. WP_023941260 (modified)							
26. WP_036887418 (modified)							
27. WP_023941260 (modified)							
28. WP_044910713 (modified)							
29. CC570554 (modified)							
30. WP_045971446 (modified)							
31. WP_039658884 (modified)							
32. WP_037385181 (modified)							
33. WP_020968726 (modified)							
34. WP_018301126 (modified)							
35. WP_044919442 (modified)							
36. WP_035798680 (modified)							
37. WP_009203010 (modified)							
38. WP_028243456 (modified)							
39. WP_027218152 (modified)							
40. WP_027109509 (modified)							
41. WP_035635841 (modified)							
42. KDN25524 (modified)							
43. KDN25524 (modified)							
44. WP_036388871 (modified)							
45. WP_027407524 (modified)							
46. WP_028830240 (modified)							
47. WP_022697740 (modified)							
48. WP_012739647 (modified)							
49. WP_022501477 (modified)							
50. WP_037975888 (modified)							
51. AG256868 (modified)							
52. WP_015504779 (modified)							
53. WP_031482824 (modified)							
54. WP_021736722 (modified)							
55. WP_005398806 (modified)							
56. WP_013282991 (modified)							
57. WP_044910712 (modified)							

FIG. 38AA

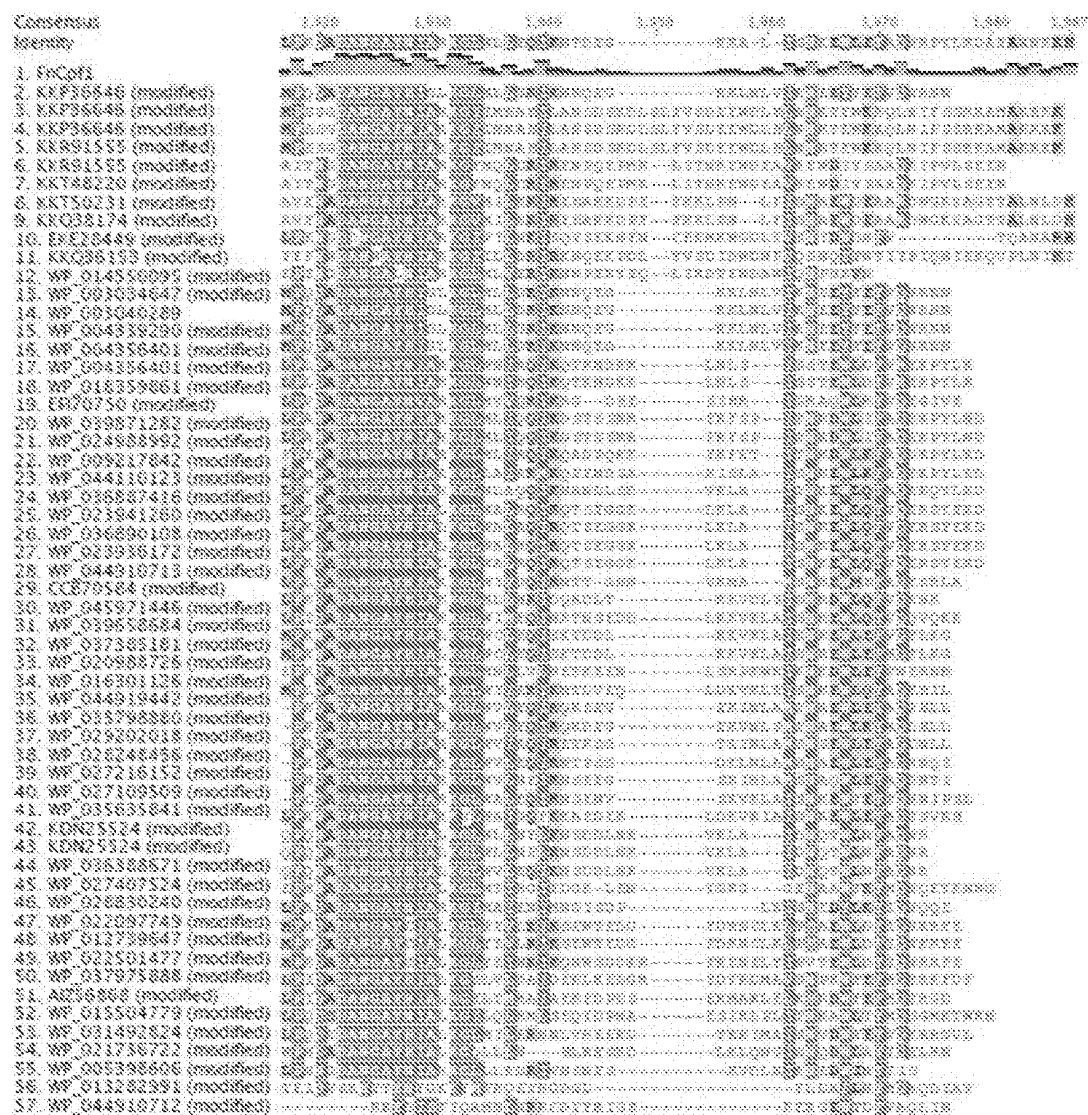


FIG. 38BB

FIG. 39A

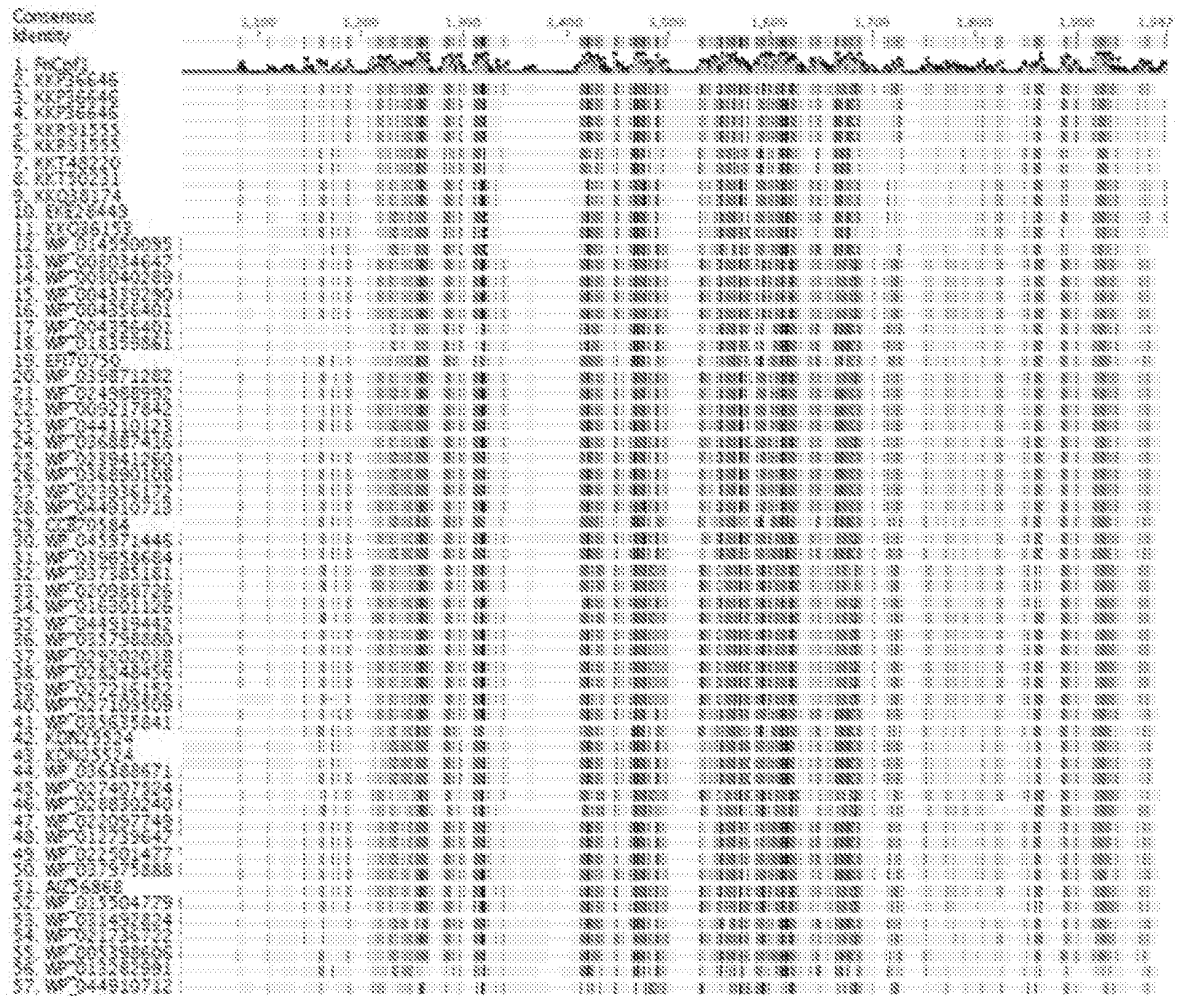


FIG. 39B

pacrc184 Pacp1 locus (p001)

[illegible]

FIG. 40A

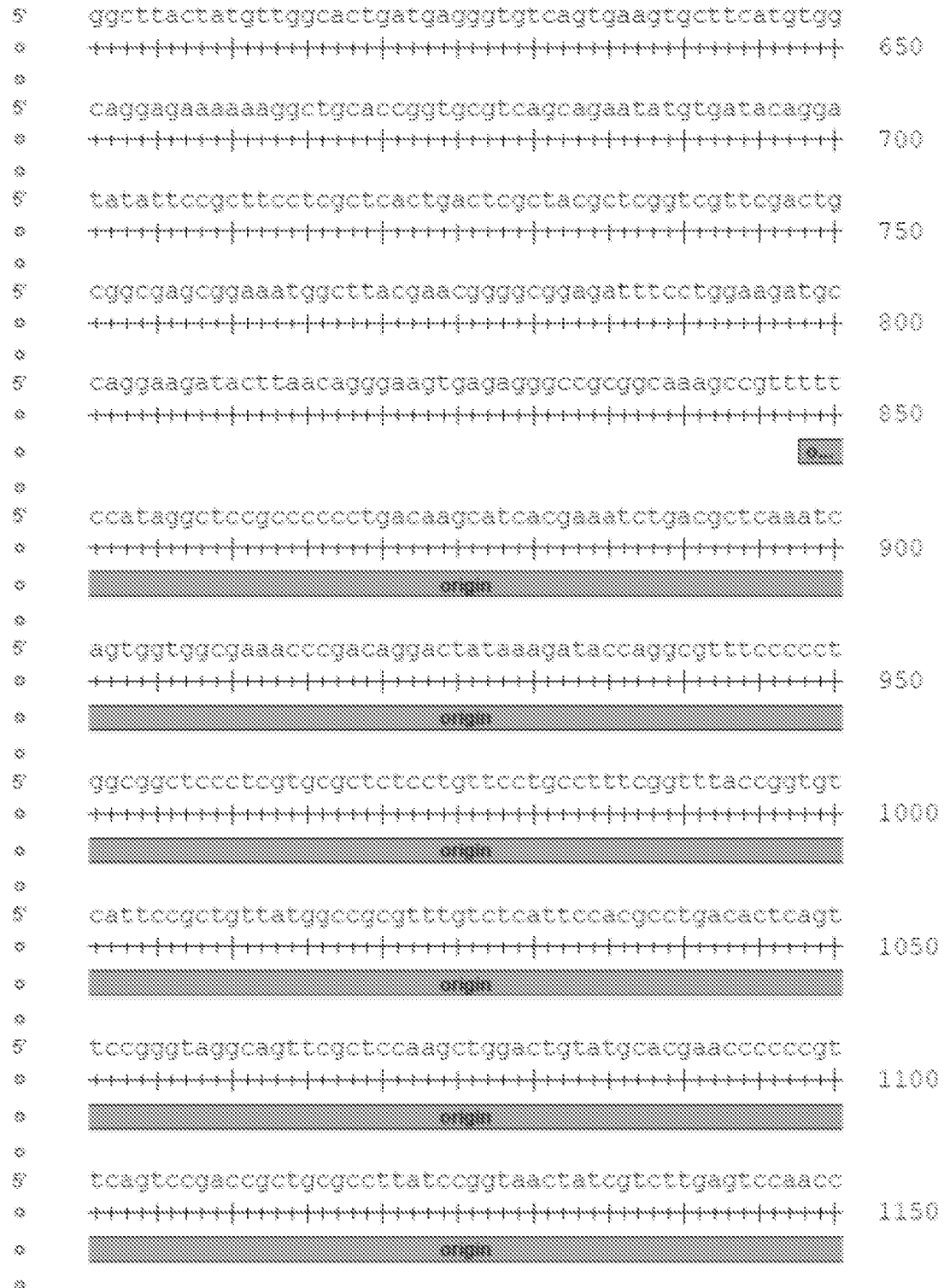
pACYC184 *FacPfl* locus (pY001)

FIG. 40B

pACYC184 FmCpf1 locus (pY001)

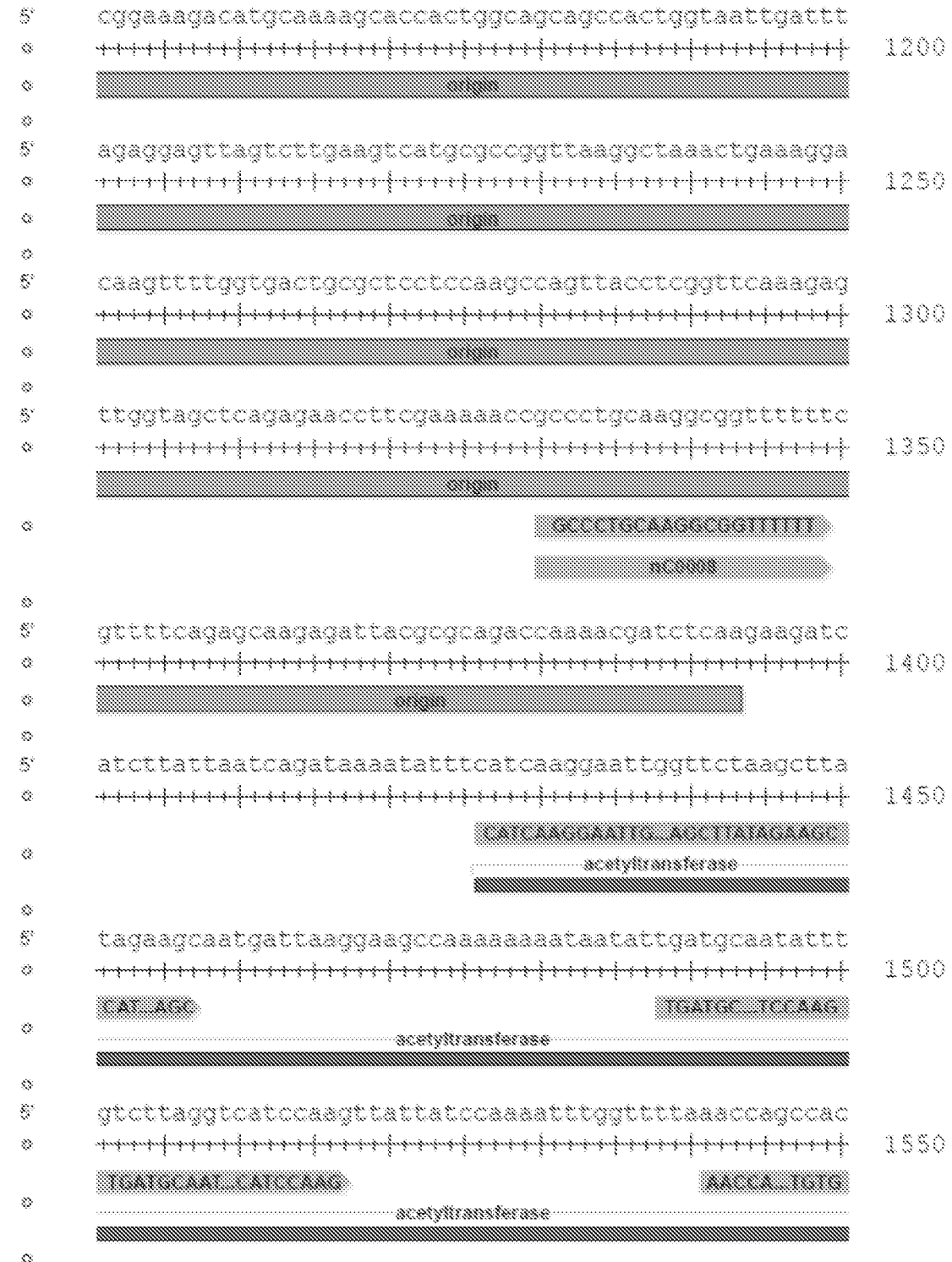


FIG. 40C

pACYC184 Fncpf1 locus (pY001)

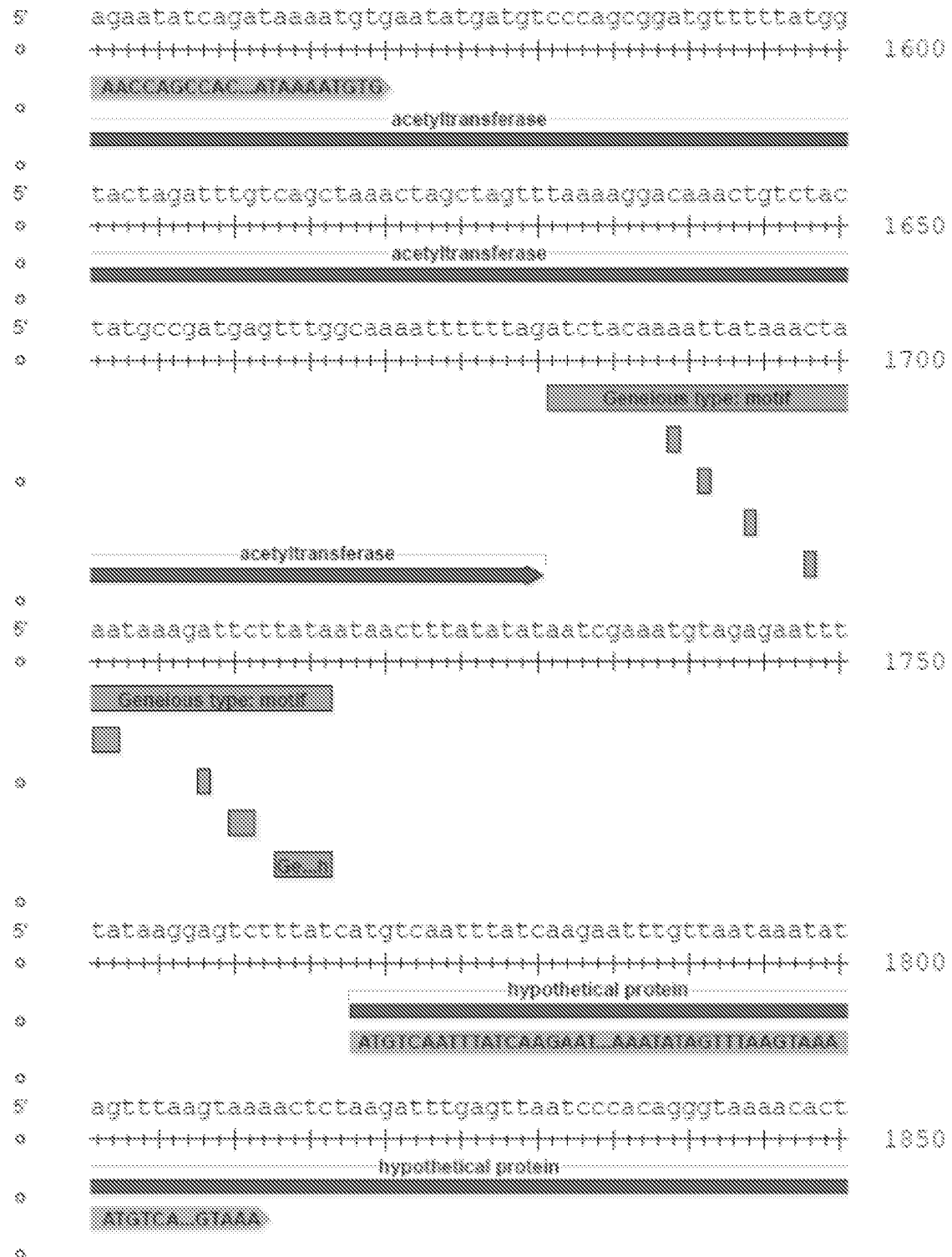


FIG. 40D

pACYC184 FncPfl locus (pY001)

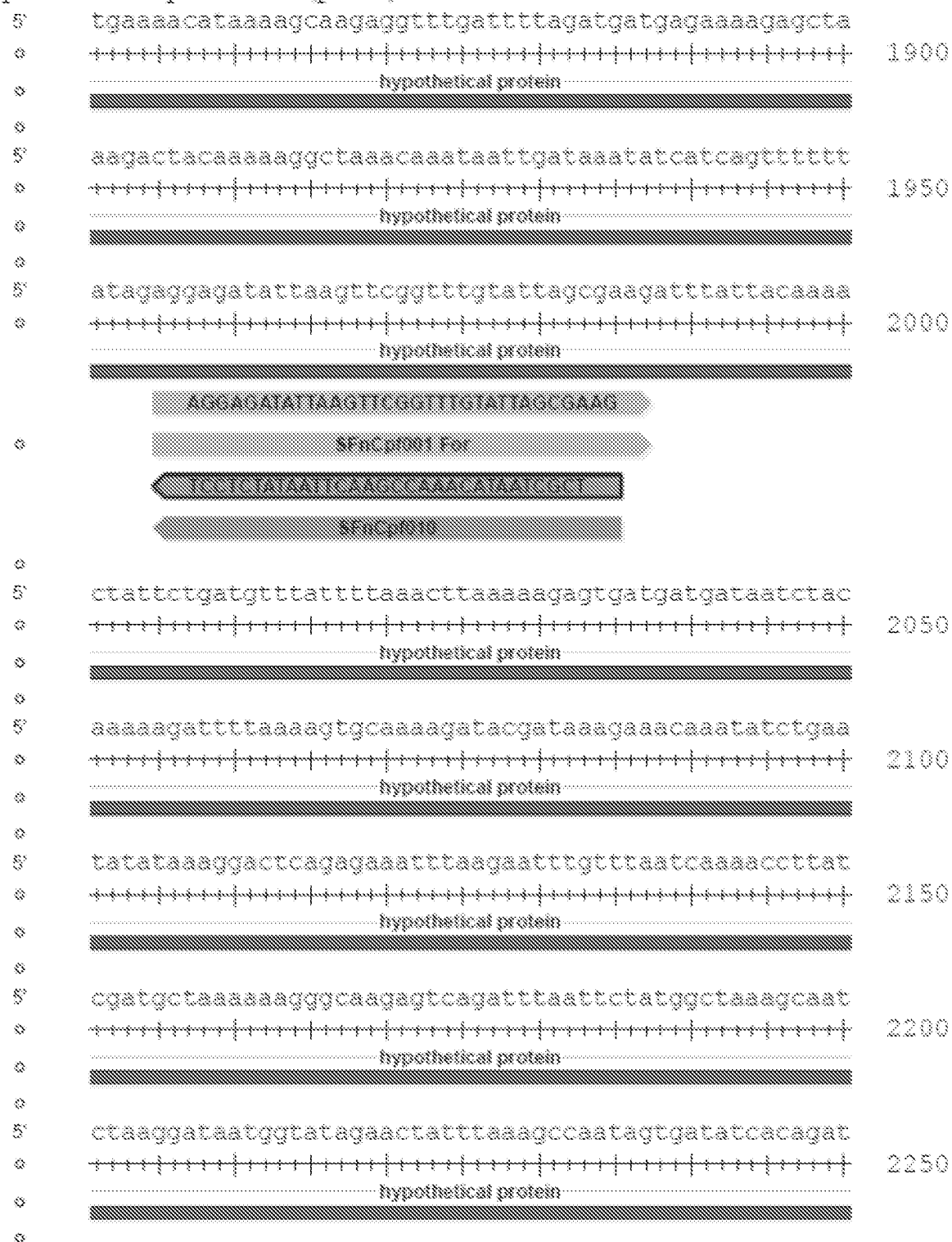


FIG. 40E

pACYC184 FncPfl locus (pY001)

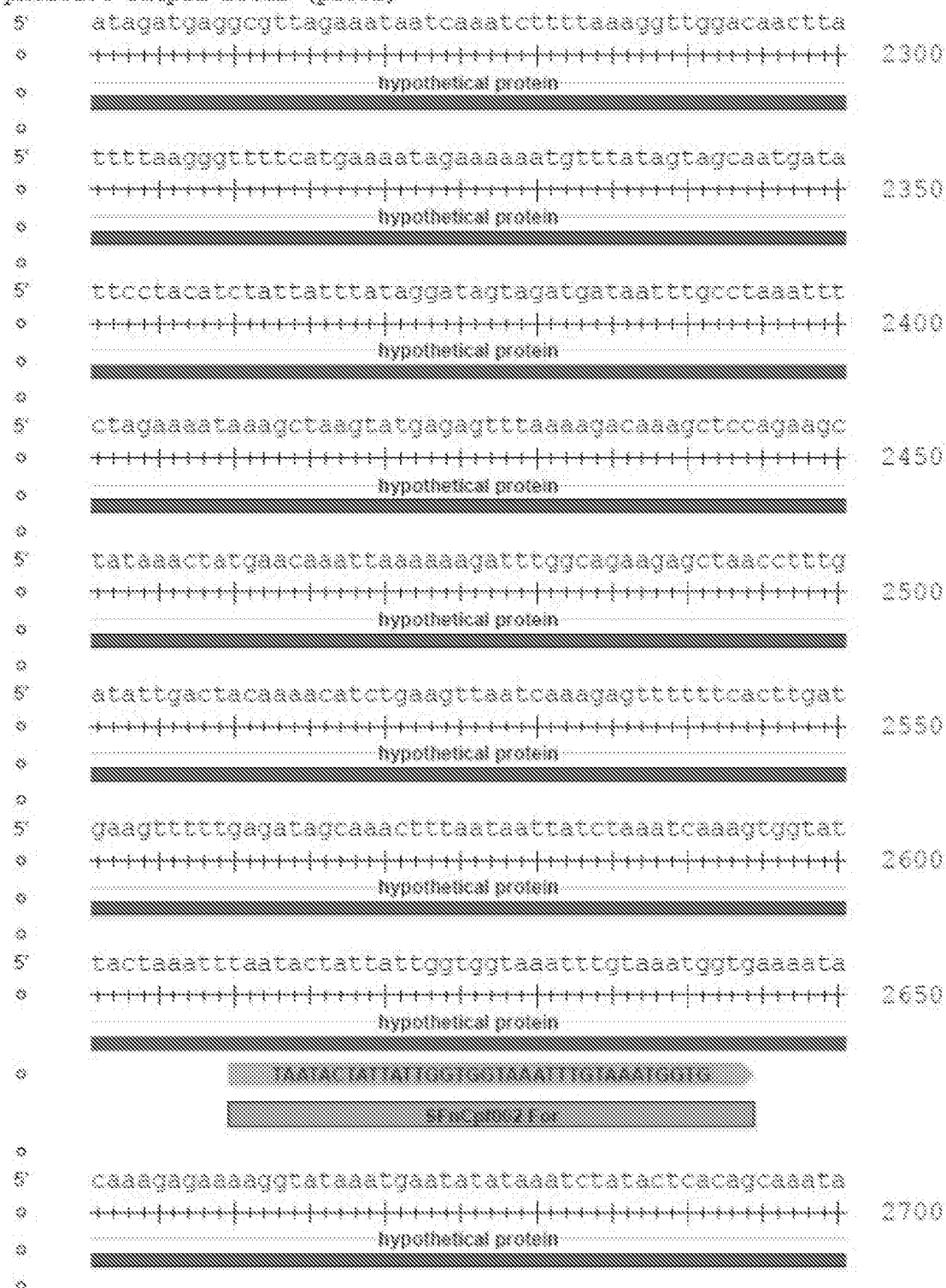


FIG. 40F

pACYC184 FncPfl locus (pY001)

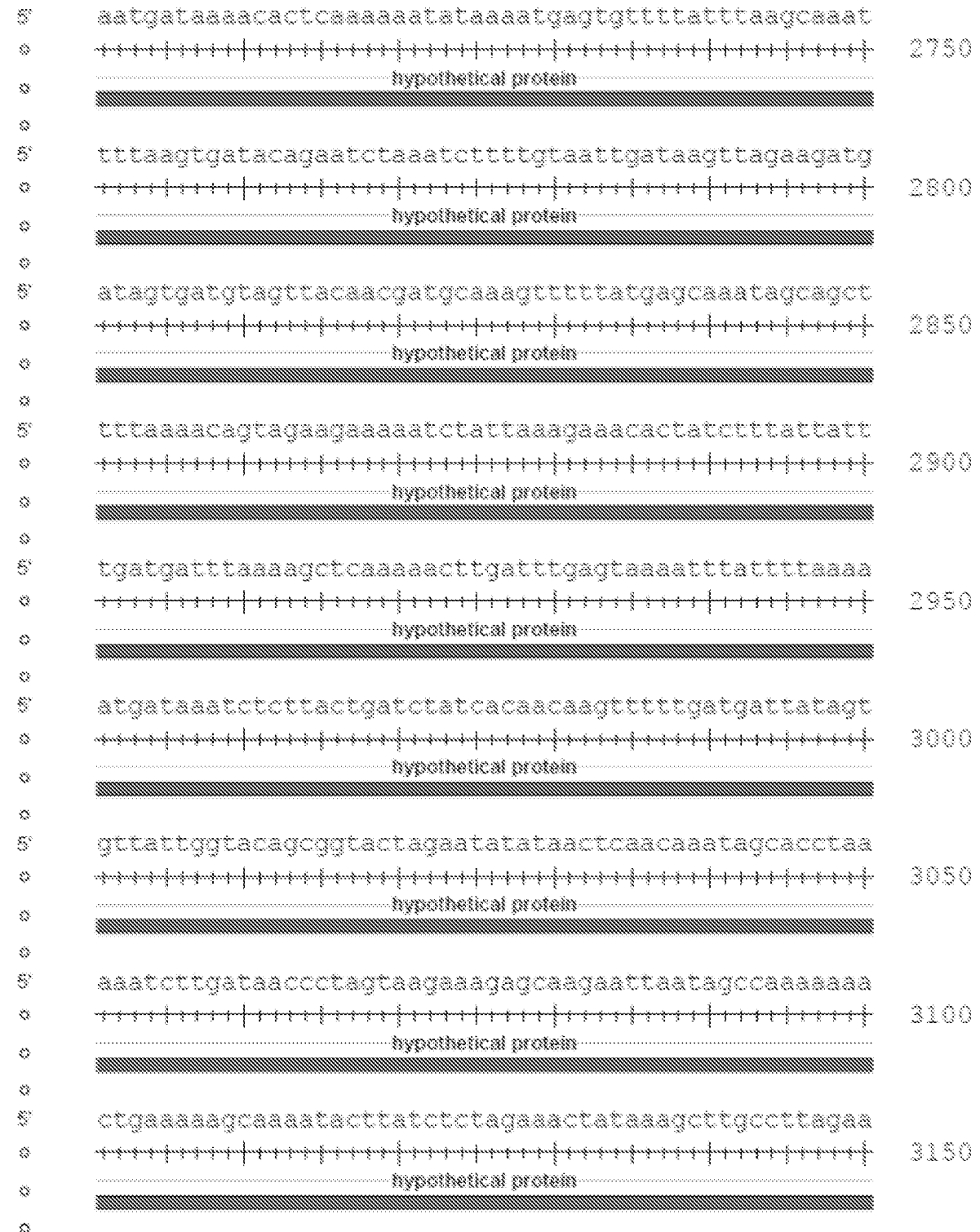


FIG. 40G

pACYC184 Fncpfl locus (pY001)

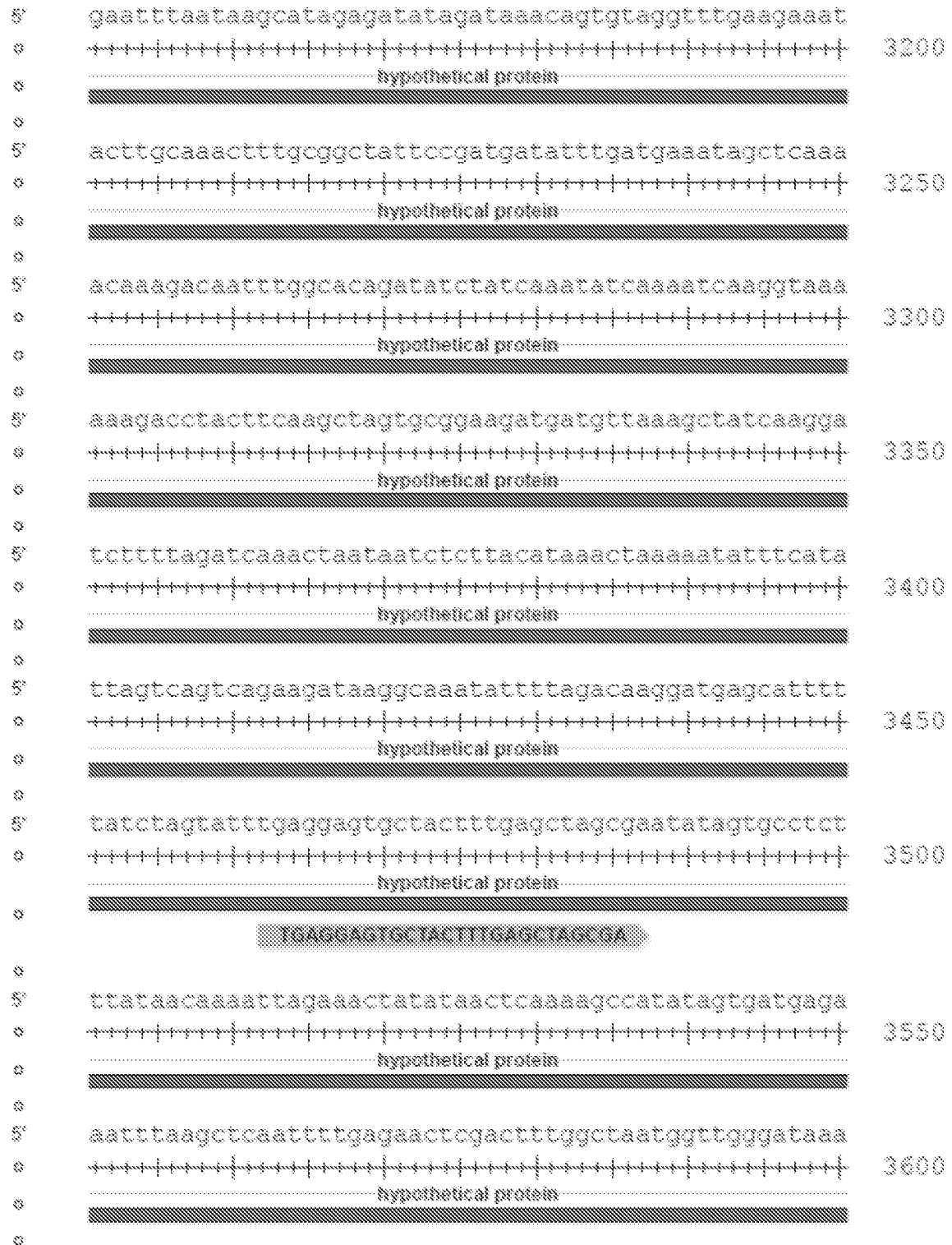


FIG. 40H

pACYC184 FnOpfl locus (pY001)

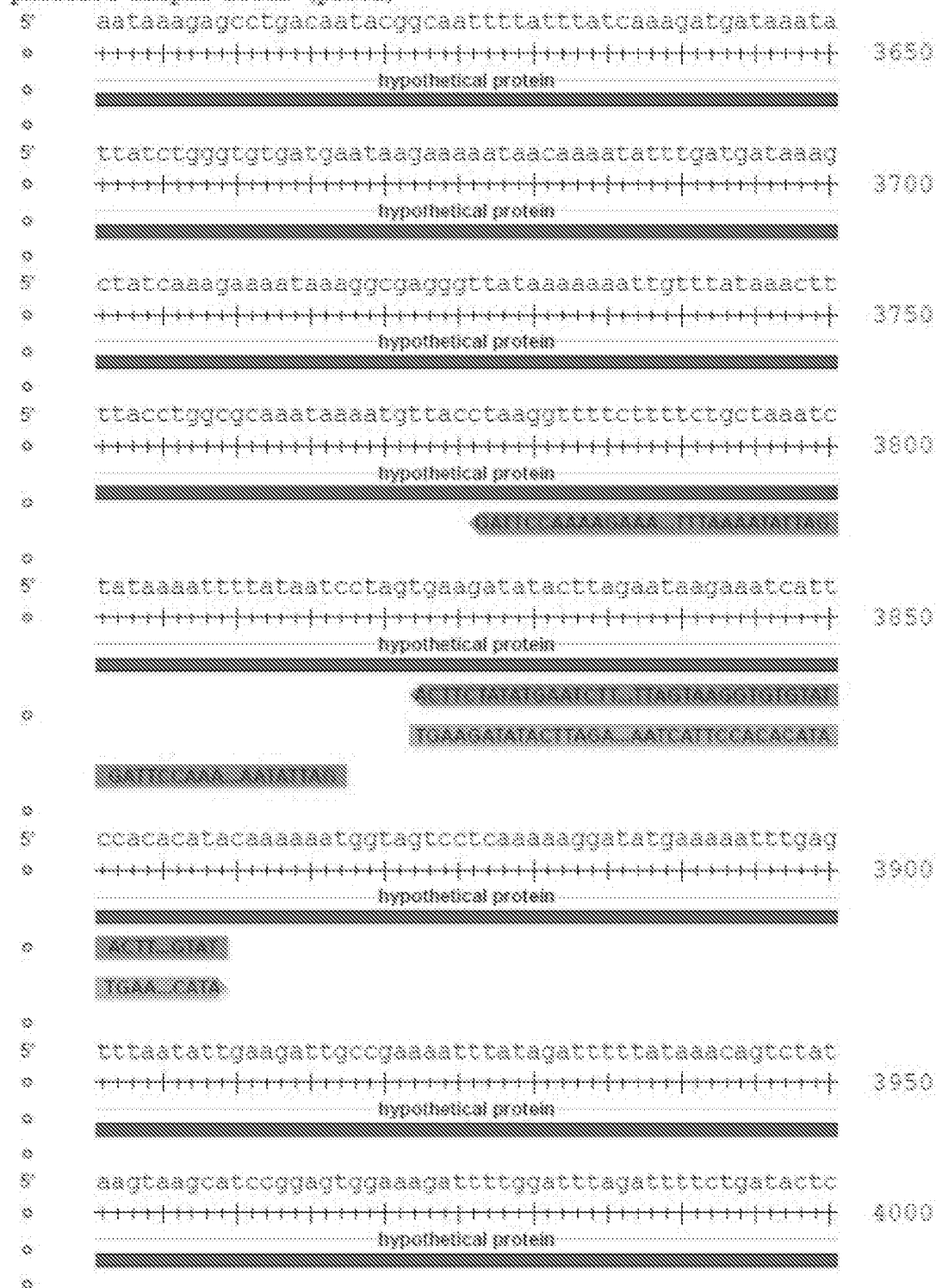


FIG. 40I

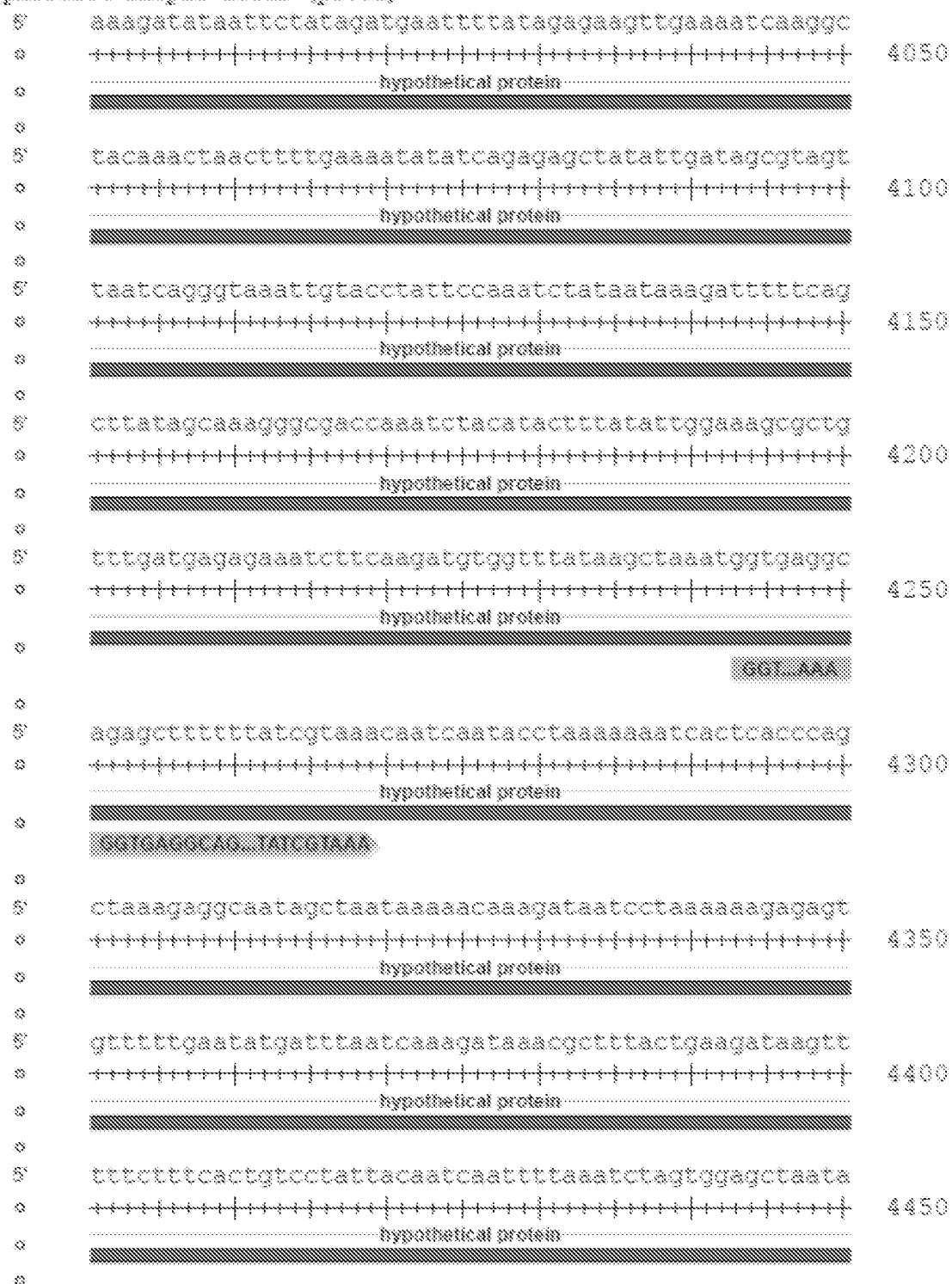
pACYC184 *FnOpfl* locus (pY001)

FIG. 40J

pACYC184 FncPfl locus (pY001)

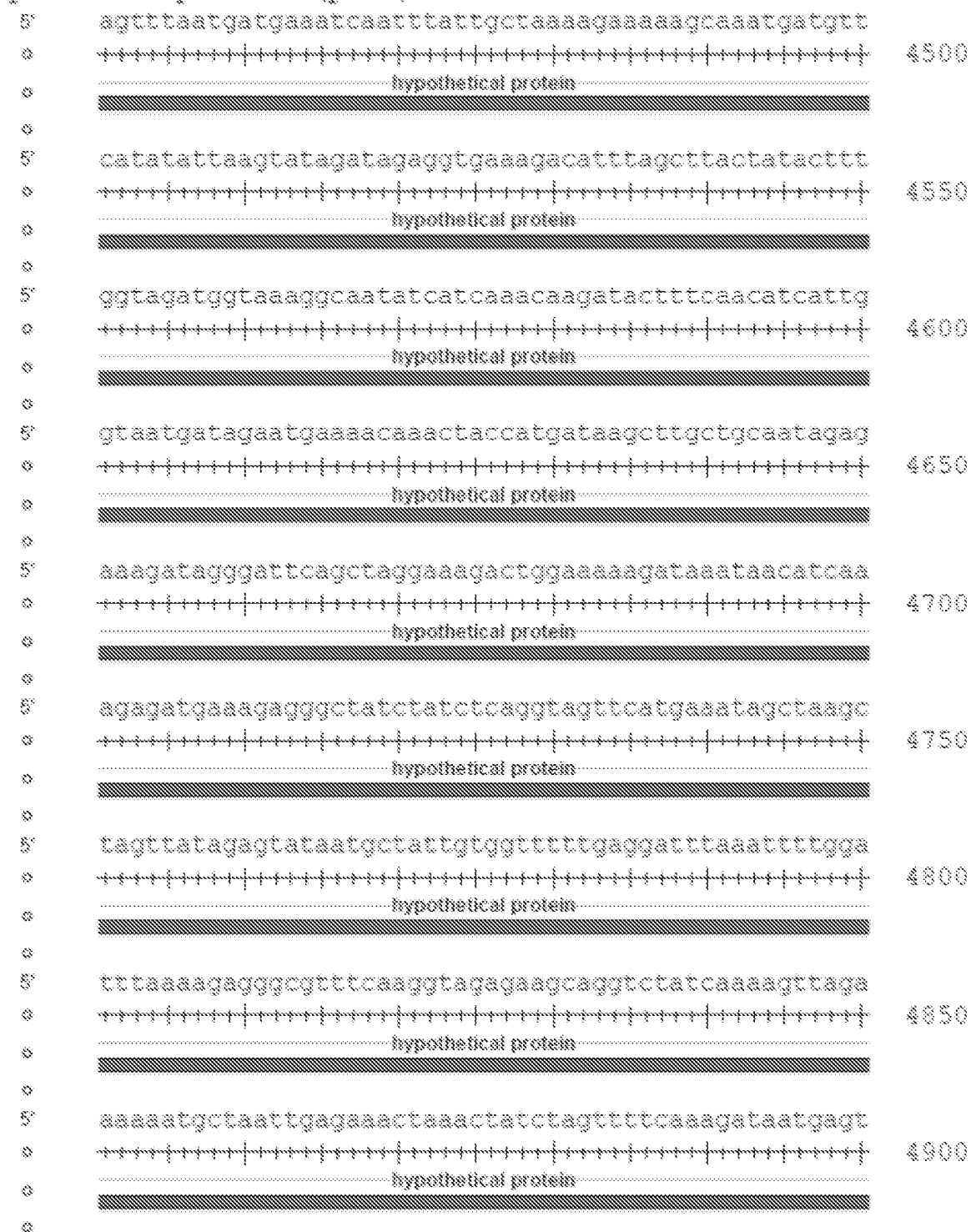


FIG. 40K

pACYC184 FncPfl locus (pY001)

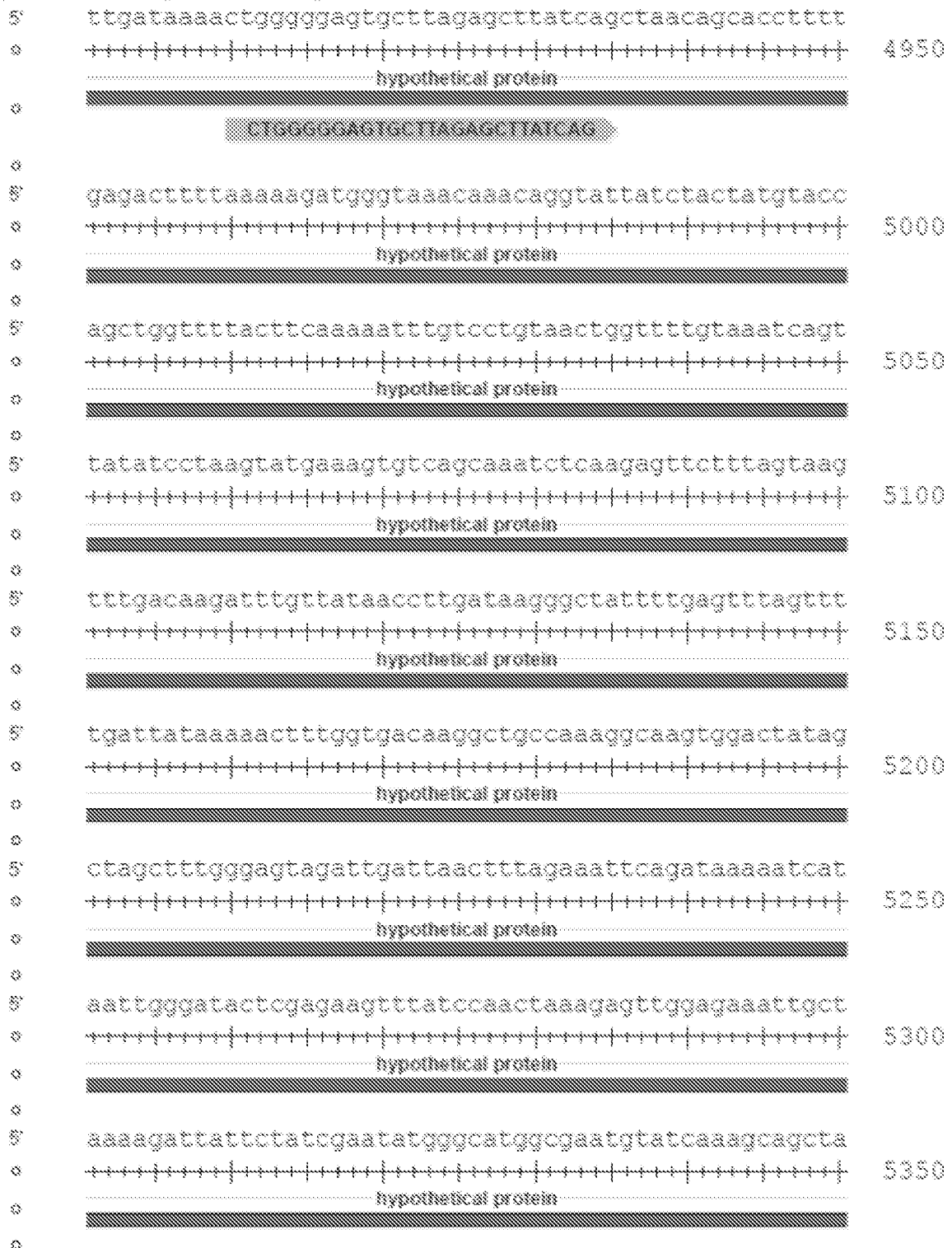


FIG. 40L

pACYC184 FmCpfI locus (pY001)

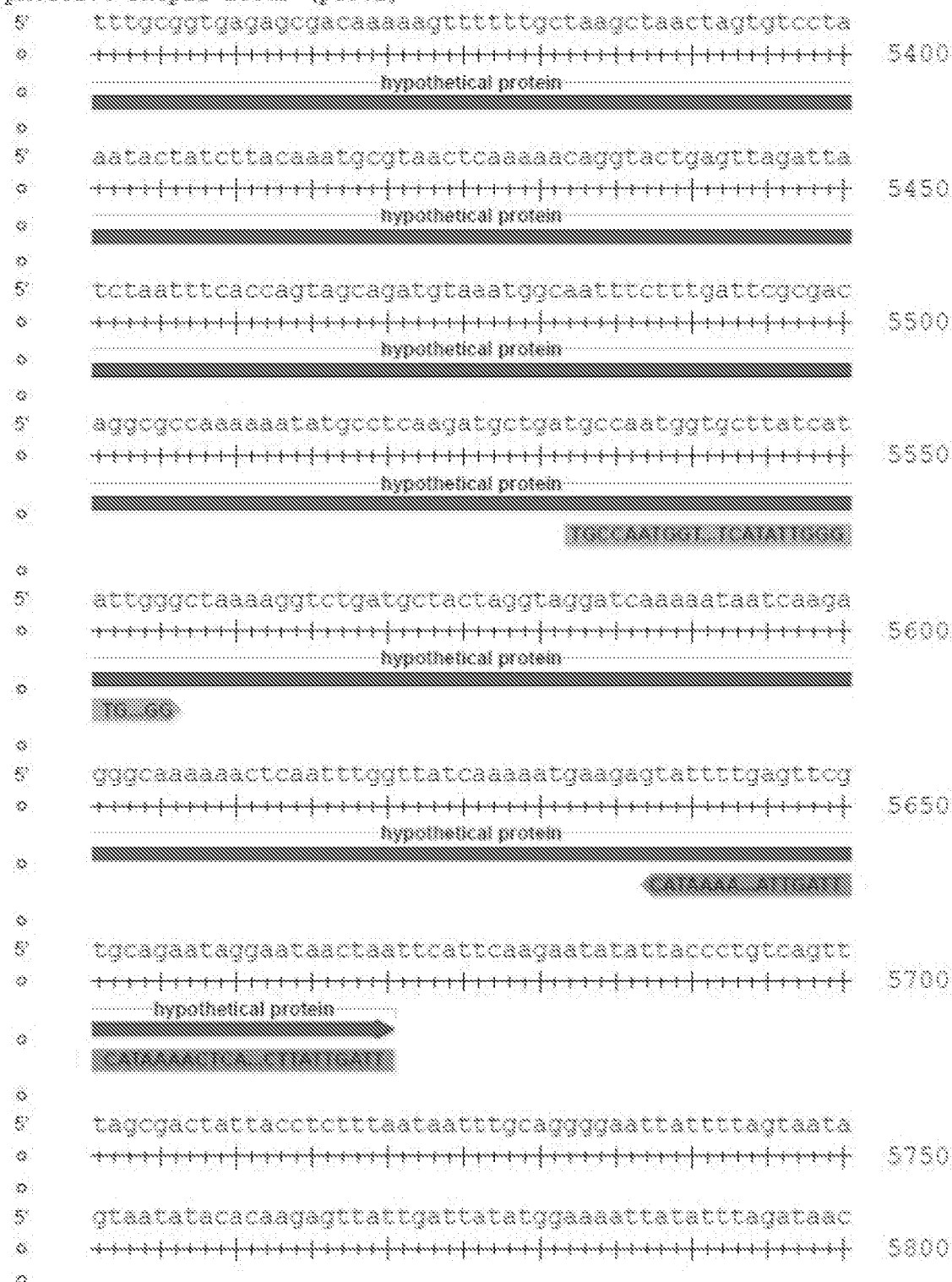


FIG. 40M

pACYC184 FmCpfI locus (pY001)



FIG. 40N

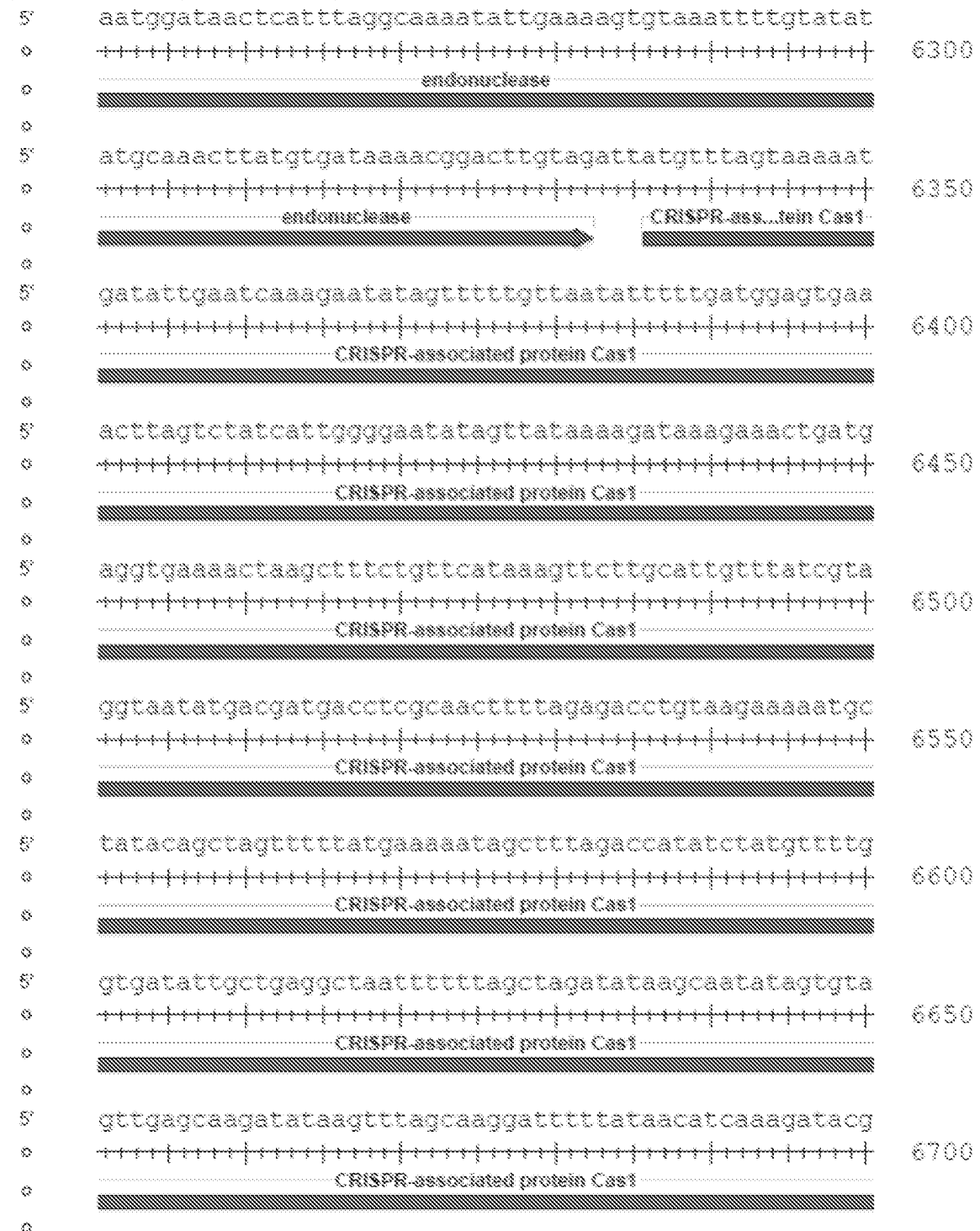


FIG. 400

pACYC184 *FnOpfl* locus (pY001)

FIG. 40P

pACYC184 FncPfl locus (pY001)

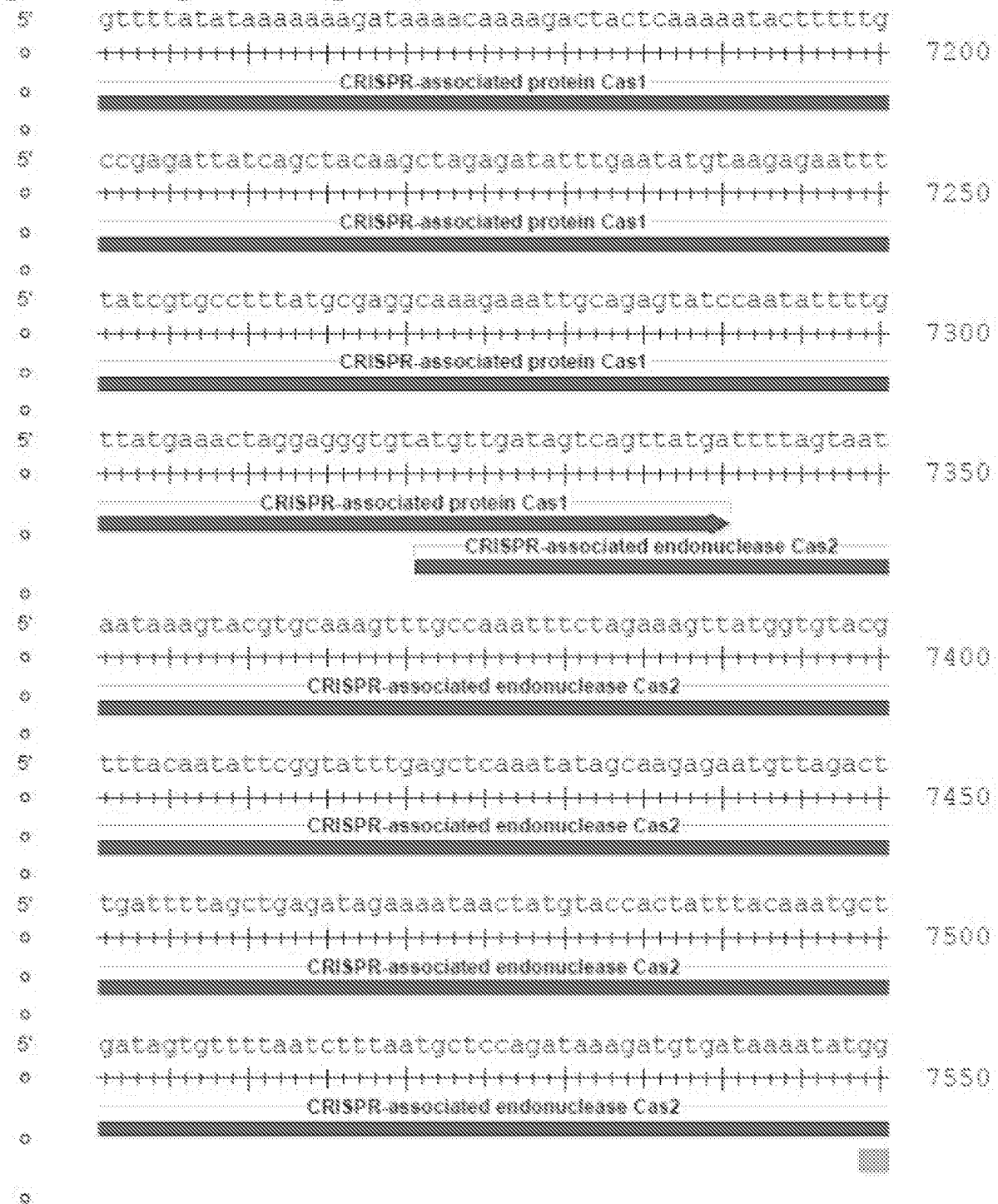


FIG. 40Q

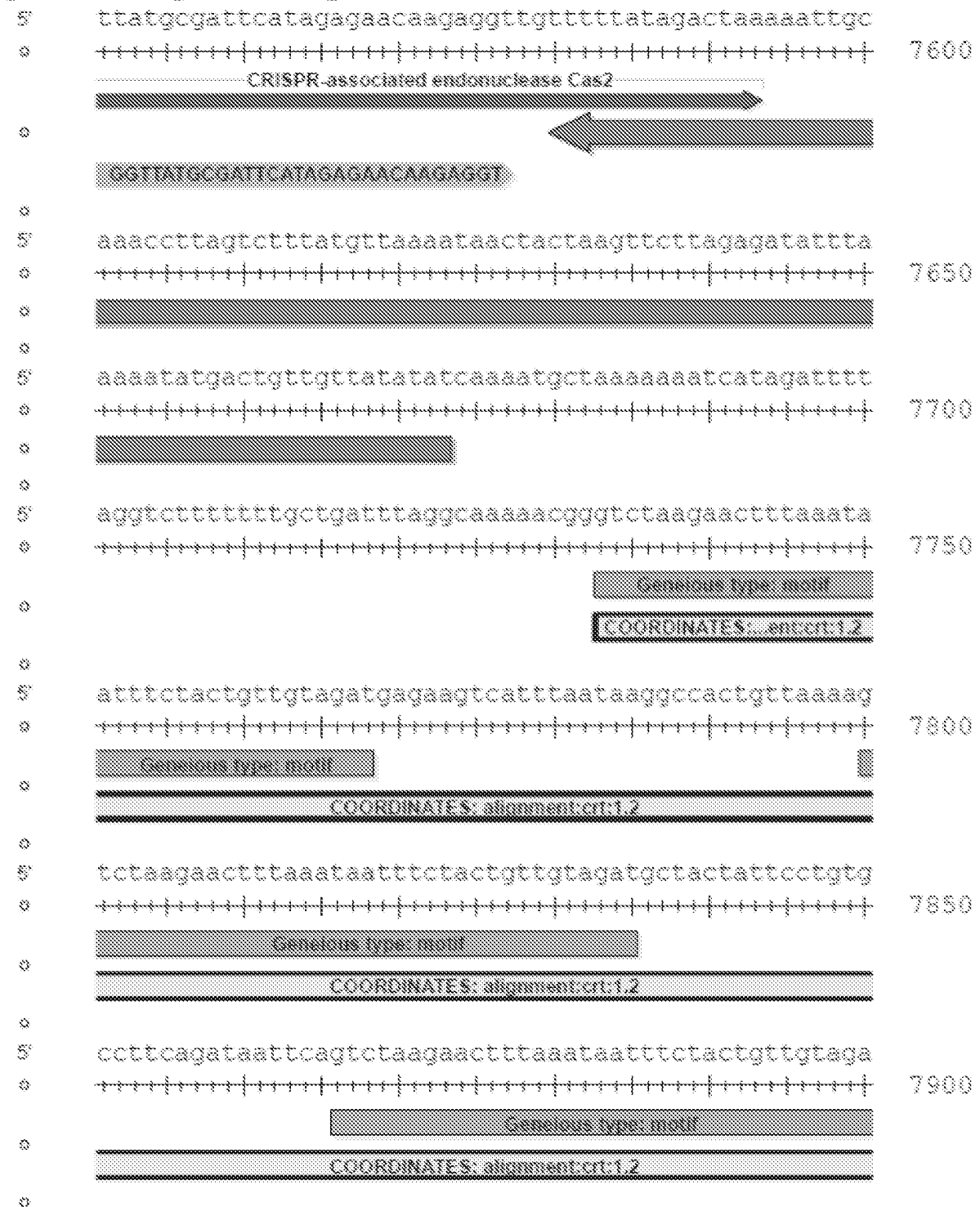


FIG. 40R

pACYC184 FrCpfI locus (pY001)

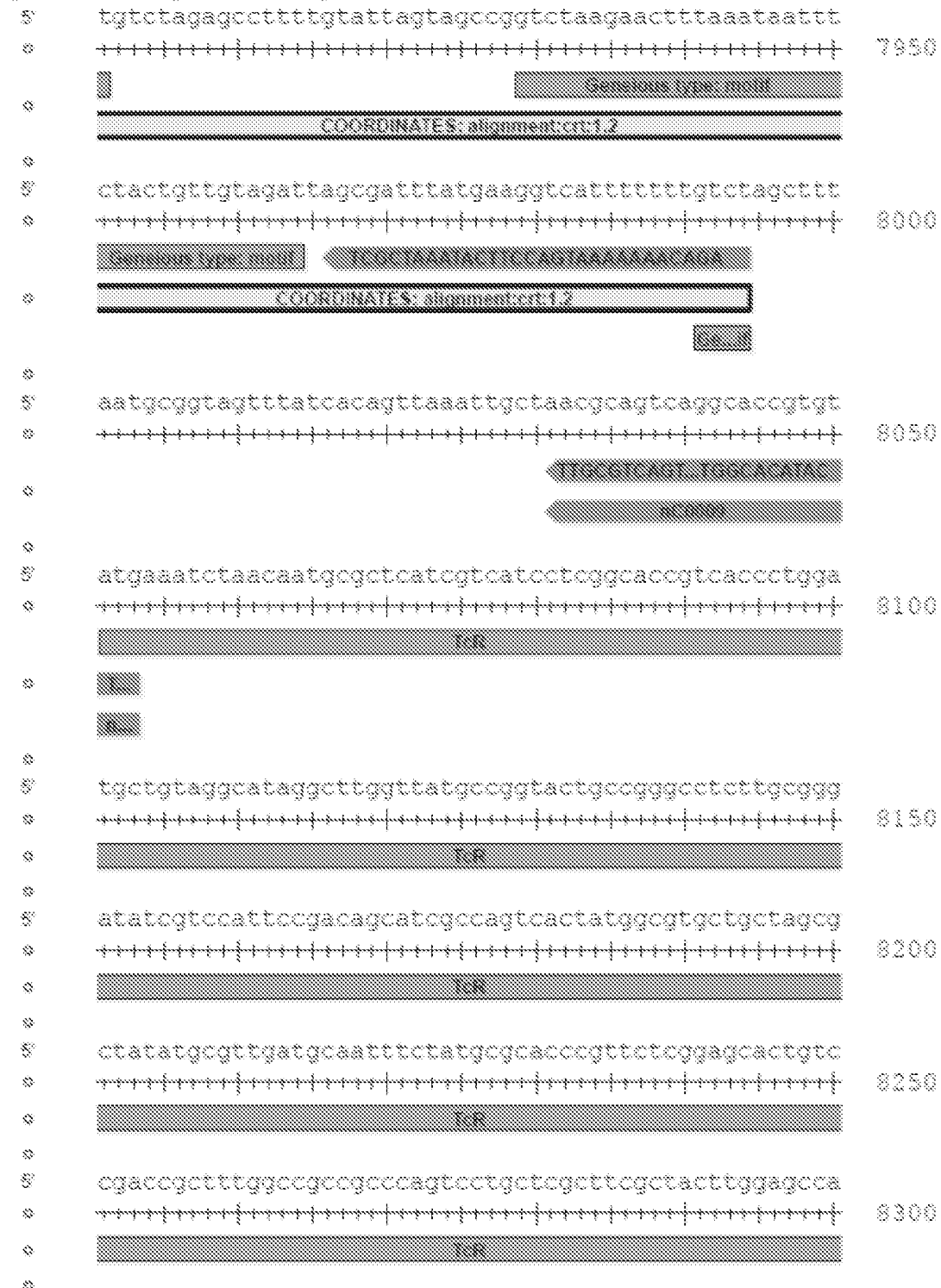


FIG. 40S

pACYC184 FncOpf1 locus (pY001)

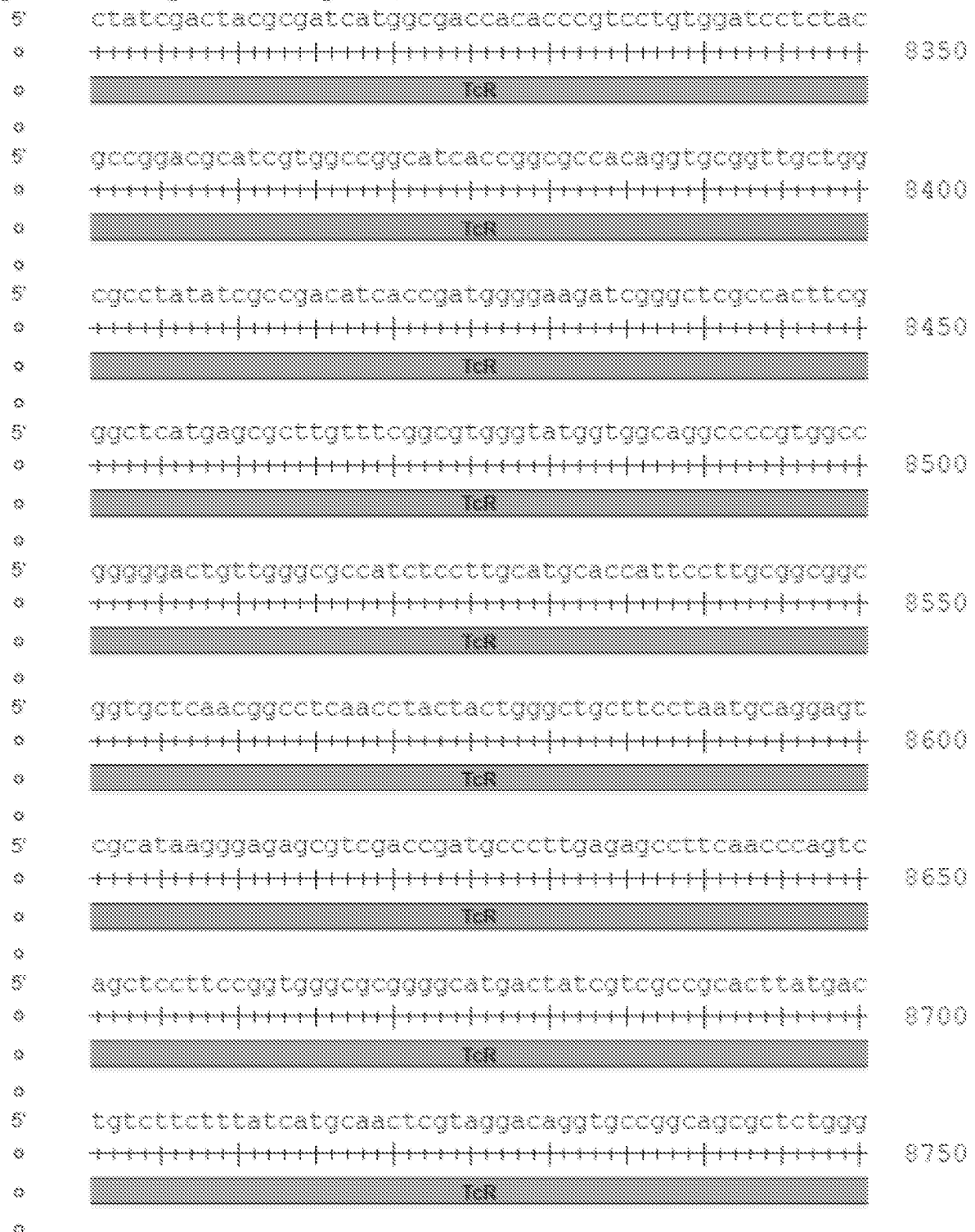


FIG. 40T

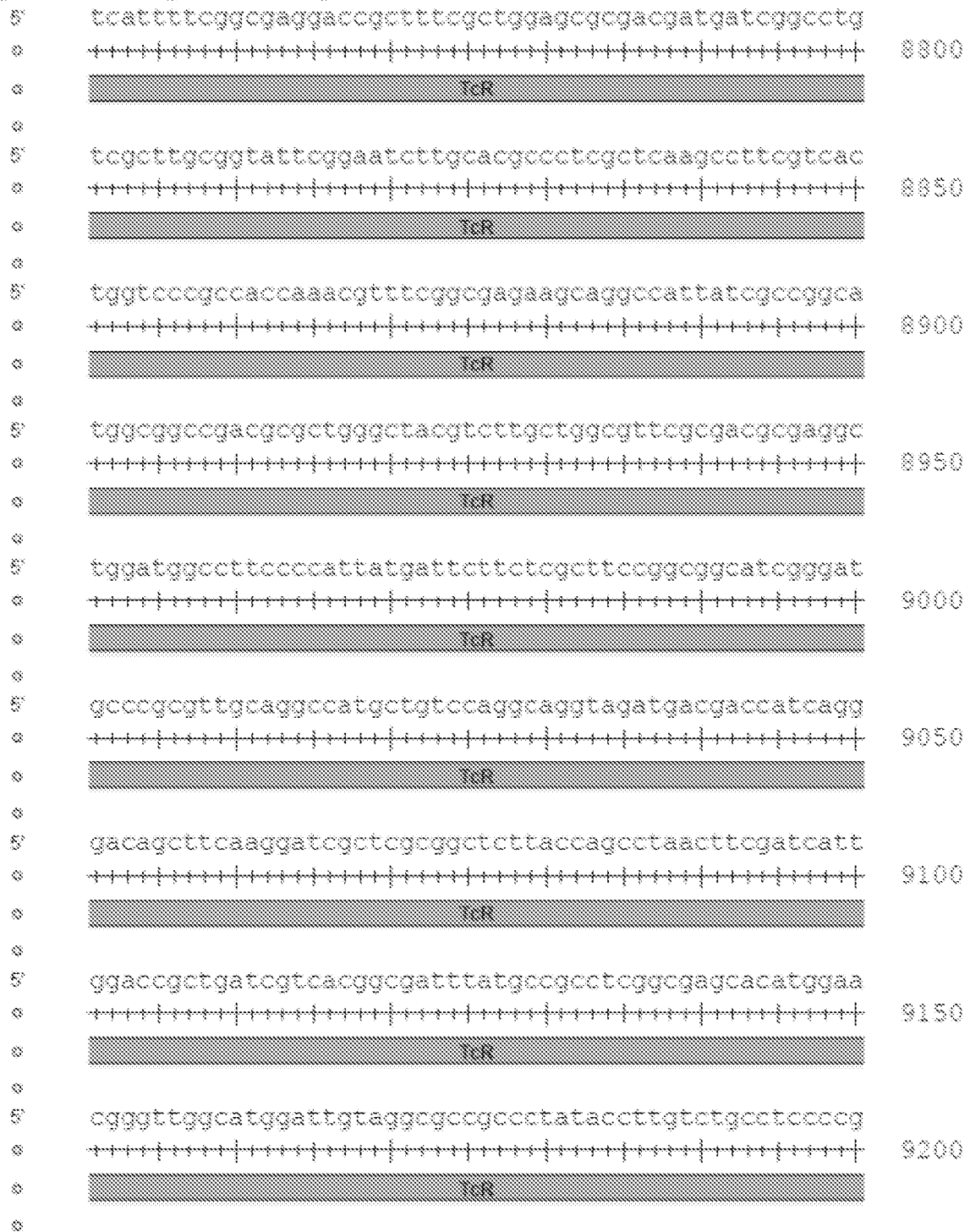
pACYC184 *FnQpf1* locus (pY001)

FIG. 40U

pACYC184 *FnCpfI* locus (pY001)

```

5'   cgttgcgctcgcggtgcatggagccgggccacctcgacctgaatggaagcc
   ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++| 9250
   +-----+-----+-----+-----+-----+-----+-----+-----+
   +
5'   ggcgggcacctcgctaacgggattcaccactccaagaattggagccaatcaa
   ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++| 9300
   +
5'   ttcttgcgggagaactgtgaatgcgcaaaccaacccttggcagaaacatatc
   ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++| 9350
   +
5'   catcgcgctccgccatctccagcagccgcacgcggcgcatctcggggcagcg
   ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++| 9400
   +
5'   ttgggtcctgggccacgggtgcgcatgatcgctgctcctgtcggtgaggacc
   ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++| 9450
   +
5'   cggctaggctggcggggttgcccttactgggttagcagaatgaatcaccgat
   ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++| 9500
   +
5'   acgcgagcgaacgtgaagcgactgctgctgcaaaaacgtctgcgacctgag
   ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++| 9550
   +
5'   caacaacatgaatggtcttcggtttccgtgtttcgtaaaagtctggaaaacg
   ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++| 9600
   +
5'   cggaagtcccctacgtgctgctgaagttgcccgcacagagagtggaaacc
   ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++| 9650
   +
5'   aaccgggtgataccacgatactatgactgagagtcaacgccatgagcggcc
   ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++| 9700
   +
5'   tcattttcttattctgagttacaacagtcgcgcaccgctgtccggtagctcc
   ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++| 9750
   +
5'   ttccgggtgggcgcgggggcatgactatcgtcgccgcacttatgactgtctt
   ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++| 9800
   +
5'   ctttatcatgcaactcgtaggacaggtgcccggcagcgcccaacagtcccc
   ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++| 9850
   +

```

FIG. 40V

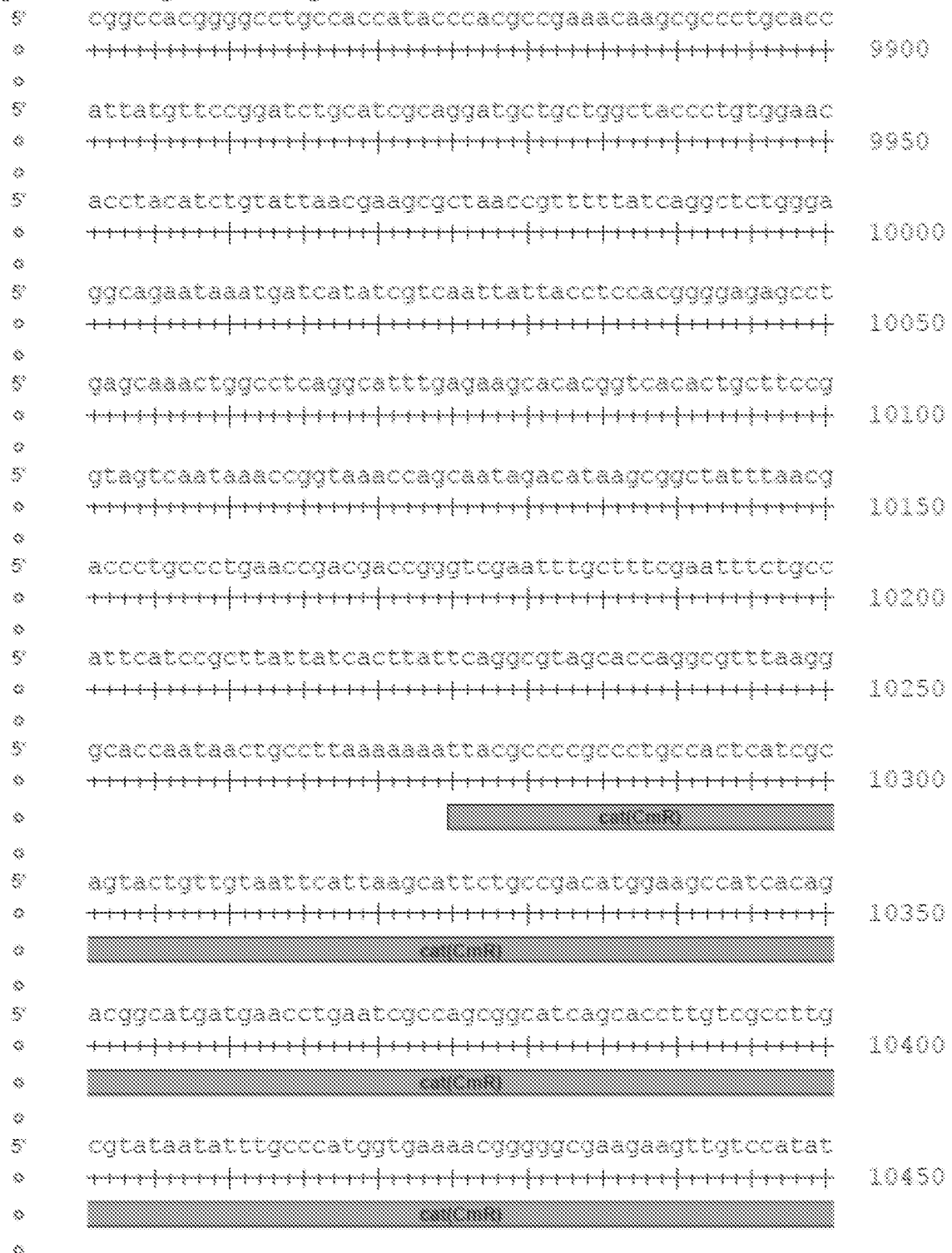
pACYC184 *FtsCpf1* locus (pY001)

FIG. 40W

pACYC184 *FnOpf1* locus (pR001)

5'	tggccacggtttaaatcaaaactgggtgaaactcaccacaggggattgggtgag	10500
+	+++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++	
	 cat(CmR)	
5'	acgaaaaacatattctcaataaaacccttttagggaaaataggccaggttttc	10550
+	+++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++	
	 cat(CmR)	
5'	accgtaacacgcccacatcttgccgaatatatgtgtagaaaactgccggaaat	10600
+	+++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++	
	 cat(CmR)	
5'	cgtcgtggtattcactccagagcgatgaaaacggtttcagtttgctcatgg	10650
+	+++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++	
	 cat(CmR)	
5'	aaaacggtgtaacaagggtgaacactatcccatatcaccagctcaccgtc	10700
+	+++++ +++++ +++++ +++++ +++++ +++++ +++++ +++++	
	 cat(CmR)	
5'	tttcattgccatacg	10715
+	+++++ +++++	
	 cat(CmR)	

FIG. 40X

1 - OpEl - NoOpEl

3' gccaccATGAACATCAAAAACCTTTACCGGGCTCTACCCCTCAGCAAAAACCTTTGGCGCTTT
 60
 A T M N I K N F T G L Y P L S K T L R F
 5' GAACTCAAGCCTATTGGCAAAACCAAGGAAAACATCGAGAAAAATGGCATCCTGACCAAG
 120
 E L K P I G K T K E N I E K N G I L T K
 5' GACGAGCAACGGGCTAAAGACTACCTCATAGTCAAAGGCTTTATTGACGAGTATCACAAAG
 180
 D E Q R A K D Y L I V K G F I D E Y H K
 5' CAGTTCATCAAAGACAGGCTTTGGGACTTTAAATTGCCCTCTCGAAAGTGAGGGGGAGAAG
 240
 Q F I K D R L W D F K L P L E S E G E K
 5' AACAGTCTCGAAGAATACCAGGAAGTGTACGAGCTCACTAAGCGCAACGATGCCCAGGAG
 300
 N S L E E Y Q E L Y E L T K R N D A Q E
 5' GCCGACTTCACCGAGATTAAAGATAACCTTCGCAGCTCTATTACCGAACAGCTCACGAAG
 360
 A D F T E I K D N L R S S I T E Q L T K
 5' TCTGGATCTCGGTACGATCGGATTTTAAAAAAGAGTTCATTAGAGAGACCTGGTCAAC
 420
 S G S A Y D R I F K K E F I R E D L V N
 5' TTCCTCGAAGATGAAAAAGATAAAAATATCGTGAAACAGTTCGAGGACTTTACTACATAT
 480
 F L E D E K D K N I V K Q F E D F T T Y

FIG. 41A

X - Cpf1 - PaCpf1

5' TTTACGGGTTTTTATGAAAATAGGAAGAACATGTACTCTAGCGAAGAGAAAGTCCACGGCC
 540
 F T G F Y E N R K N M Y S S E E K S T A
 5' ATCGCATACCGGCTTATCCATCAGAATCTGCCAAAATTCATGGACAACATGAGAAGTTTT
 600
 I A Y R L I H Q N L P K F M D N M R S F
 5' GCCAAAATTGCAAATTCAGTGTTCGAGGCACTTTAGCGACATCTATGAAAGCTGGAAG
 660
 A K I A N S S V S E H F S D I Y E S W K
 5' GAATATCTGAATGTAAATAGCATCGAGGAAATCTTCCAGCTCGACTATTTTAGCGAAACC
 720
 E Y L N V N S I E E I F Q L D Y F S E T
 5' TTGACTCAGCCACATATTGAGGTGTATAACTATATTATCGGGGAAGAAAGTCCCTGGAAGAC
 780
 L T Q P H I E V Y N Y I I G K K V L E D
 5' GGAAACCGAGATAAAGGGCATCAACGAGTATGTGAACCTCTACAATCAGCAGCAGAAAGAT
 840
 G T E I K G I N E Y V N L Y N Q Q Q K D
 5' AAGASTAAACGACTGCCTTTCCTGGTGCCACTGTATAAGCAAATTTTGTCTGATAGGGAA
 900
 K S K R L P F L V P L Y K Q I L S D R E
 5' AAACCTCTCCTGGATTGCTGAAGAGTTGACAGCGACAAGAAGATGCTGAGCGCTATCACC
 960
 K L S W I A E E F D S D K K M L S A I T

FIG. 41B

I - Cpfl - Pacpfl

```

5' GAGTCTTACAACCACTGACACAACGTGTTGATGGGTAAACGAGAACGAAAGCCTGCGAAAT
+-----+
E S Y N H L H N V L M G N E N E S L R N
1020

5' CTGCTGCTGAATATTAAAGGACTATACCTGAGAGAAATTAATATCACAAACGACTTGTCT
+-----+
L L L N I K D Y N L E K I N I T N D L S
1080

5' CTCACCGAAATCTCCAGAAATCTTTTGGCCGATATGATGTATTACAAATGGGATCAAA
+-----+
L T E I S Q N L F G R Y D V F T N G I K
1140

5' AACAAAGCTGAGAGTGTGACTCCAAGGAAGAAAAGGAGACGGACGAAATTTTGAGGAC
+-----+
N K L R V L T P R K K K E T D E N F E D
1200

5' CGCATTAACAAAATTTTAAAGACCCAGAAAGTCCTTCAGCATCGCTTTTCTGAACAAGCTG
+-----+
R I N K I F K T Q K S F S I A F L N K L
1260

5' CCTCAGCCCGAAATGGAGGATGGGAAGCCCCGGAACATTGAGGACTATTTTATTACACAG
+-----+
P Q P E M E D G K P R N I E D Y F I T Q
1320

5' GGGGCGATTAAACACCAAATCTATACAGAAAGAAAGATATCTTCGCCCAAATTGAGAAATGCA
+-----+
G A I N T K S I Q K E D I F A Q I E N A
1380

5' TACGAGGATGCACAGGTGTTCTGCAAAATTAAGGACACCGACAACTTAGCCAGAAC
+-----+
Y E D A Q V F L Q I K D T D N K L S Q N
1440

```

FIG. 41C

~~SECRET~~

5' AAGACGGCGGTGGAAAAAGATCAAAAACCTTTGCTGGACGCGCTTGAAGGAACTCCACGACTTC
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1500
K T A V E K I K T L L D A L K E L Q H F
5' ATCAAACCGCTGCTGGGCTCTGSGGAGGAGAAACGAGAAAGACGAACTGTTCTACGGTTCC
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1560
I K P L L G S G E E N E K D E L F Y G S
5' TTCTGGCCATCTGGGACGAACTGGACACCATTACACCACTTTATAACAAAGTGAGAAAT
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1620
F L A I W D E L D T I T P L Y N K V R N
5' TGGCTGACCCGAAAACCATATTCAACAGAAAAAATCAAATTGAATTTGACACGCTCAG
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1680
W L T R K P Y S T E K I K L N F D N A Q
5' CTGCTGGGAGGGTGGGATGTCAATAAAGAACACGACTGTGCAGGTATCTTGTTGCGGAAA
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1740
L L G G W D V N K E H D C A G I L L R K
5' AACGATAGCTACTATCTCGGAATTATCAATAAGAAAACCAACCACATCTTTGATACGGAT
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1800
N D S Y Y L G I I N K K T N H I F D T D
5' ATTACGCCATCAGATGGCGAGTGTATGACAAAATCGACTACAAGCTCCTTCCCGGGGCG
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1860
I T P S D G E C Y D K I D Y K L L P G A
5' AACAAAATGCTTCCAAAGGTGTTTTTTAGTAAGTCCCGAATCAAAGAGTTCGAGCCATCA
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1920
N K M L P K V F F S K S R I K E F E P S

FIG. 41D

I - Opfl - PaOpfl

5' GAGGCCATAATCAATTGCTATAAGAAGGGGACACACAAAAAGGAAAAACTTTAACCTG
 ++++++
 1 E A I I N C Y K K G T H K K G K N F N L 1980

5' ACGGACTGTCAACCGCCTGATCAACTTTTTTAAGACCTCAATCGAGAAACACGAGGATTGG
 ++++++
 1 T D C N R L I N F F K T S I E K H E D W 2040

5' TCAAAATTCGGATTCAAGTTCTCCGATACCGAAACGTATGAGGATATTAGCGGTTTTFIAT
 ++++++
 1 S K F G F M F S D T E T Y E D I S G F Y 2100

5' AGAGAGGTCGAGCAGCAGGGATACAGGCTGAACGAGCCATCCAGTCAGTGCCAGCTATATA
 ++++++
 1 R E V E Q Q G Y R L T S H P V S A S Y I 2160

5' CATAGTCTGGTCAAGGAAGGAAAACTGTACCTCTTCCAAATCTGGAACRAGGACTTTTCT
 ++++++
 1 H S L Y K E G K L Y L F Q I W N K D F S 2220

5' CAATTCTCCAAGGGGACCCCTAACTTGCACACTCTCTATTGGAAGATGCTGTTTGACAAA
 ++++++
 1 Q F G K G T P N L H T L Y W K M L F D K 2280

5' CGGAATCTTAGCGATGTGGTTTATAAGCTGAATGGCCAGGCTGAAGTGTCTATAGAAAAG
 ++++++
 1 R N L S D V V Y K L N G Q A E V F Y R K 2340

5' AGCTCCATTTGAACACCAGAACCGAATTATCCACCCCGCTCAGCATCCCATCACAATAAG
 ++++++
 1 S S I E H Q N R I I H P A Q H P I T N K 2400

FIG. 41E

I ~ Cpfl ~ PaCpfl

```

5' AATGAGCTTAAACAAAAGCACACTAGCACCTTCAAATACGATATCATCAAGATCGCAGA
+-----+
2460
N E L N K K H T S T F K Y D I I K D R R

5' TACACGGTGGATAAAATTCAGTTCATGTGCCCATTAATAAAATTTTAAGGCGACCGGG
+-----+
2520
Y T V D K F Q F H V P I T I N F K A T G

5' CAGAACAACATCAACCCATCGTCCAAGAGGTGATTTCGCCAAPACGGTATCACCCACATC
+-----+
2580
Q N N I N P I V Q E V I R Q N G I T H I

5' ATAGGCATCGATCGAGGTGAACGCCATCTTCTGTACCTCTCTCTCATCGATTTGAAAGGC
+-----+
2640
I G I D R G E R H L L Y L S L I D L K G

5' AACATCATCAAGCAGATGACTCTCAACGAAATTATTAATGAGTATAAGGGTGTGACCTAT
+-----+
2700
N I I K Q M T L N E I I N E Y K G V T Y

5' AAGACCAACTACCATAACCTCCTGGAGAAGAGGGAGAAGGAGCGGACCGAGCCCGACAC
+-----+
2760
K T N Y H N L L E K R E K E R T E A R H

5' TCCTGGAGTAGTATTGAAAGCATAAAAGAACTGAAGGATGGATACATGTCACAGGTGATT
+-----+
2820
S W S S I E S I K E L K D G Y M S Q V I

5' CACAAAATTACGGACATGATGGTTAAGTACAATGCGATTGTGGTCTCTGGAGGACCTCAAC
+-----+
2880
H K I T D M M V K Y N A I V V L E D L N

```

FIG. 41F

I - Cpf1 - PaCpf1

5' GGGGGGTTTATGCCGAGGCGCCAGAAAGGTCGAGAAGCAGGTGTACCAGAAATTTGAAAA
 ++++++ 2940
 G G F M R G R Q K V E K Q V Y Q K F E K
 5' AAGTTGATCGACAAGCTGAACTATCTCGTTGACAAGAACTCGACGCTAACGAGGTGGGC
 ++++++ 3000
 K L I D K L N Y L V D K K L D A N E V G
 5' GGAGTACTGAATGCTTATCAGCTGACCAACAAGTTCGAGTCTTTCAAGAAGATTGGGAAA
 ++++++ 3060
 G V L N A Y Q L T N K F E S F K K I G K
 5' CAAAGCGGATTTTGTCTACATCCCGCGCTGGAACACAAAGCAAAATCGATCCTATAACA
 ++++++ 3120
 Q S G F L F Y I P A W N T S K I D P I T
 5' GGGTTCGTTAATCTGTTCAACACCAGGTACGAGTCTATCAAGGAGACAAAAGTTTTTTGG
 ++++++ 3180
 G F V N L F N T R Y E S I K E T K V F W
 5' TCTAAGTTTGATATTATCCGATACAATAAAGAGAAAGAAATTGGTTCGASTTCGTCTTCGAT
 ++++++ 3240
 S K F D I I R Y N K E K N W F E F V F D
 5' TACAATACCTTTACGACTAAAGCGGAGGGGACACGCACTAAGTGGACTCTGTGCACCCAC
 ++++++ 3300
 Y N T F T T K A E G T R T K W T L C T H
 5' GGCACCTCGCATCCAGACATTCGGGAAGCCAGAAAAGAAATGCCAGTGGGACAATAAAGAG
 ++++++ 3360
 G T R I Q T F R N P E K N A G W D N K E

FIG. 41G

I -- Cpf1 - PaCpf1

5' ATCAATTTGACTGAGTCCTTCAAAGCTCTGTTTGAAAAGTACAAGATCGATATCACCAGT 3420
 ++++++
 I N L T E S F K A L F E K Y K I D I T S
 5' AATCTCAAGGAATCCATCATGCAGGAAACCGAGAAGAAGTTCTTCCAGGAACGCATAAT 3480
 ++++++
 N L K E S I M Q E T E K K F F Q E L H N
 5' CTGCTCCACCTGACCCTGCAGATGAGGAATAGCGTTACTGGAACCGACATAGACTATTTG 3540
 ++++++
 L L H L T L Q M R N S V T G T D I D Y L
 5' ATCAGCCCCGTTGCCGATGAGGATGGAAATTTCTATGATAGTCGCATAAATGGCAAAAT 3600
 ++++++
 I S P V A D E D G N F Y D S R I N G K N
 5' TTTCCGAGAAATGCCGATGCCAATGGCCCGTACAACATCGCACGAAAGGGTCTGATGCTT 3660
 ++++++
 F P E N A D A N G A Y N I A R K G L M L
 5' ATTCCGCAGATCAAGCAAGCAGATCCACAGAAGAAATTCAAGTTTGACACAATCACC AAT 3720
 ++++++
 I R Q I K Q A D P Q K K F K F E T I T N
 5' AAAGACTGGCTGAAATTCGCCCAAGACAAGCCCTATCTTAAAGATggcagcgggAAAAGG 3780
 ++++++
 K D W L K F A Q D K P Y L K D G S G K R
 5' CCGGCGGCCACGAAAAAGGCCGCCAGGCCAAAAAAGAAAAAGggatcccTACCCATACGAT 3840
 ++++++
 P A A T K K A G Q A K K K K G S Y P Y D

FIG. 41H

$$I = \text{CPI} - \text{DeCPI}$$
[illegible]

FIG. 411

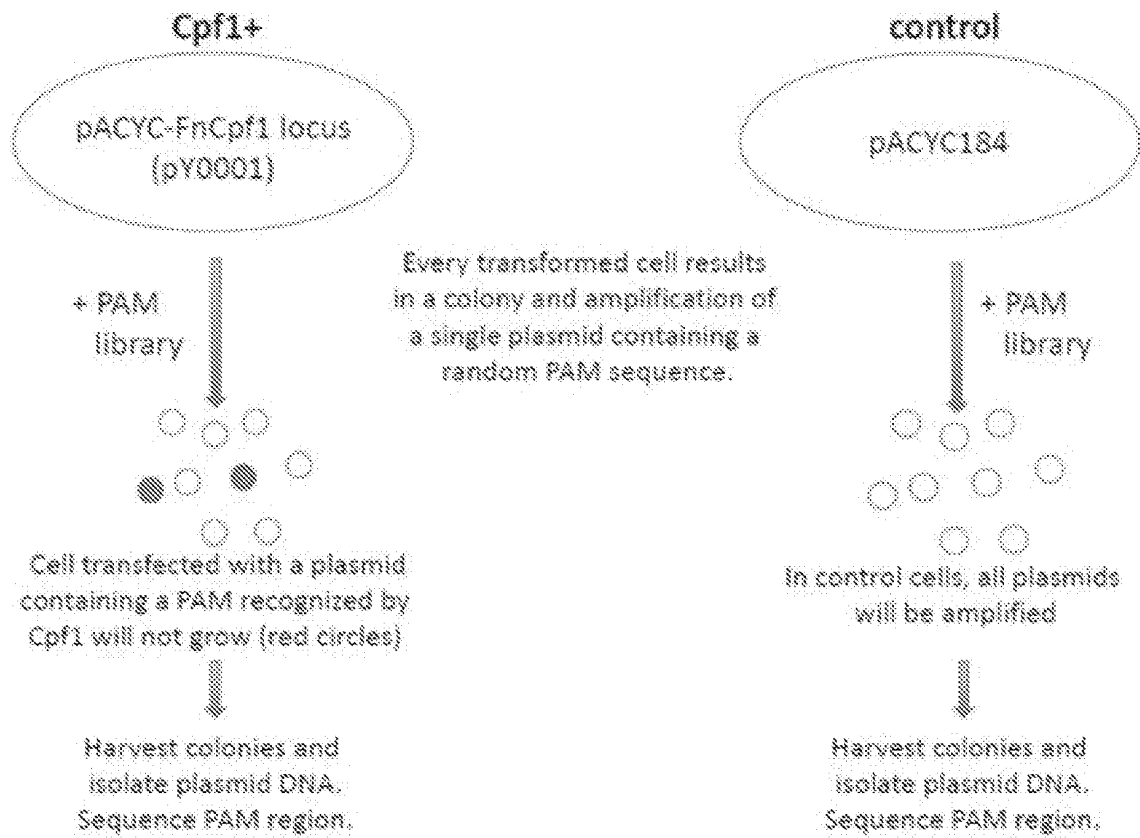
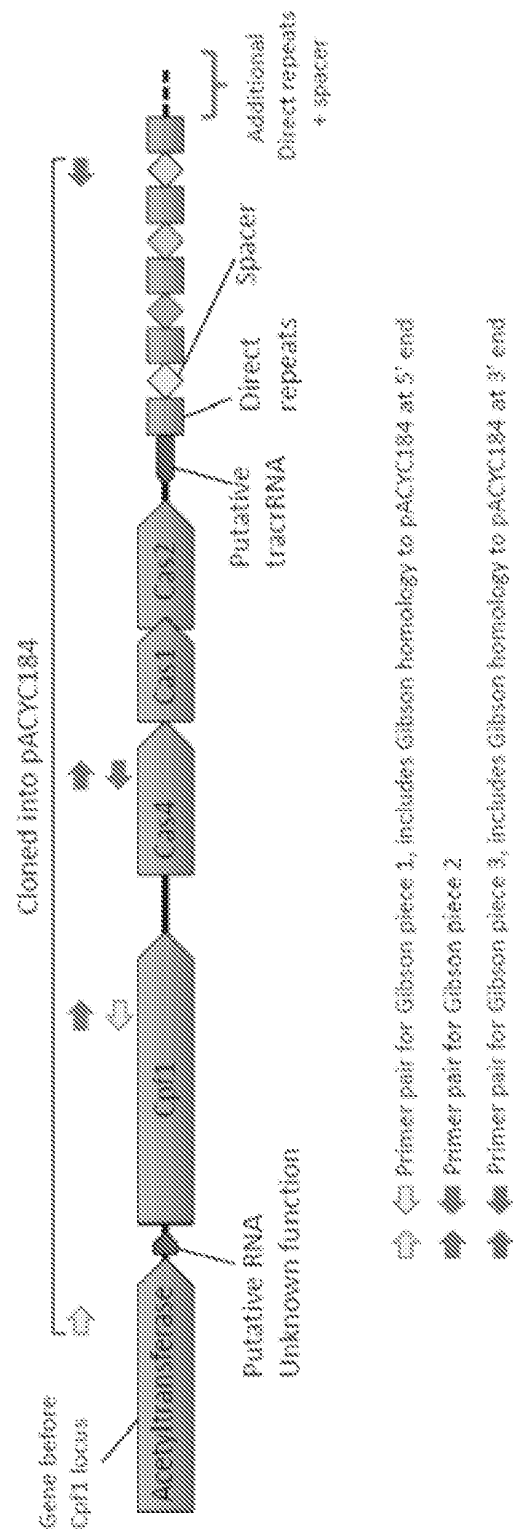


FIG. 42



43

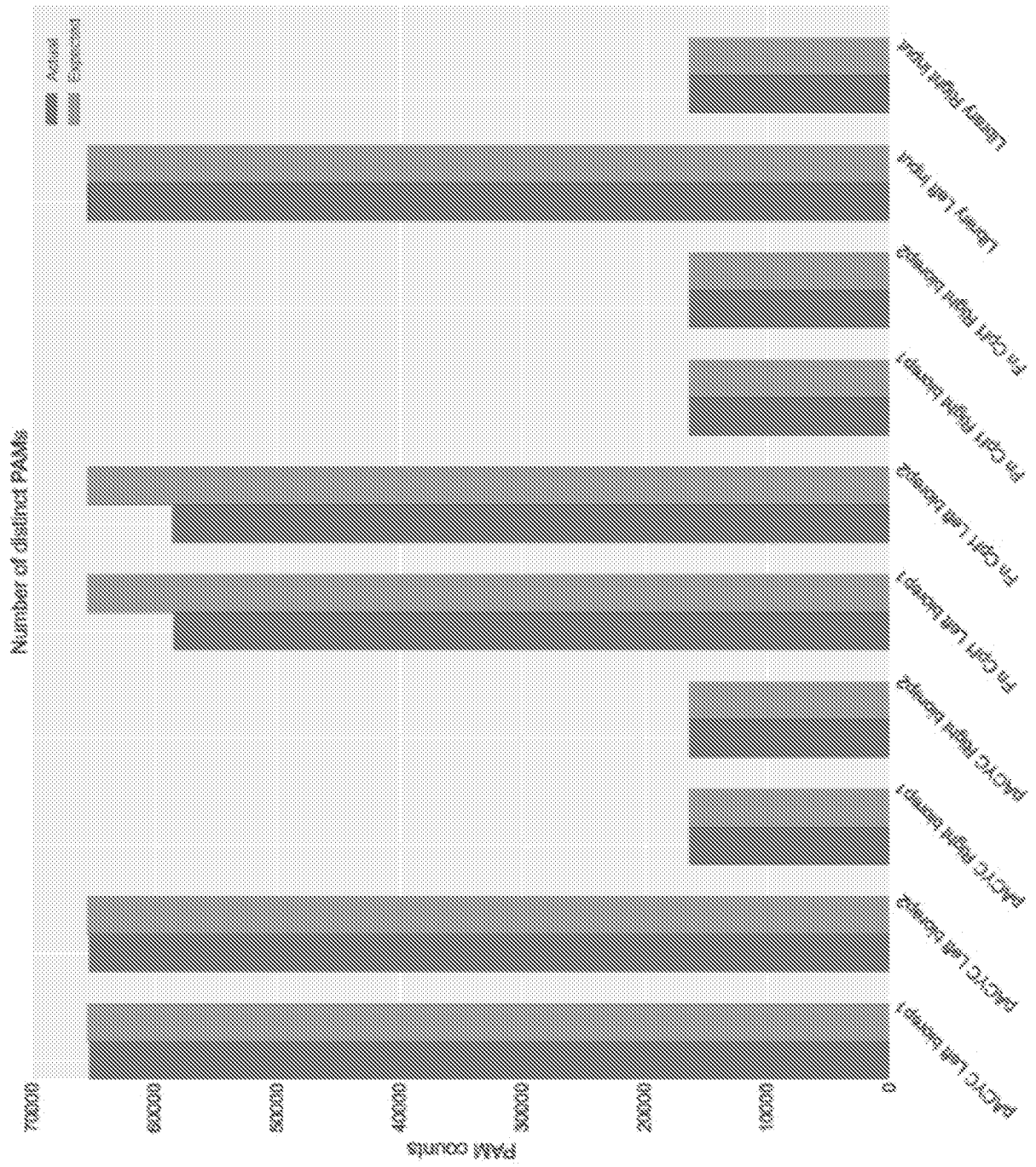


FIG. 45A

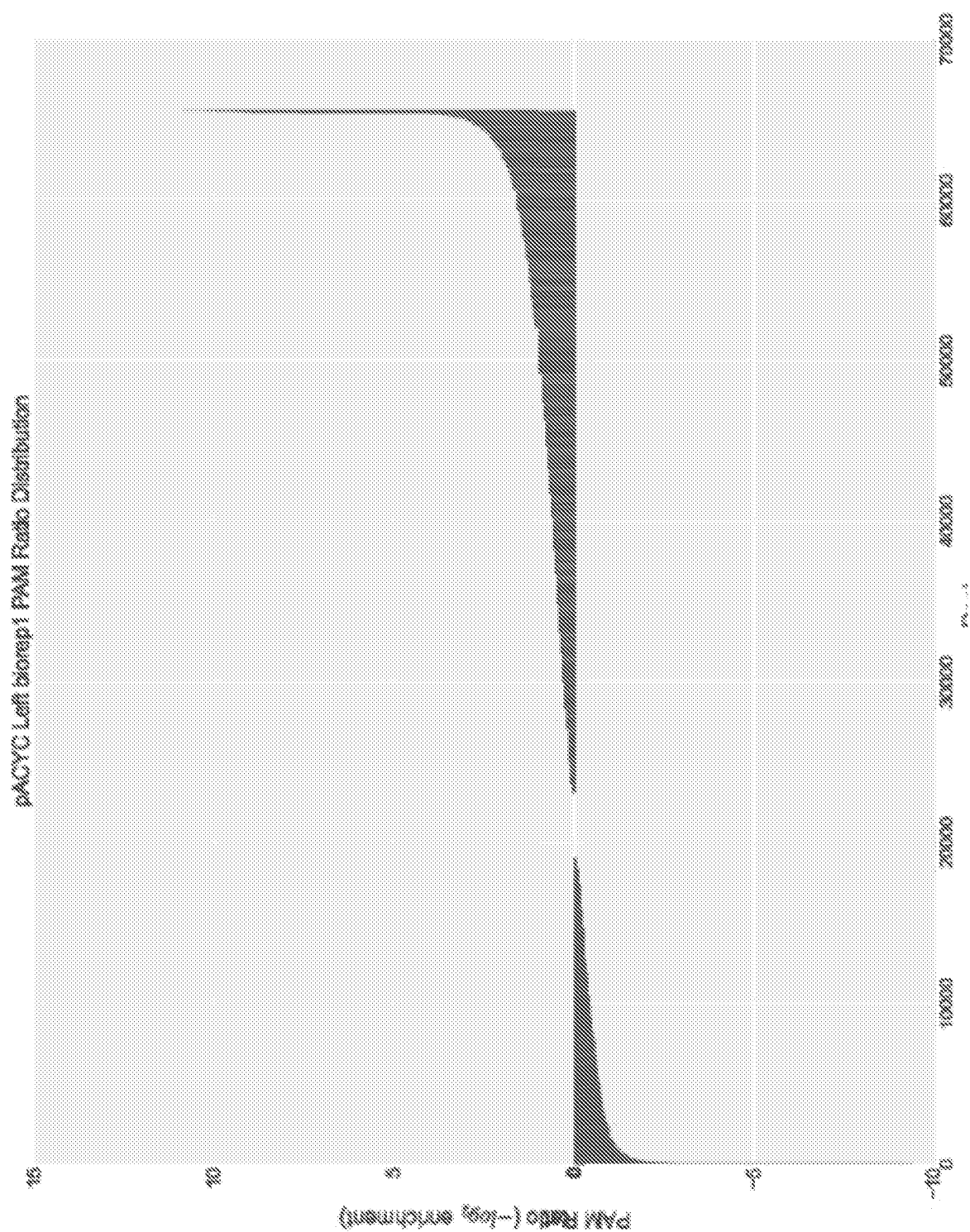


FIG. 45B

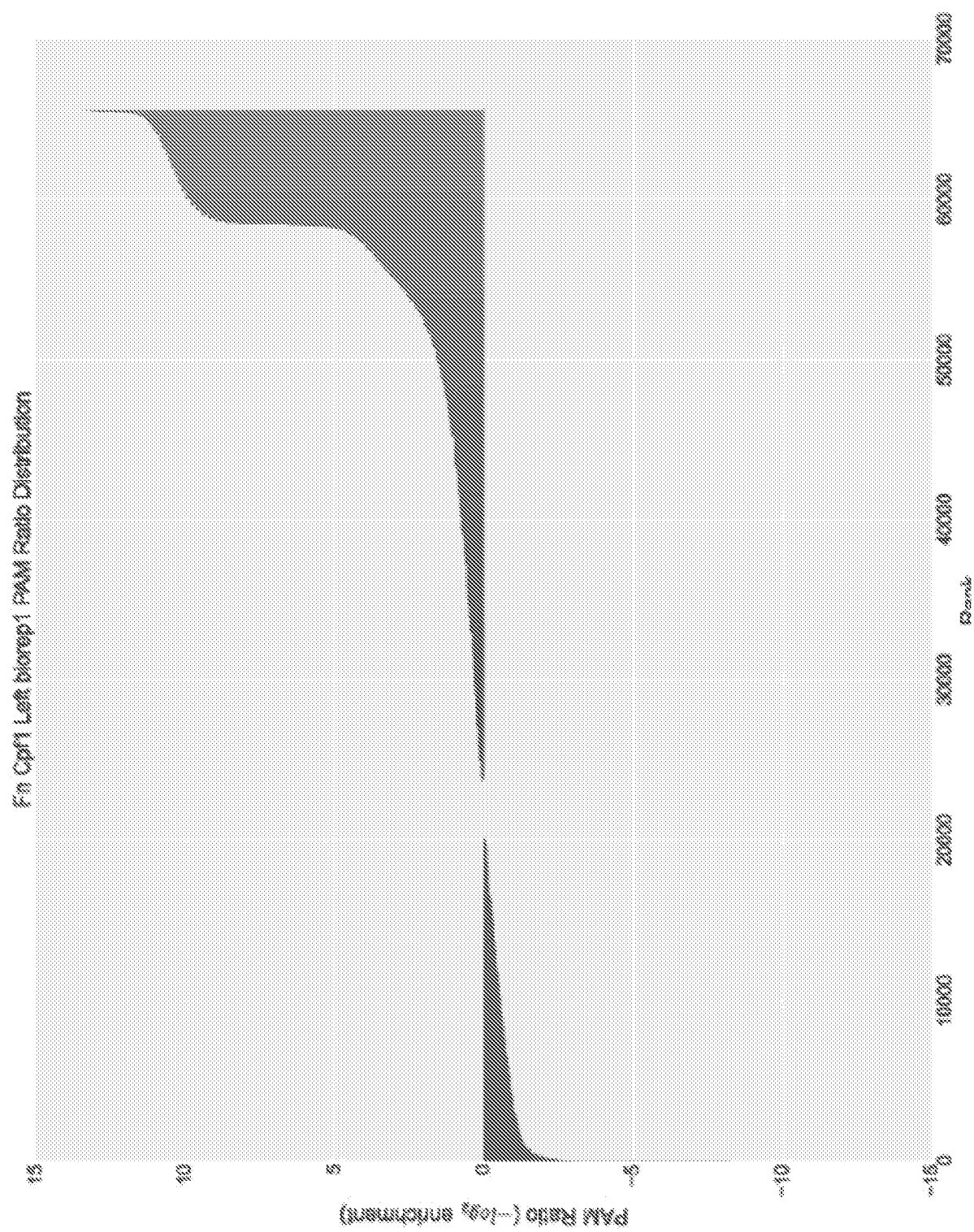
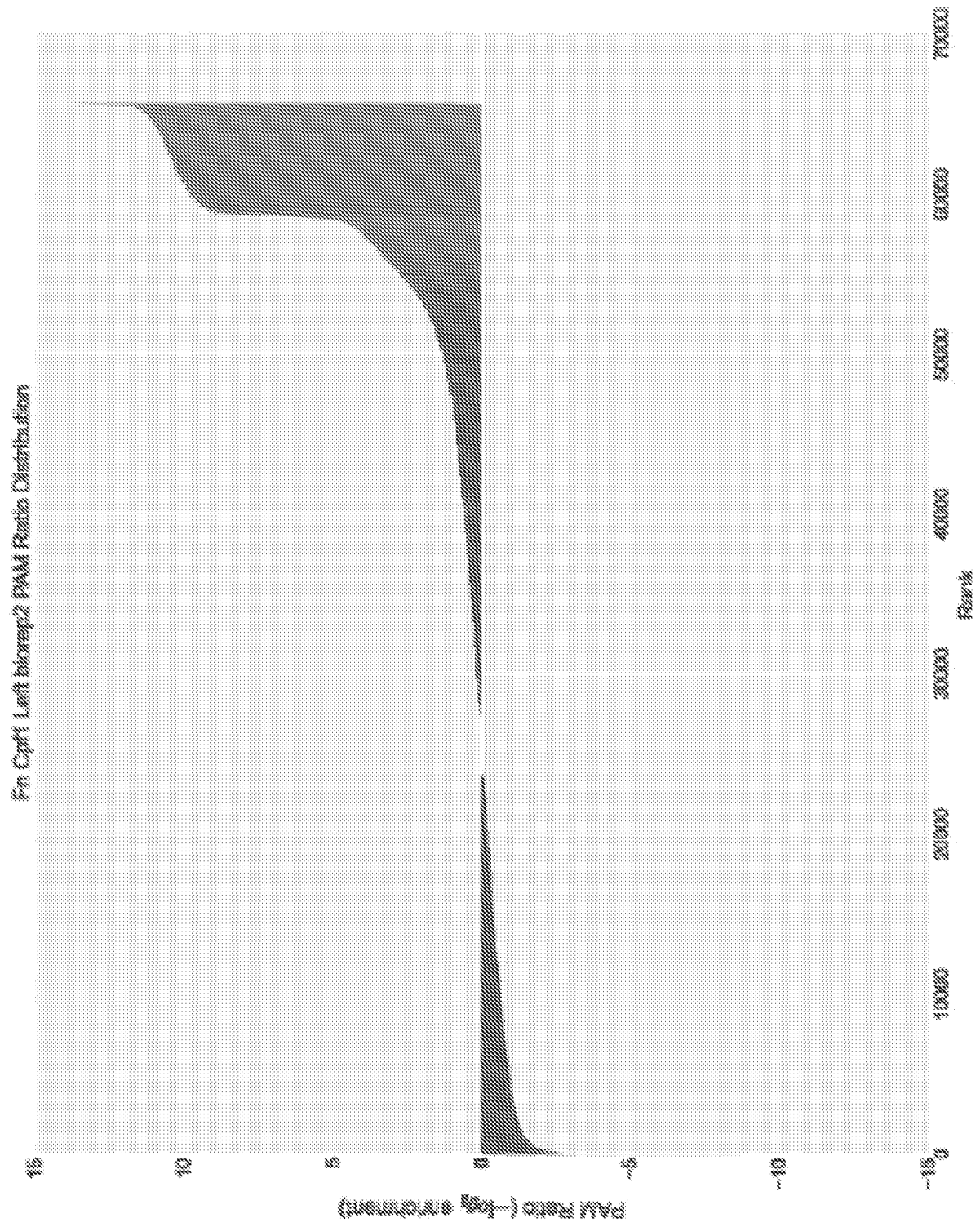


FIG. 45C



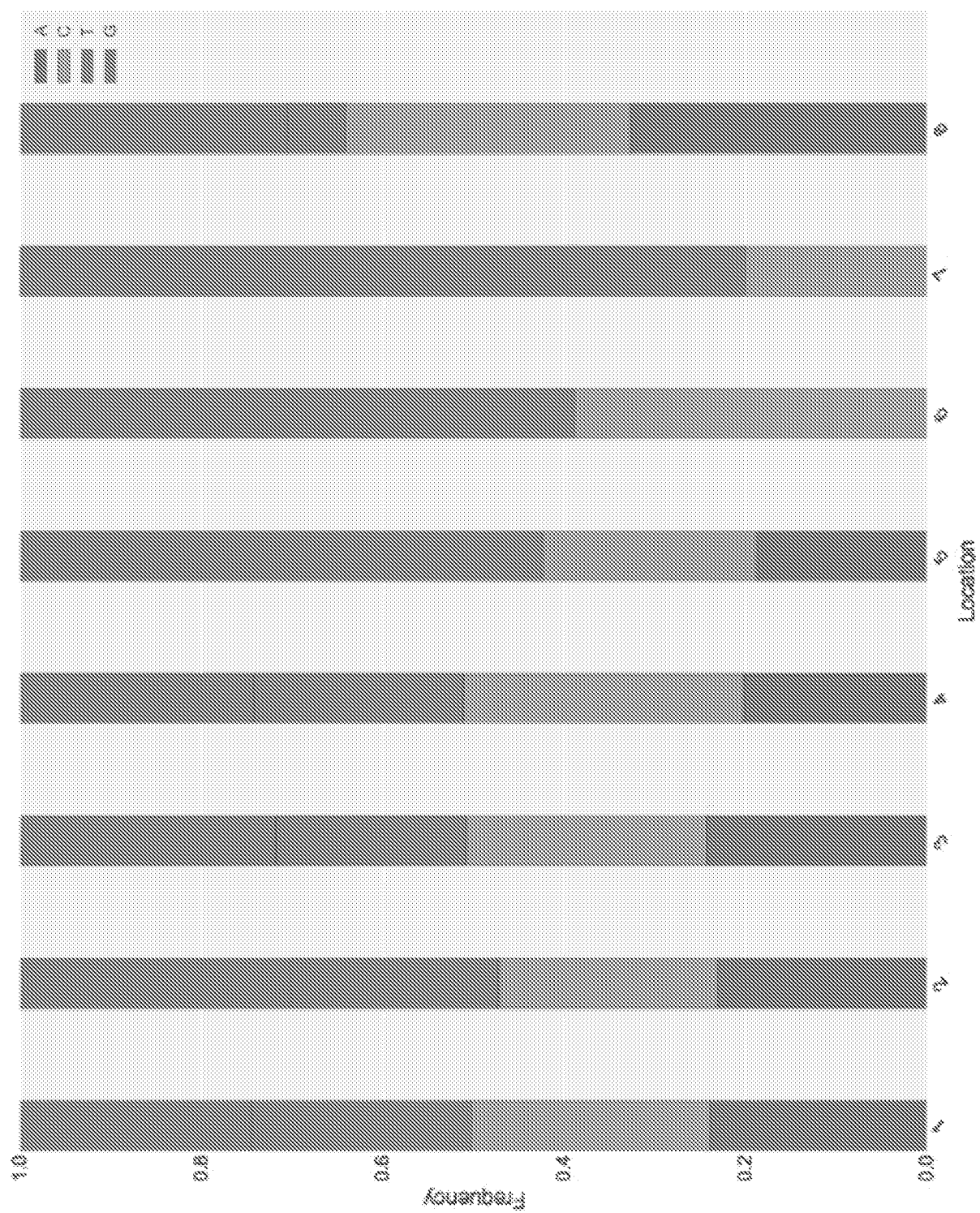


FIG. 45E

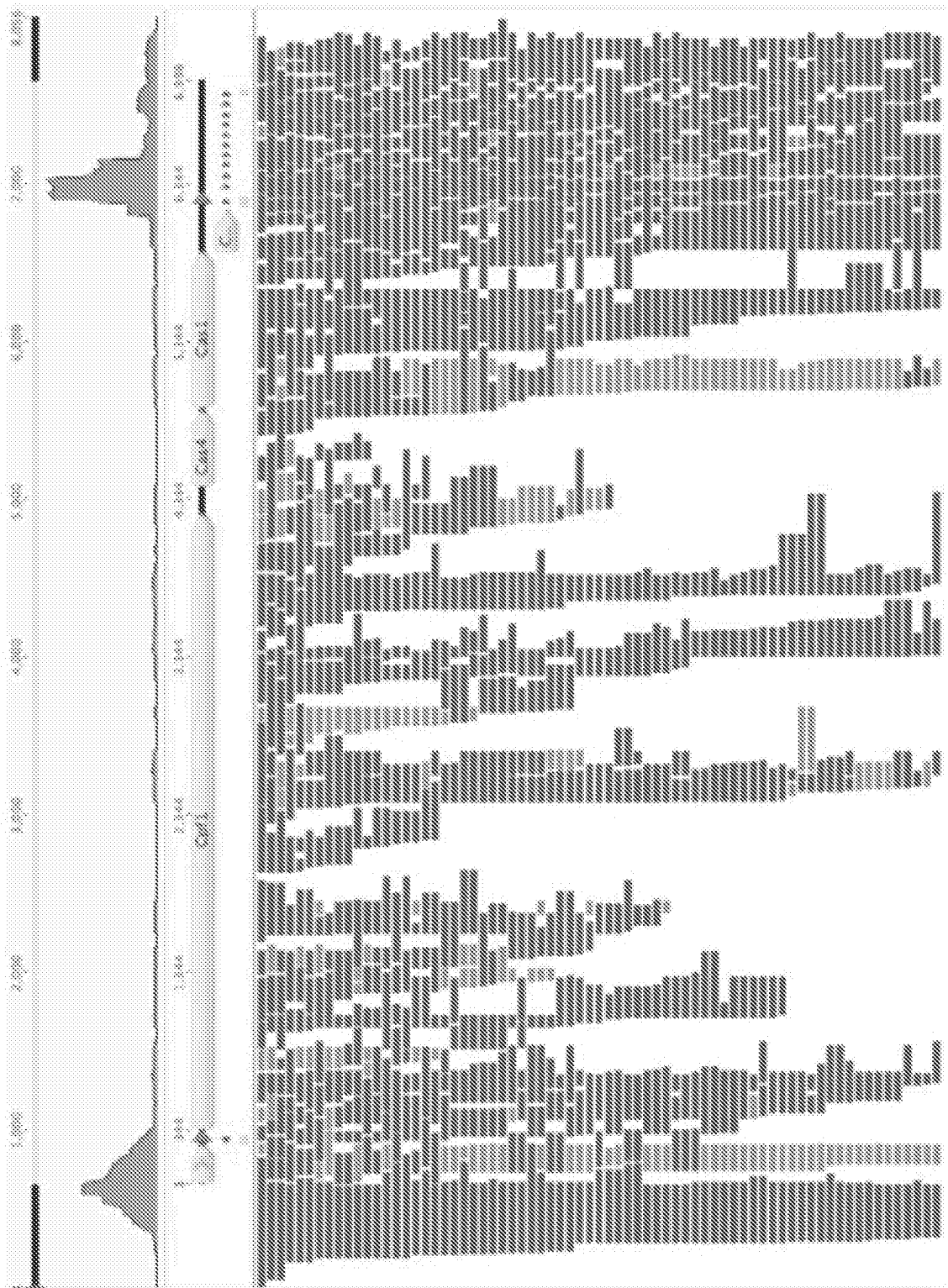


FIG. 46

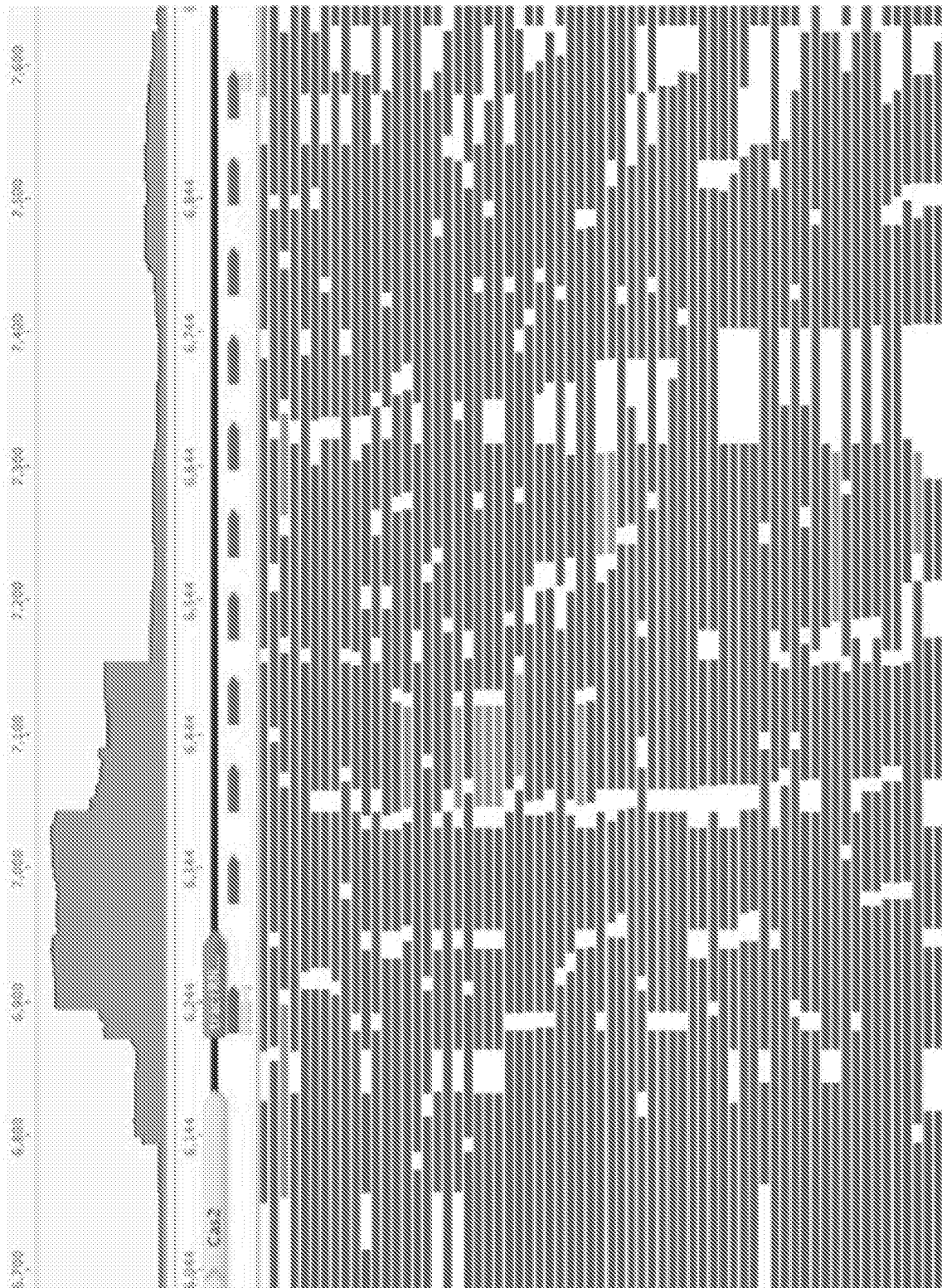


FIG. 47

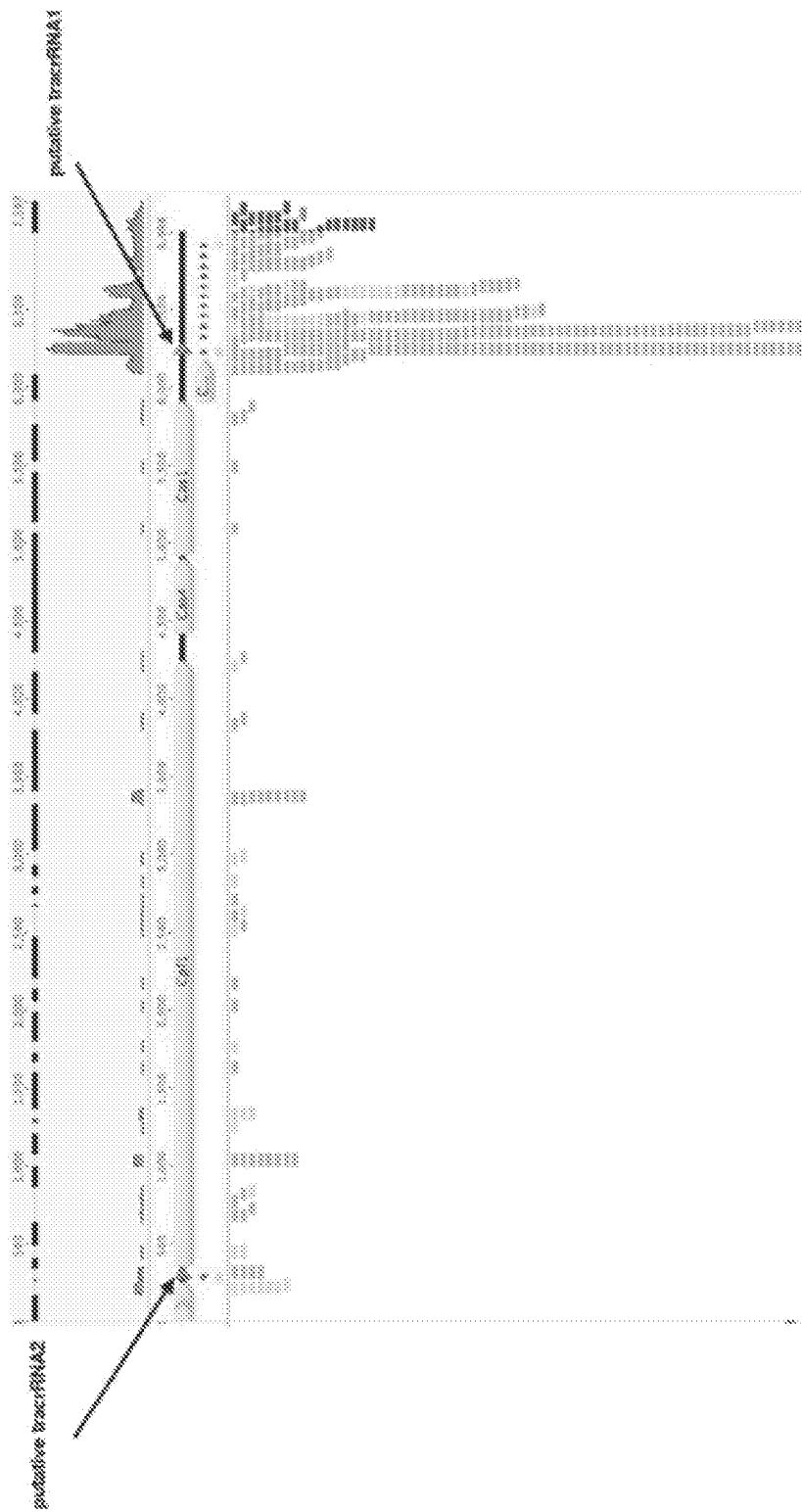


FIG. 48

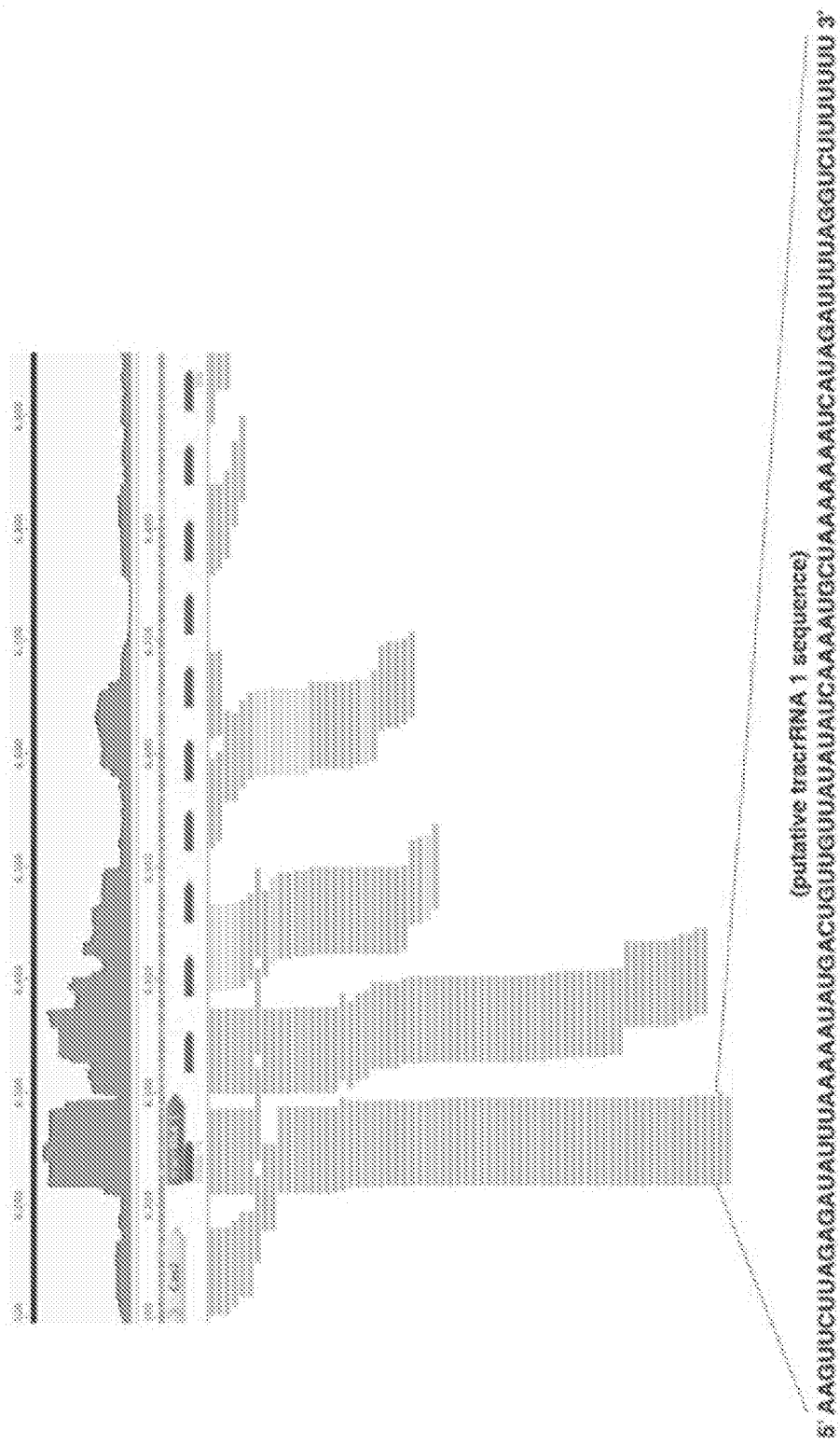


FIG. 49

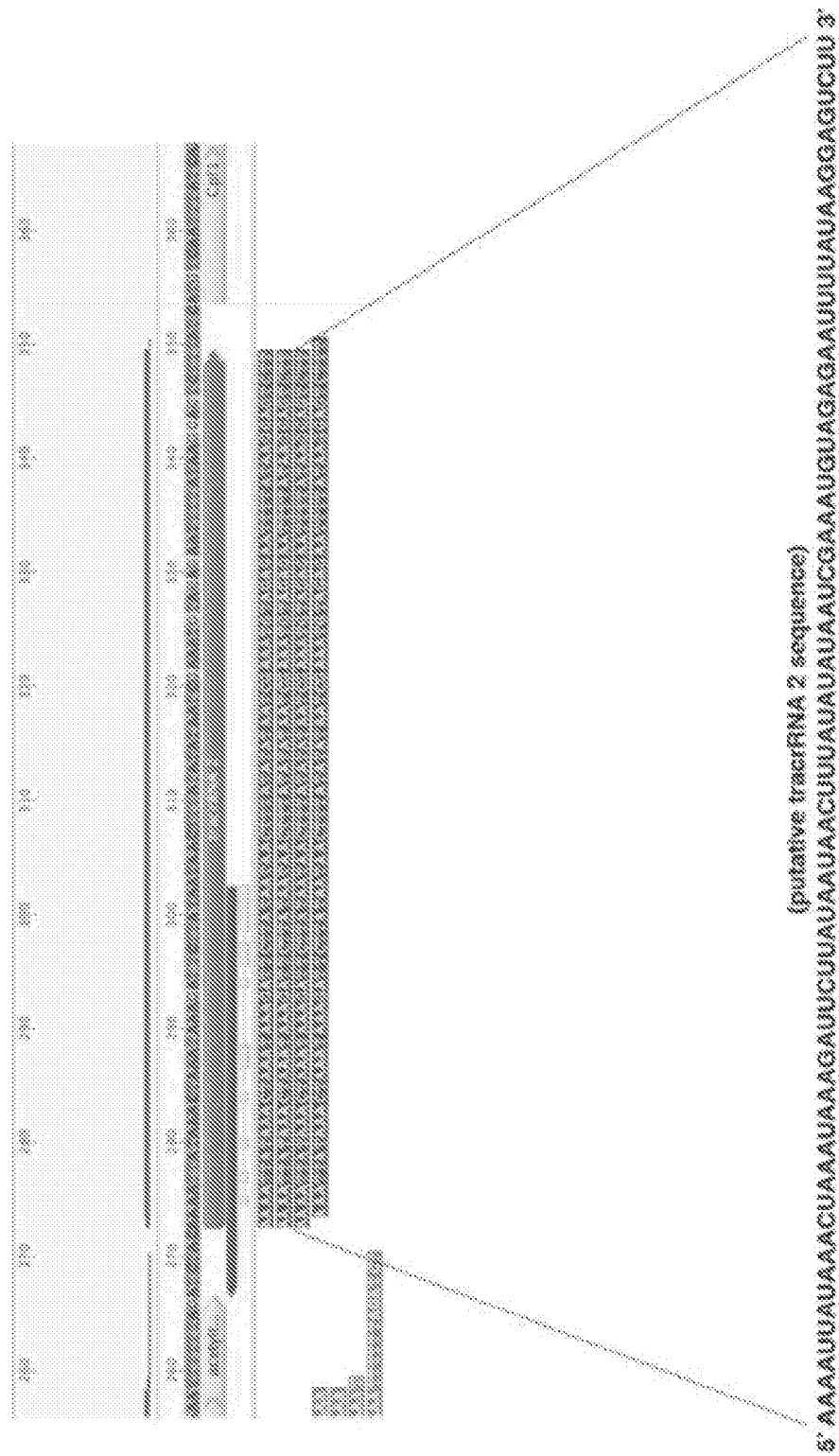


FIG. 50

Version 1: 5' ACTTTAAATAATTCTACTGTTGTAGATGAGAAAGTCATTTAAGGCCACTGTTAAAGAGTCTAAGAACTTTTAAATA 3'

Version 2: 5' CTACTATTCTGTGCTTCAGATAATTCAGTCTAAGAACTTTTAAATAATTCTACTGTTGTAGATGTTCTAGAGCCCTT 3'

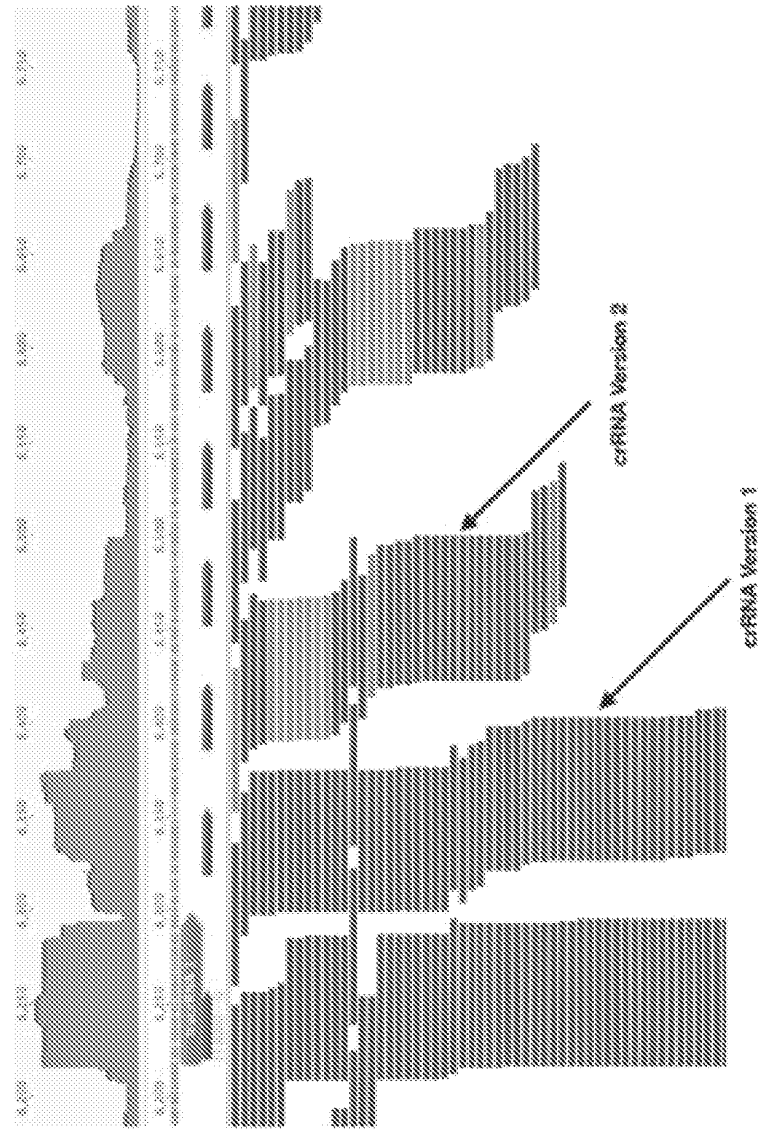


FIG. 51

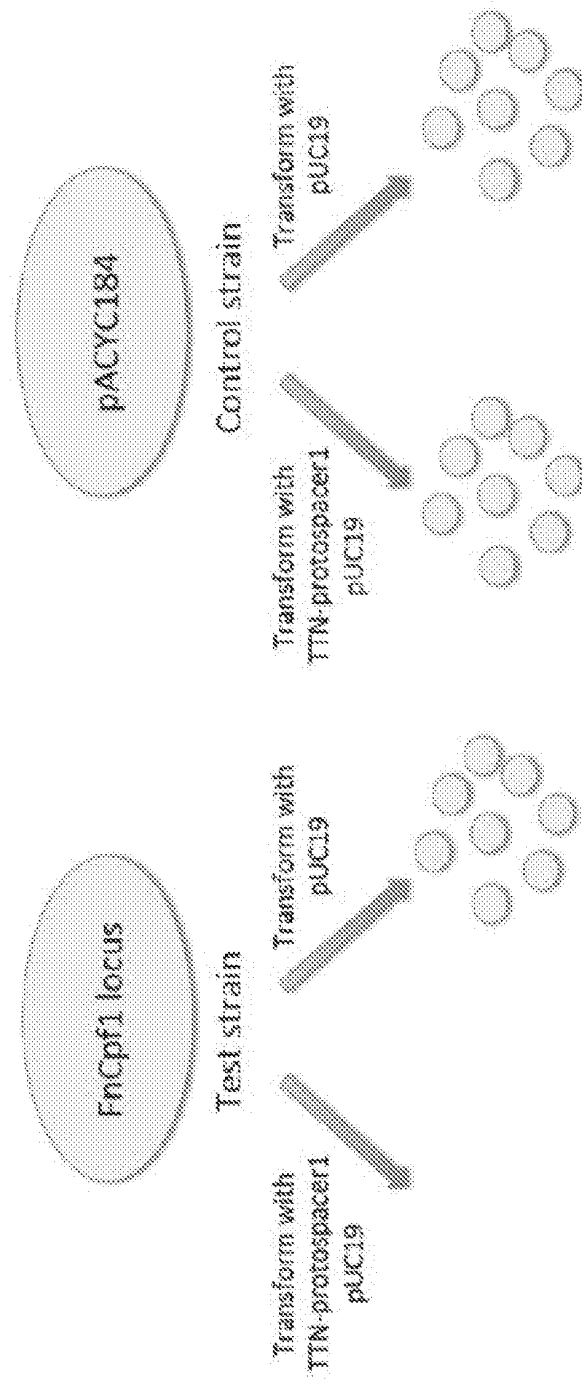


FIG. 52

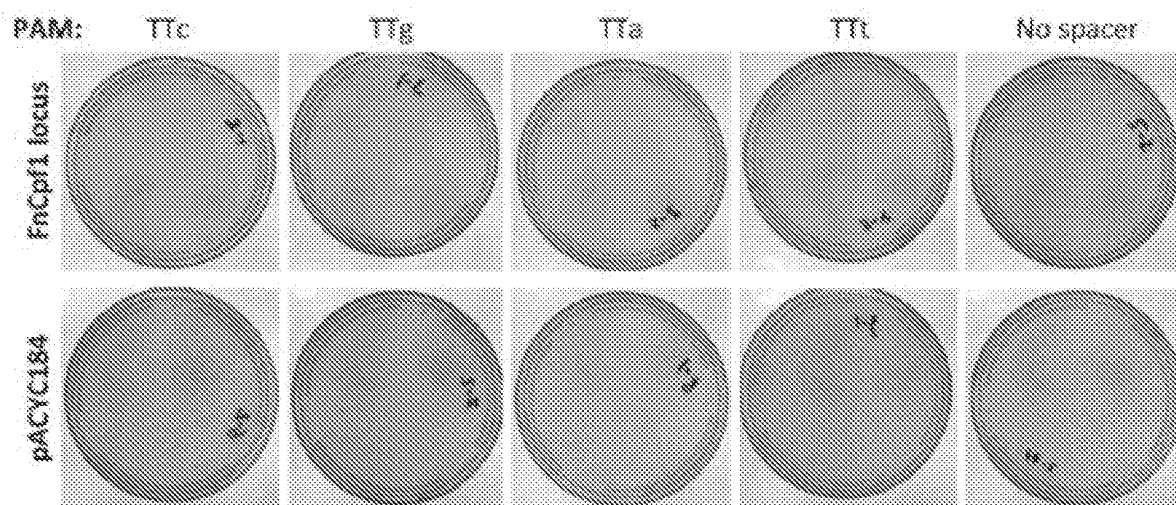


FIG. 53

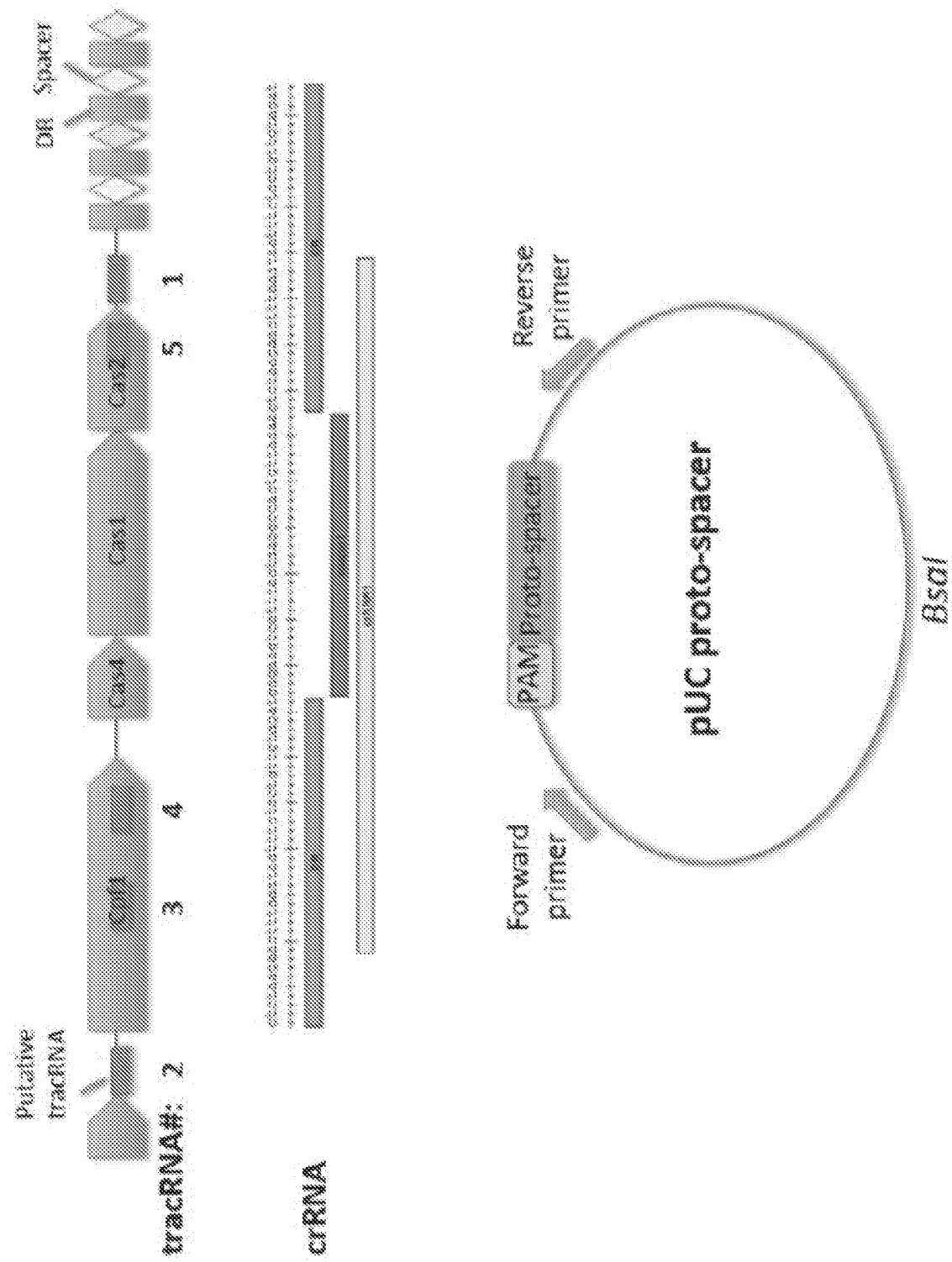


FIG. 54

FnCpf1	+	+	+	+	+	+	-	-
crRNA	+	+	+	+	+	+	+	+
tracRNA	1	2	3	4	5	-	1	2

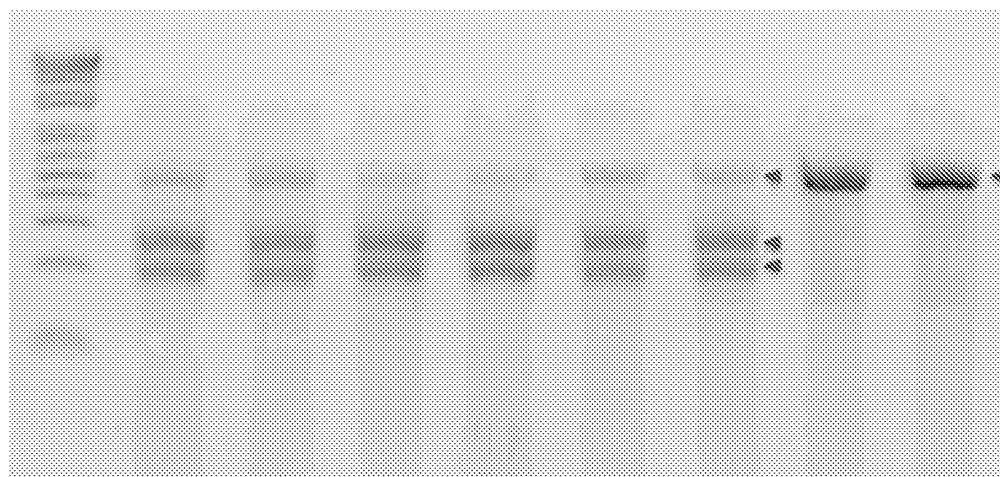


FIG. 55

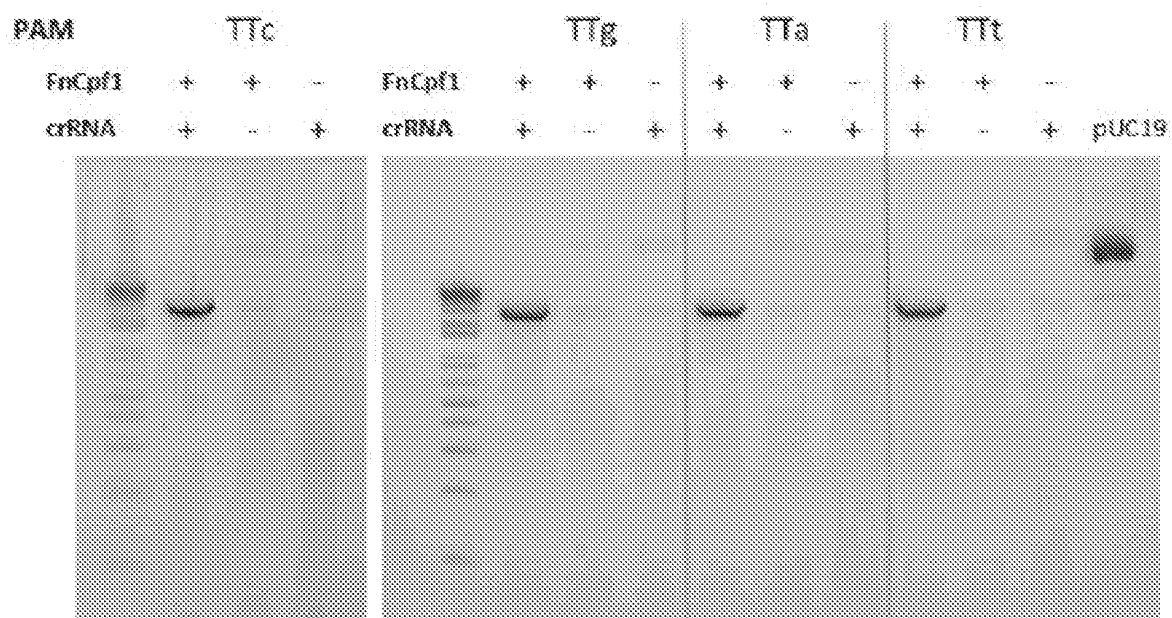


FIG. 56

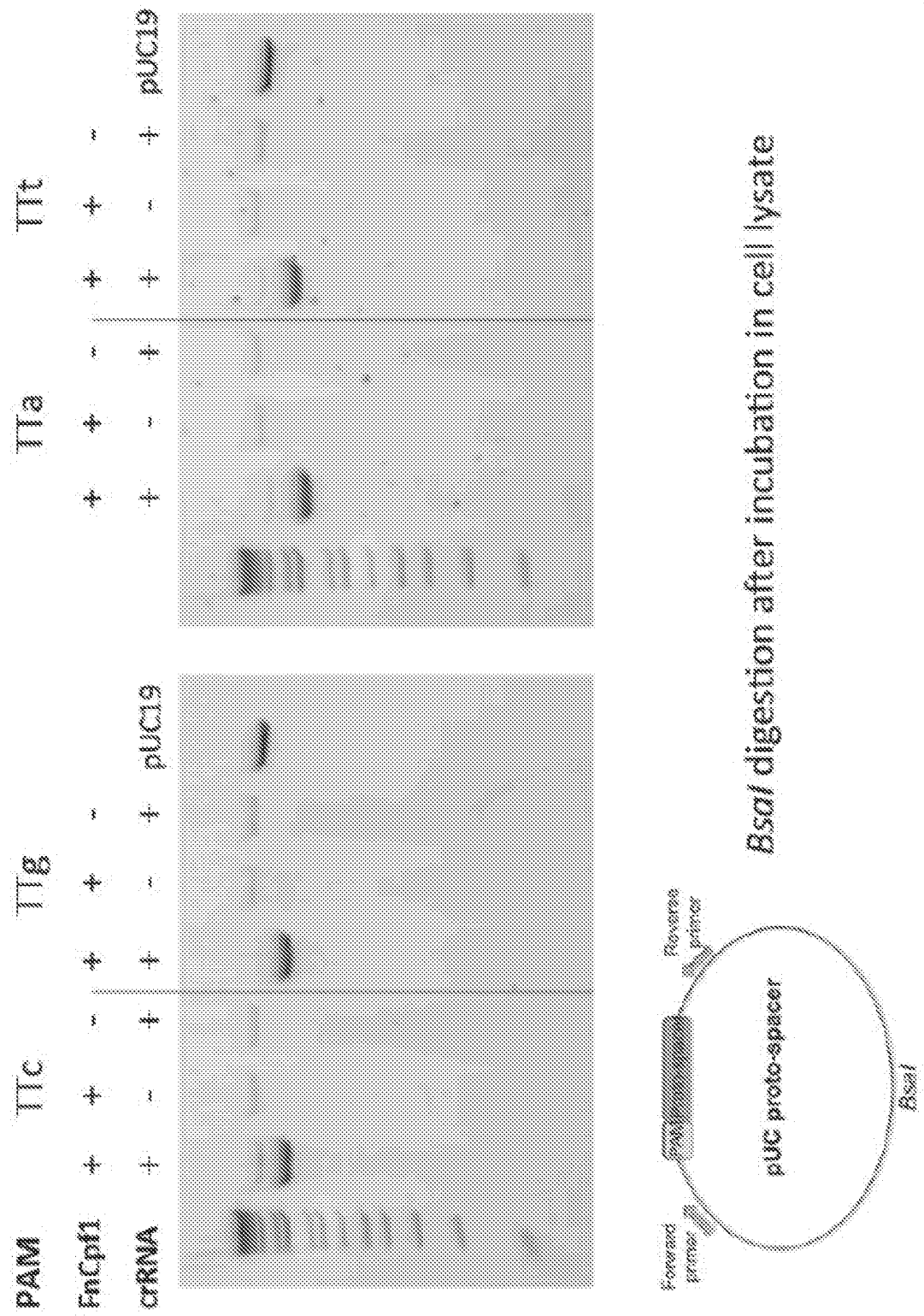


FIG. 57

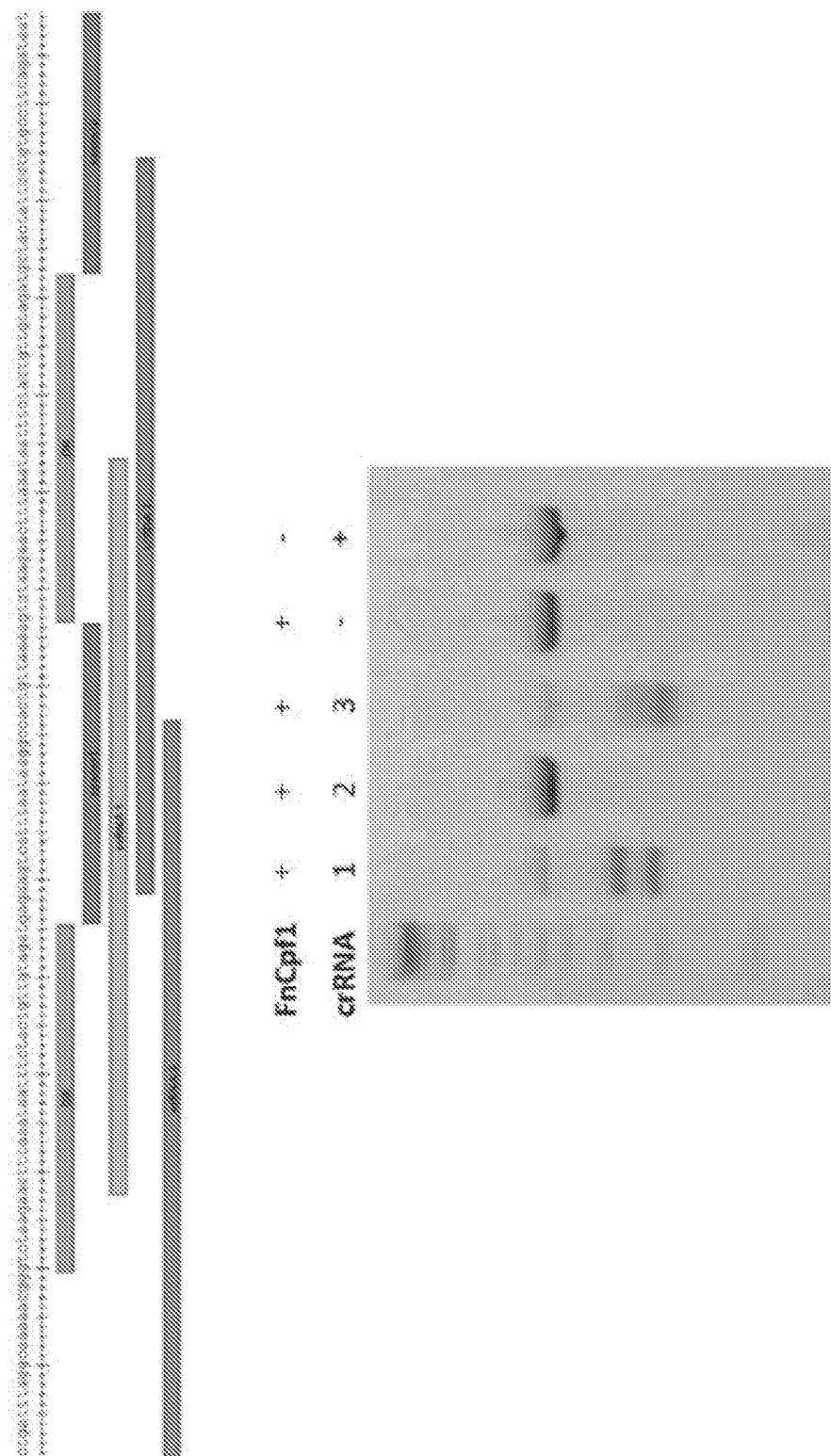
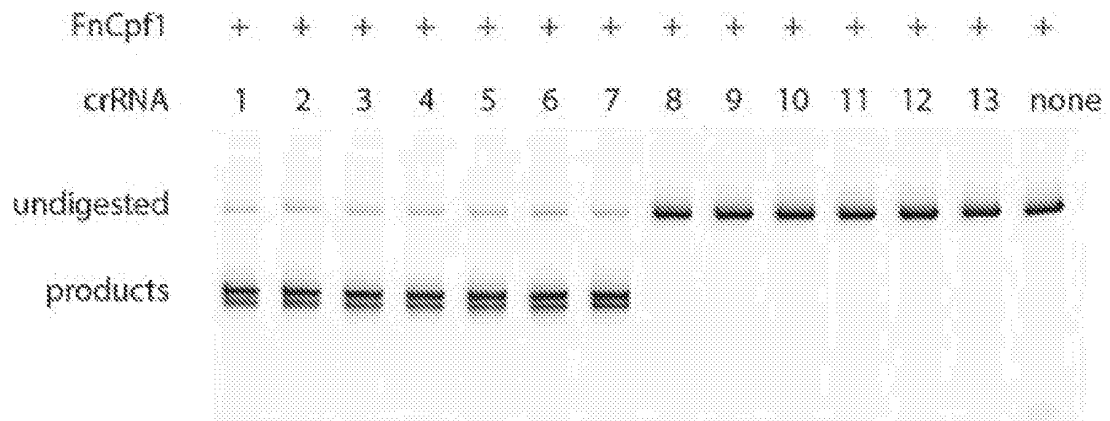


FIG. 58



crRNA			
1	acuuuaaaauaaauuuaacuguuugagau	gagaaagucuuuaaauaaaggccacuguuuaaaagcu	guuaaaggaacuuuaaaau
2	acuuuaaaauaaauuuaacuguuugagau	gagaaagucuuuaaauaaaggccacuguuuaaa	guuaaaggaacuuuaaaau
3	acuuuaaaauaaauuuaacuguuugagau	gagaaagucuuuaaauaaaggccacuguuuaaa	guuaaaggaacuuuaaaau
4	acuuuaaaauaaauuuaacuguuugagau	gagaaagucuuuaaauaaaggccacuguu	guuaaaggaacuuuaaaau
5	acuuuaaaauaaauuuaacuguuugagau	gagaaagucuuuaaauaaaggccacuu	guuaaaggaacuuuaaaau
6	acuuuaaaauaaauuuaacuguuugagau	gagaaagucuuuaaauaaaggcc	guuaaaggaacuuuaaaau
7	acuuuaaaauaaauuuaacuguuugagau	gagaaagucuuuaaauaaagg	guuaaaggaacuuuaaaau
8	acuuuaaaauaaauuuaacuguuugagau	gagaaagacuuuaa	guuaaaggaacuuuaaaau
9	acuuuaaaauaaauuuaacuguuugagau	aguuuaaauaaaggccacuguuuaaa	guuaaaggaacuuuaaaau
10	acuuuaaaauaaauuuaacuguuugagau	cauuuaaaggaaggccacuguuuaaa	guuaaaggaacuuuaaaau
11	acuuuaaaauaaauuuaacuguuugagau	uaauaaaggccacuguuuaaa	guuaaaggaacuuuaaaau
12	acuuuaaaauaaauuuaacuguuugagau	uaaggccacuguuuaaa	guuaaaggaacuuuaaaau
13	acuuuaaaauaaauuuaacuguuugagau	gucuaaaggaacuuuaaa	guuaaaggaacuuuaaaau

FIG. 59

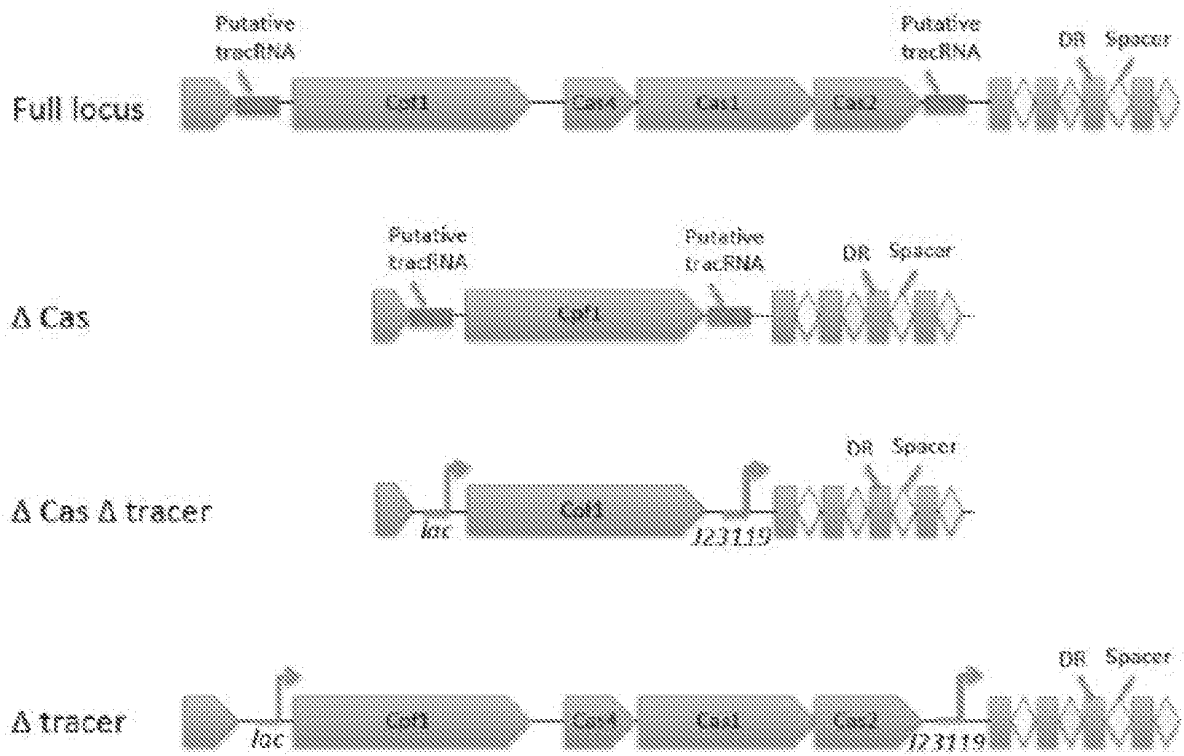


FIG. 60

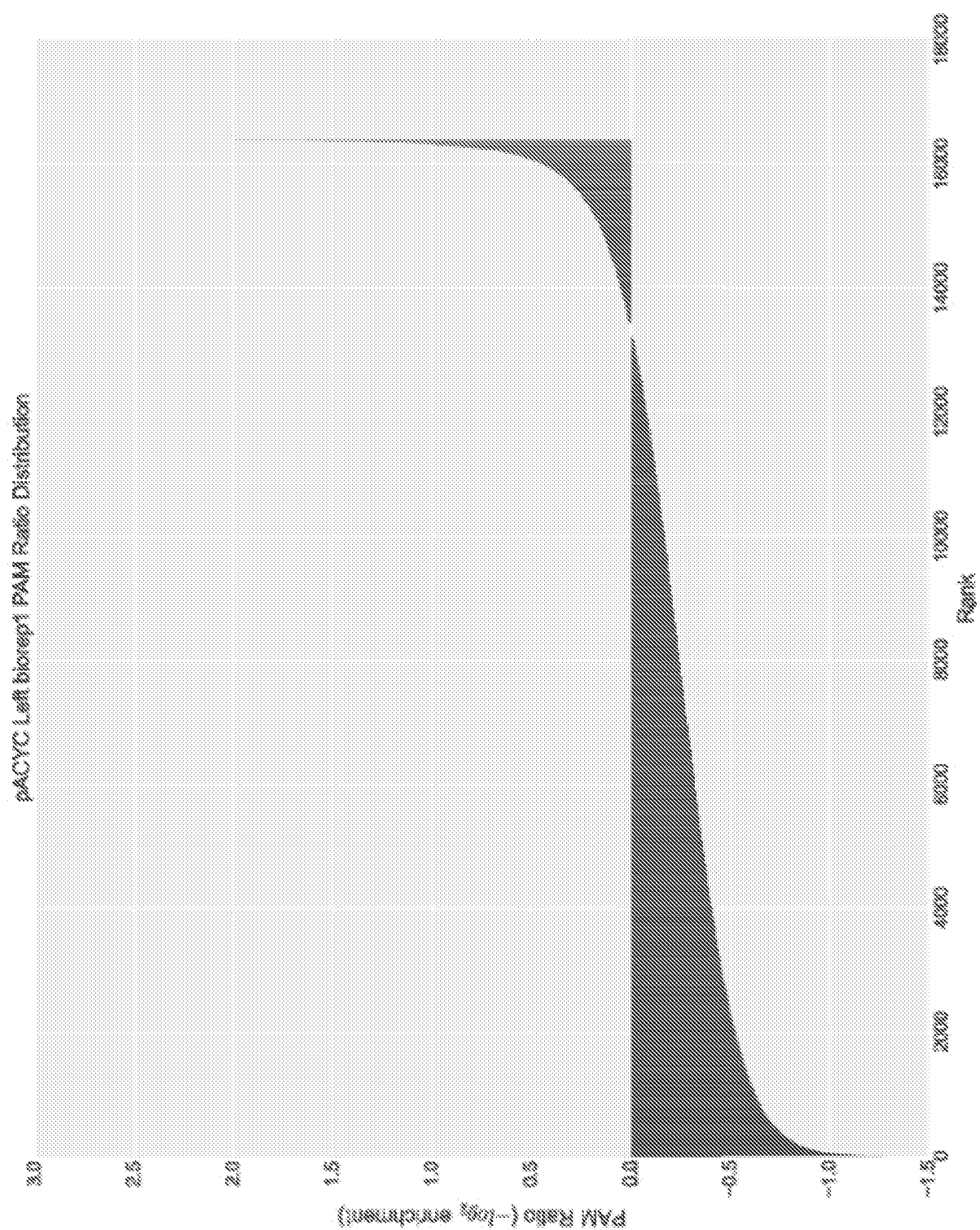


FIG. 62B

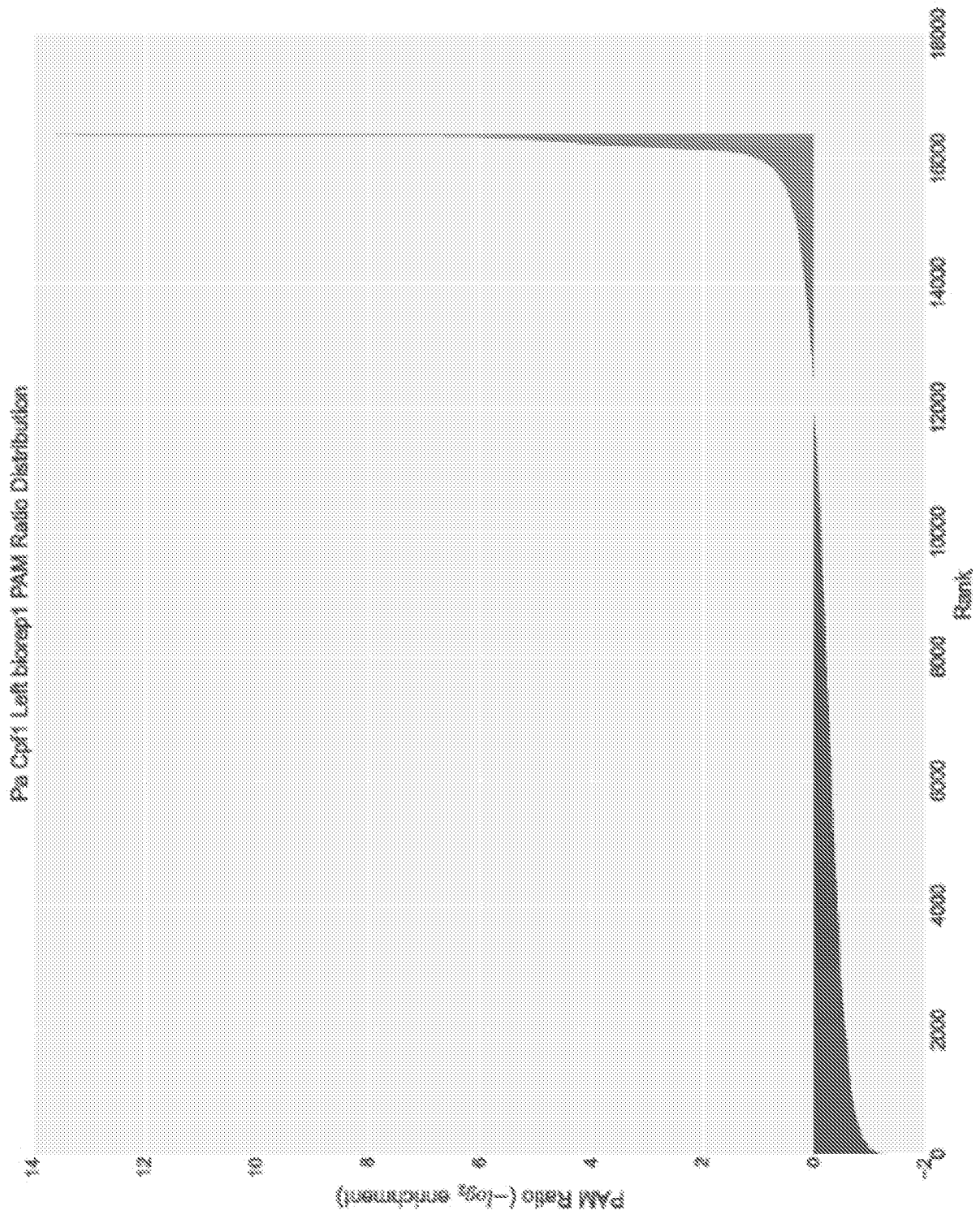


FIG. 62C

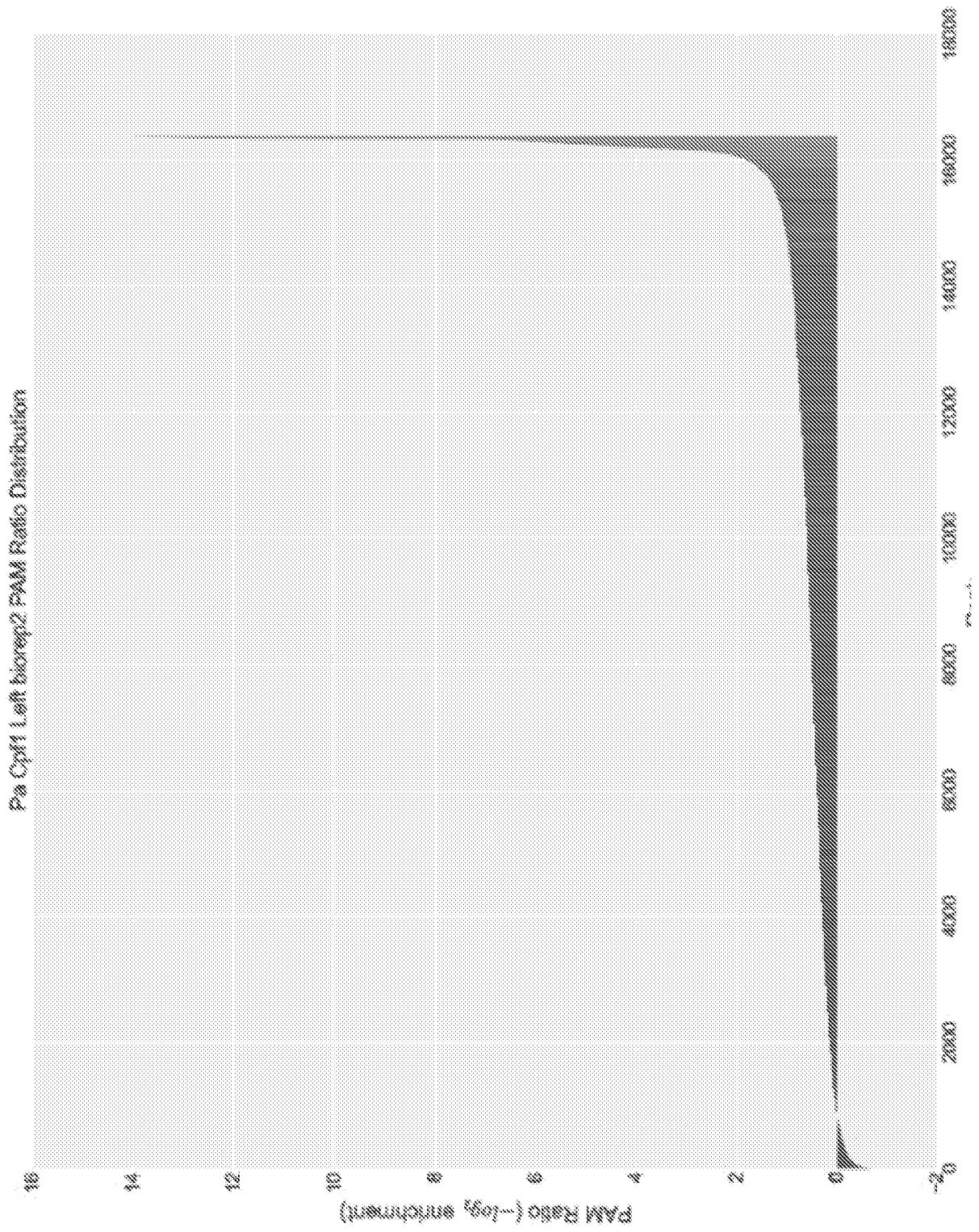


FIG. 62D

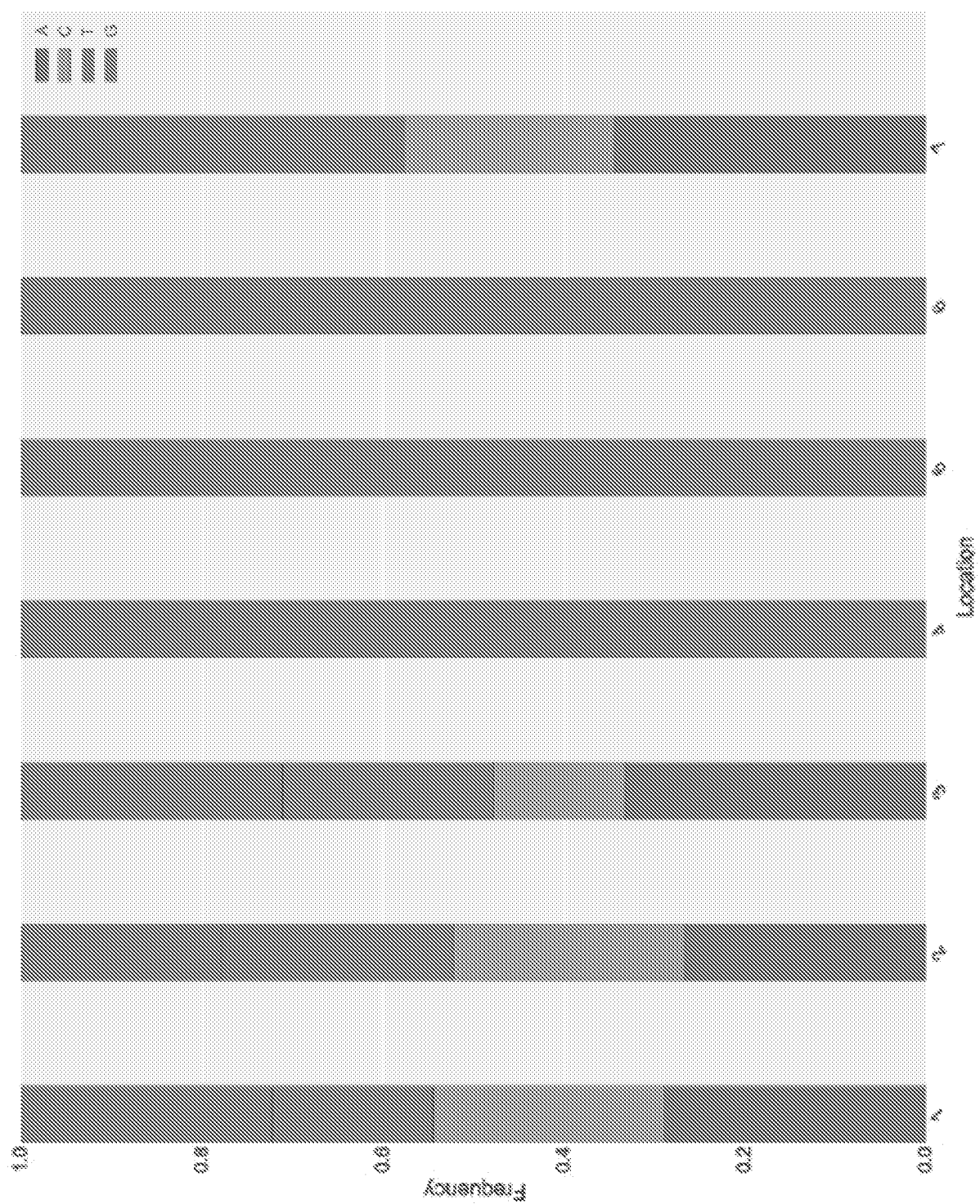


FIG. 62E

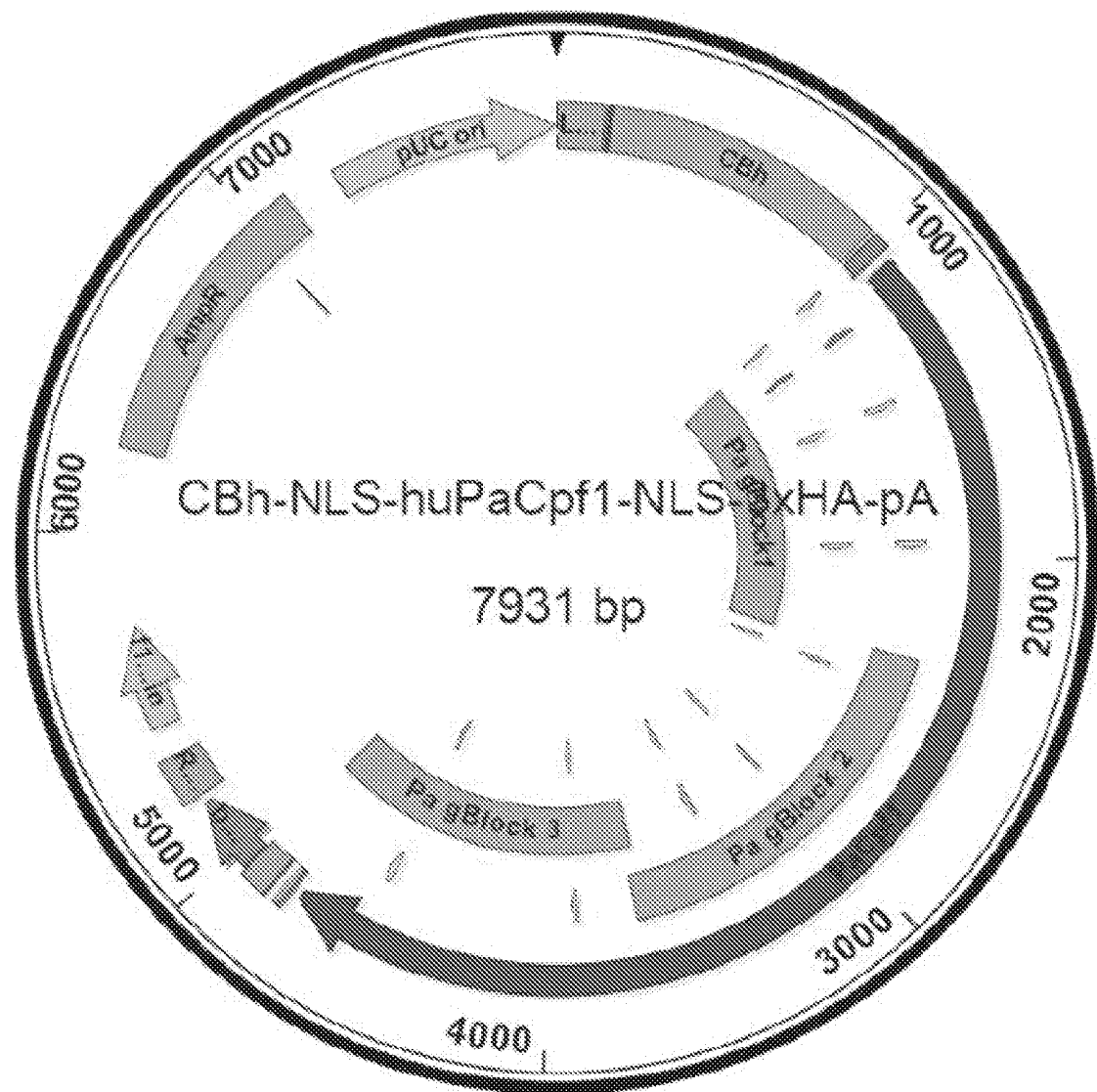


FIG. 63

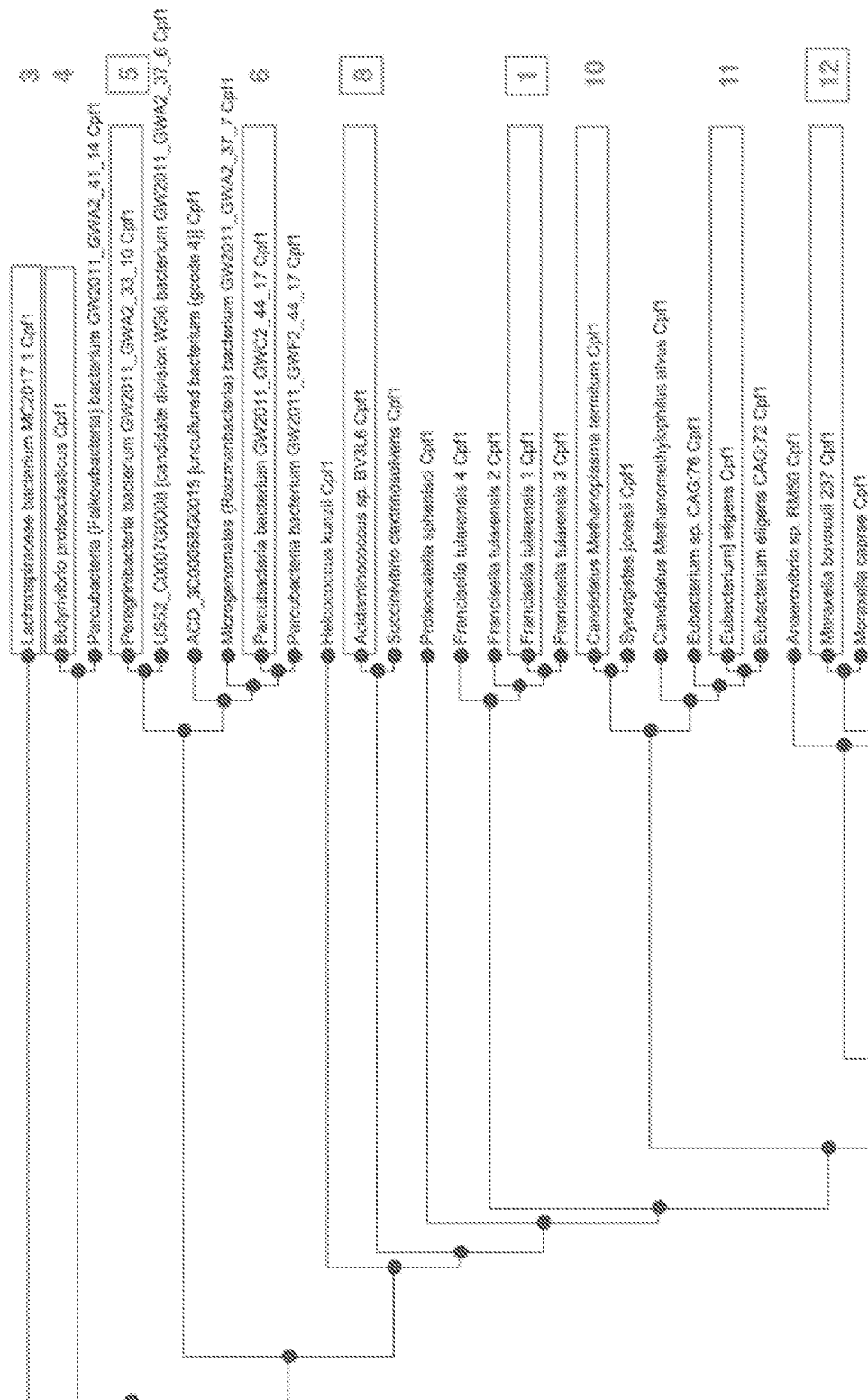


FIG. 64A

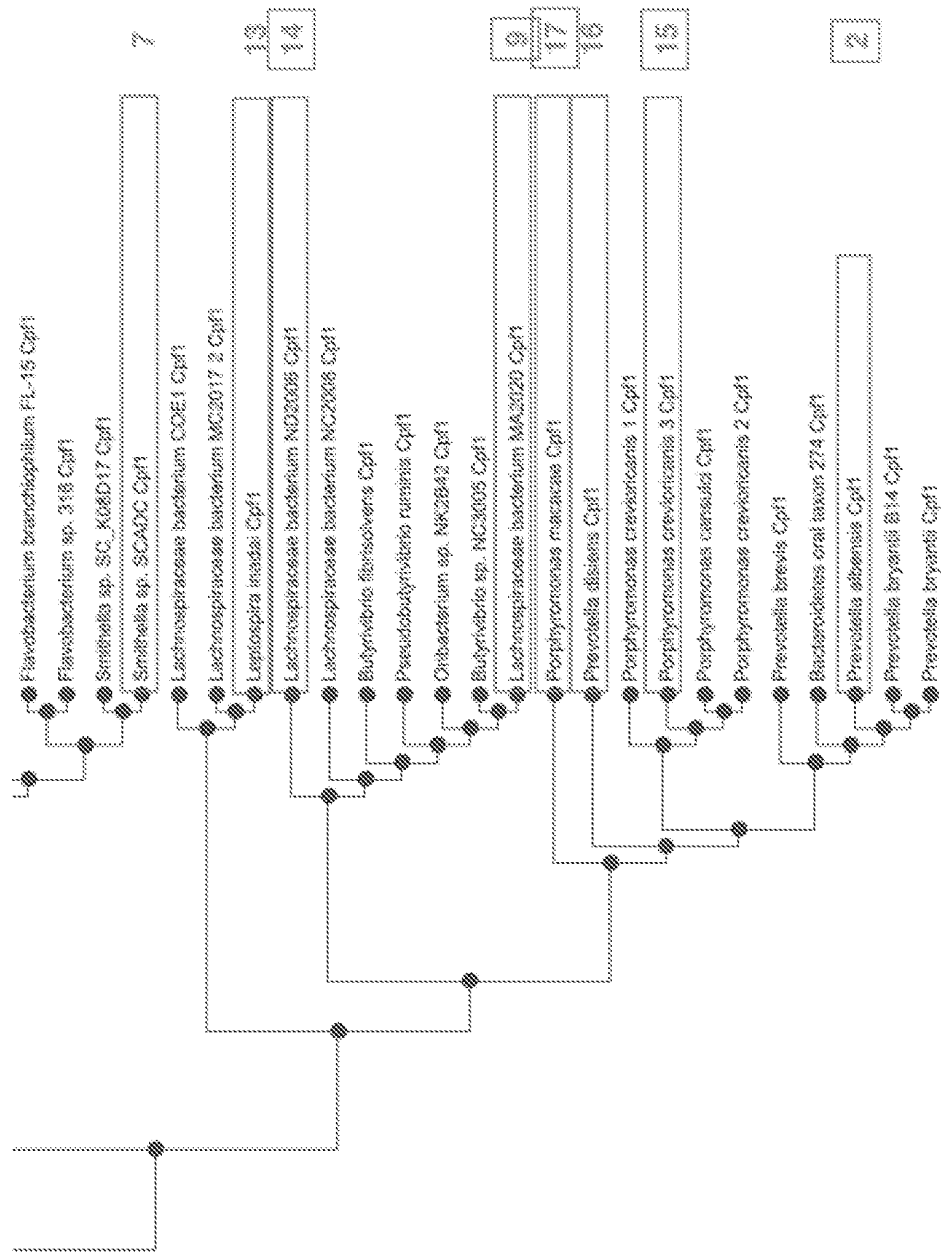


FIG. 64B

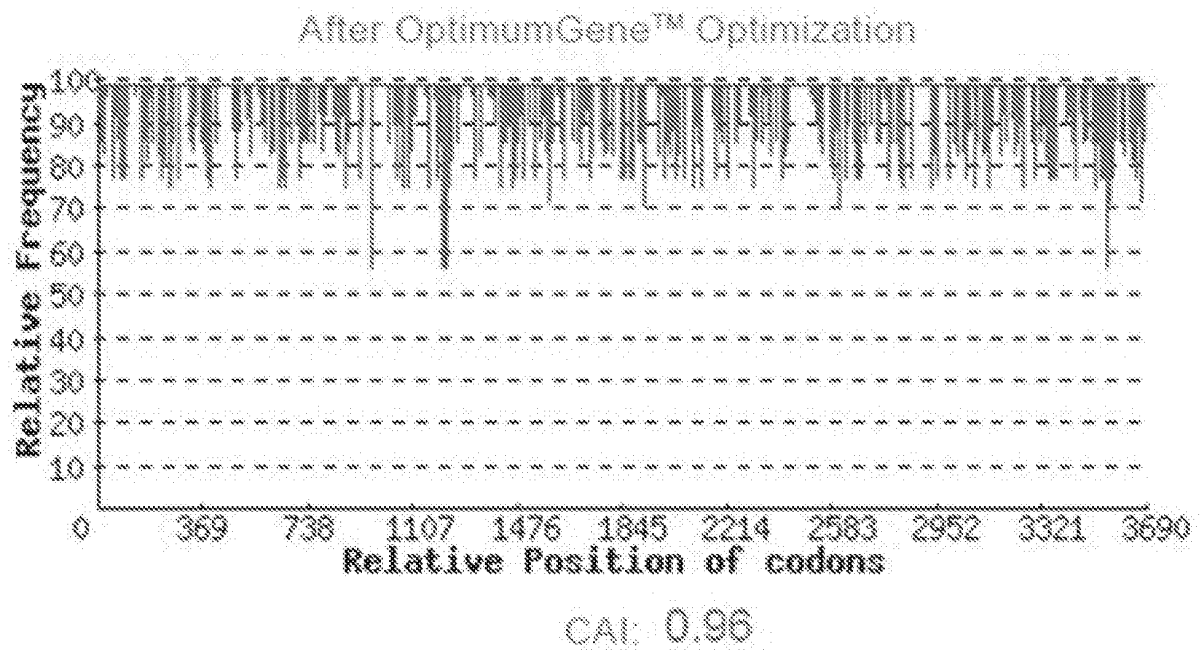


FIG. 65A

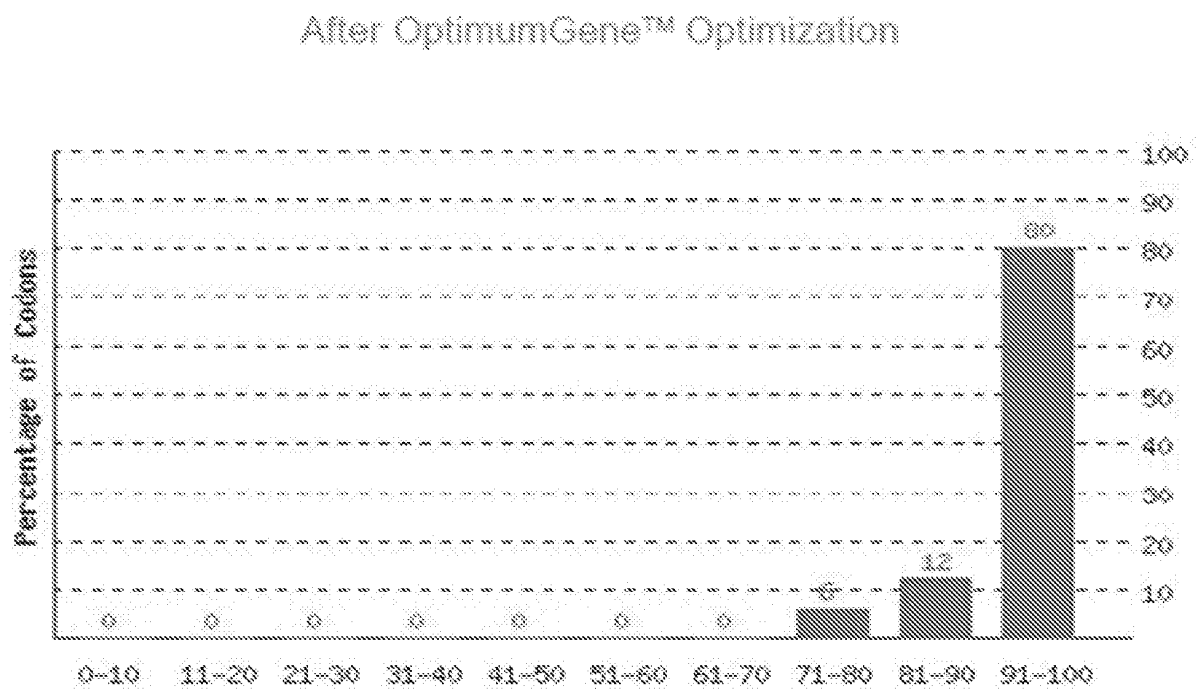


FIG. 65B

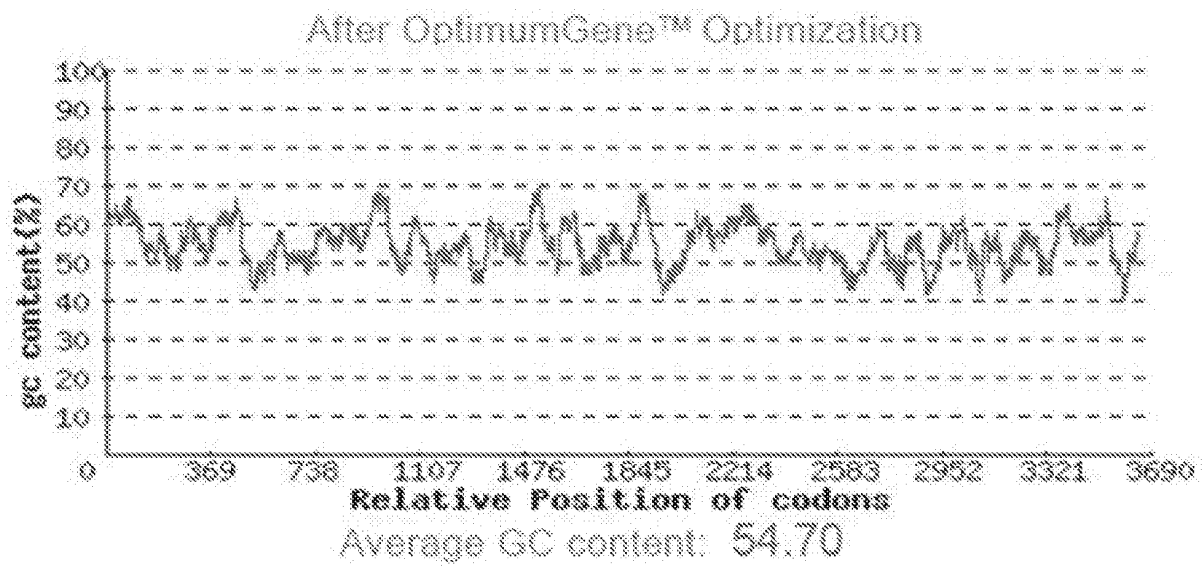


FIG. 65C

Restriction Enzymes Optimized

* Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.

BamHI(GGATCC)	1(3754)
EcoRI(GAATTC)	1(3844)
HindIII(AAGCTT)	0
BsmBI(CGTCTC)	0
BsaI(GGTCTC)	0
BbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(3818)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BarGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0

CIS-Acting Elements Optimized

Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0

Antiviral Motifs Optimized

0

FIG. 65D

After Optimization

Max Direct Repeat: Size:15 Distance:2688 Frequency:2
 Max Inverted Repeat: None
 Max Dyad Repeat: Size: 13 Tm: 49.0 Start Positions: 1666, 1908

FIG. 65E

SEQUENCE

ATGGAATTACGGCAACCGGCCAGTTTGGACCGGAGAGGCCCCCTGACCAAGACAATCAACCTGCGCTGAAGCCTATC
 GGCGAGACACCGGAGACAATCCGCGAGCAGAGGCTGCTGGAGCAGGACCGCCGCTTCAGGAAGCTGGTGGAGACA
 GTGACCCCTATCTGTGACGATTGTATCAAGGAAGATCGCCGATAACCGCCCTGTGCCACTTTGGCACCGAGTATGAC
 TTCAGCTGTCTGGGCAACGCCATCTCTAAGGATGACAGCAAGGCCATCAAGAAGGAGACAGAGAAGGTGGAGAAAG
 CTGCTGGCCAAAGGTGCTGACCGAGAATCTGCCAGATGGCCTGCGCAAGGTGAACGACATCAATTCCGCCGCTTT
 ATCCAAGGATACACTGACCTCTTTCTGTGACGAGCGATGCCGACAAGCGGGTCTGATCCAGGAGCTGAAGGGCAAG
 ACCGTGCTGATGACAGCGGTTCTTGACCAACACGGATCACAGCCCTGACCTGTGGCTGCCCGACAGAGTGTTCGAG
 AACTTTAATATCTTTCATCGAGAACGCCGAGAGAGATGAGAAATCCTGCTGACTCCCTCTGAATGAGAGATCATG
 AAGTTTGAACCCAGATGCCGAGCACTACCCCTCTCTGGAGTTCTATGGCCAGTGGCTGTCTCAGAAGGACATGAT
 AGCTACAACCTGATCATCTCTCGGCATCTATGCGGACGATGAGGTGAAGAACCTTGGCATCAATGAGATCGTGAAG
 GAGTACAATCAGGAGATGCCGGGGGACAAAGGATGAGTCCCCACTGCCCAAGCTGAAGAAGCTGCACAAGGAGATC
 CTGATGCCAGTGGAGAAAGGCTTCTTTGTGCGCGTGTCTTCTAACGACAGCGATGCCCGGAGCATCTCTGGAGAAAG
 ATCCTGAAGGACACAGAGATGCTGCCCTCCAAGATCATGAGGGCCATGAAGGAAGCAGATGACGGGCGACATGCC
 GTGTACGGCAGCCCGCTGCAAGAGCTGAGCCACGTTGATCTACGGCGATCACGGCAAGCTGTCCAGATCATCTAT
 GACAAAGAGTCCAGAGAGGATCTCTGAGCTGATGGAGACACTGTCTCCAAAGGAGCGCAAGGAGAGCAAGAAAGCGG
 CTGGAGGGGCTTGGAGGAGCAGATCAGAAAGTCTACATACACCTTGGACGAGCTGAACAGGTATGCCGAGAGAAAT
 GTGATGGCAGCATACATGCGAGCAGTGGAGGAGTCTTGTGCCGAGATCATGAGAAAGGAGAAAGATCTGAGGAGC
 CTGCTGAGCAAGGGGAGCTGAGAGATCCGGGGCAACAGACACAAATACACTGATCGTGAAGAACTACTTTAATGCC
 TGGACCGTGTTCGGAAACCTGATCAGAAATCCTGAAGGCGCAAGTCCGAGGGCGAGATCGACTCTGACTTCTACGAT
 GTGCTGGACGATTCGGTGGAGGTGCTGTCTCTGACNTACAAGGGGAGAGAACTGTGTGCCGAGCTATATCAACCAAG
 AAGATCGGCTCCGACCTGAAGCCCGAGATCGCCACATACGGCAGCGCCCTGAGGCCAACAGCCCTGCTGTGTGCC
 CCGAGAGAGAAAGTTAATGTGAAGTTCCACACCATCGTGGCGAGAGATGCCCGGCTGTACTATTTTCATCTGTGCC
 AAGGGCGCAAGCCTGTGGAGCTGAGGACATGGATGGCGACATCGAGTGTCTGAGATGAGAAAGATCCCTAAC
 CCAACAATCTTTCTGCCCAAGCTGTGTCTCAAGGACCTGAGGCCCTTCTTTAGGGATAATCCAGAGGCCGAGCGAG
 TTCTGTGTTCTGAGCGGCATGAAGGCCCGGTGACAAATCAACAGAGAGACATACGAGGCCCTACAGGTATAAGCTG
 TATACCGTGGGCAAGCTGCGGATGCGGAGGTGTCCGAGAGGAGTACAGCGGGGCGCTGCTGACAGGTGCTGACC
 GCCTACAAGGAGTTTCTGGAGAACAGAAATGATCTATGCCGACCTGAATTTGCGCTTTAAGGATCTGGAGGAGTAT
 AAGGACAGCTCCGAGTTTATCAAGCAAGTGGAGACACACAACACCTTCAATGTGCTGGGCCAAGGTGTCTAGCTCC
 CAGCTGGACGATCTGTGAAATCTGGCAACCGGCTGTGTCTGAGATCTGGAGCGAGCGCTGGAGTCTTACTAT
 AAGTACGGCAATGAGAAAGTGTGCGGGCTATGAGGGCTGTGCTGAGCATCTGAGAGGATGAGAACCTGTGT
 TCCATGGGAGCCCTGCTGAACAGCCCGCCCATGCTGTGTATCCCGCCAAAGGAATCTAGCAAGCCTATGCTGTGT
 CACCGGATGGCAGCAGAGTGTGTGACAGTTTGATAAGGACGGCAAGTACATCCCCCTGAGGTGCAACGAG
 CTGTATCCCTTCTTTAACAATCTGCTGATCAAGGAGAGGCTGGGCGAGAAAGCCCGGAAGATCTTGACAAACAAG
 AAGGTGAAGGTGAAGGTGCTGGAGAGCGAGAGAGTGAAGTGGTCCAAAGTTCTACGATGAGCAGTTTGGCGTGACC
 TTCAGCGTGAAGAGAAACGCCGATTGTCTGGACACCAACAAGGACCTGAATGCCGAGTGAAGAGCAGTATAGC
 GAGTCCAACAGACTGATCTGTATCAGGAATACCACAGATATCTGTACTATCTGGTGTGGACAAGAATGGCAAG
 GTGCTGAAGCAGAGATCCCTGAACATCATCAATGACCGCGCCAGGGATCTGACTGGAAGGAGAGGTTCCGCCAG
 GTGACAAAGGATAGAAACGAGGGCTACAATGAGTGGGATTTATCCAGGACCTCTAACGACCTGAAGGAGGTGTAC

FIG. 65F

FIG. 65G

MDYGNGQFERRAPLTKTITLRLKPIGETRETIREQKLLSEQDAAFRKLIVETVTPIVDDCIR
 KIADNALCHFGTEYDFSCLGNAISKNDKAIKKETEKVEKLLAKVLTENLPDGLRKVNDI
 NSAAFIQDTLTSFVQDDADKRVLIQELKGKTVLMQRFLTTRITALTVWLPDRVFENFNIF
 IENAEKMRILLDSPLNEKIMKFDPAEQYASLEFYGQCLSQKDIDSYNLIISGIYADDEV
 KNPGINEIVKEYNQQIRGDKDESPLPKLKKLHKQIILMPVEKAPFVRVLSNDS DARSILEK
 ILKDTEMLPSKIIIEAMKEADAGDIAVYGSRLHELSHVVIYGDHGKLSQIIYDKESKRISL
 METLSPKERKESKKRLEGLEEHIRKSTYTFDELNRYAEKNVMAAYIAAVEESCAEIMRKE
 KDLRTLKSKEDVKIRGNRHNTLIVKNYFNAWTVFRNLIRILRRKSEAEIDSDFYDVLDD
 VEVLSLTYKGENLCRSYITKKIGSDLKPEIATYGSALRPNSRWWSPGEKFNVKFHTIVRR
 DGRLYYFIFLPGKAKPVELEDMDGDIECLQMRKIPNPTIFLPKLVPKDPEAFPRDNPEADE
 FVPLSGMKAPVTITRETYEAYRYKLYTVGKL RDGEVSEEEYKRALLQVLTAYKEPLENRM
 IYADLNFGPKDLEEYKDSSEFIKQVETHNTFMCAKVVSSSQLDDLVKSGNGLLF EIWSER
 LESYYKYGNKVLRGYEGVLLSILKDENVLMRTLLNSRPMLVYRPKESSKPMVVHRDGS
 RVVDRFDKDGKYIPPEVHDELYRFFNNLLIKEKLGEKARKILDNKKVKVKVLESERVKWS
 KFYDEQPAVTFSVKKNADCLDTTKDLNAEVMEQYSESRLILIRNTTDILYYLVLDKNGK
 VLKQRSINIINDGARDVDWKERFRQVTKDRNEGYNEW DYSRTSNDLKEVYLNALKEIAE
 AVIEYNAILIIEKMSNAFKDKYSFLDDVTFKGFETKLLAKLSDLHFRGIKDGEPCSP TNP
 LQLCQNDSENKILQDGVIFMVPNEMTRSLDPDTGFI FAINDHNI RTKKAKLNFLSKFDQLK
 VSSEGCLIMKYSGDSLPTHNTDN RVWNCCCNHPITNYDRETKKVEFIEEPVEELSRVLEE
 NGIETDTELNKLNERENVPGKVVD AIYSLVLN YL RGT VSGVAGQRAVYYSPVTGKKYDIS
 FIQAMNLRKCDYYRIGSKERGEWTD FVAQLIN

FIG. 65H

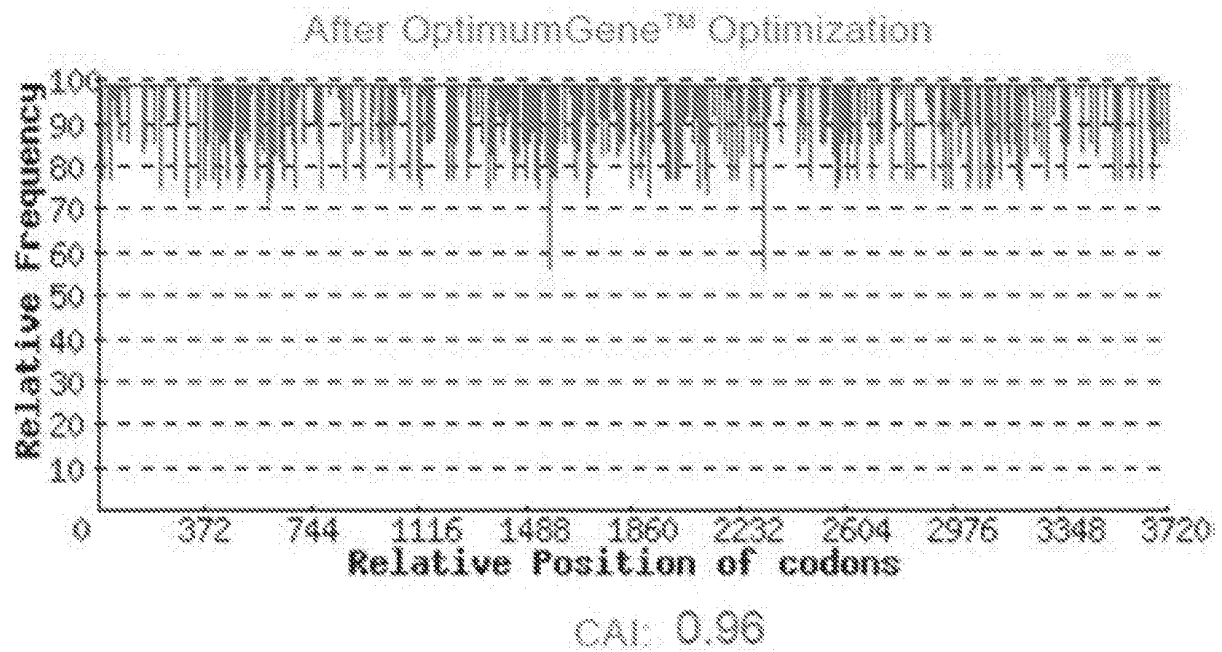


FIG. 66A

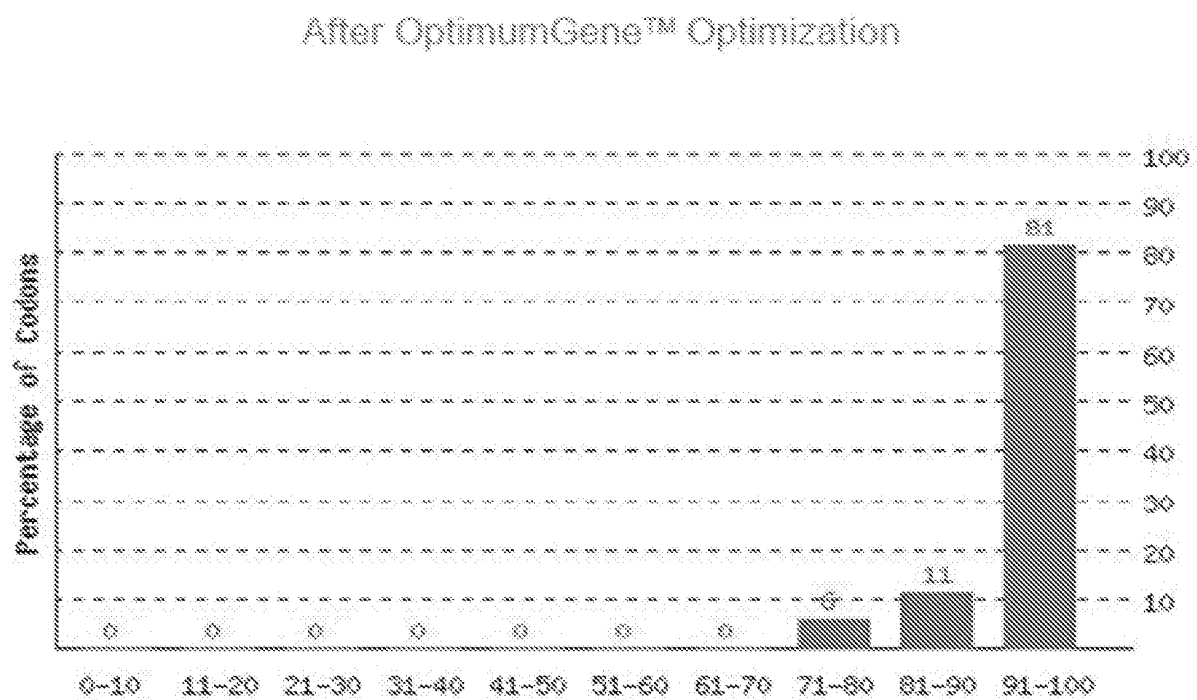


FIG. 66B

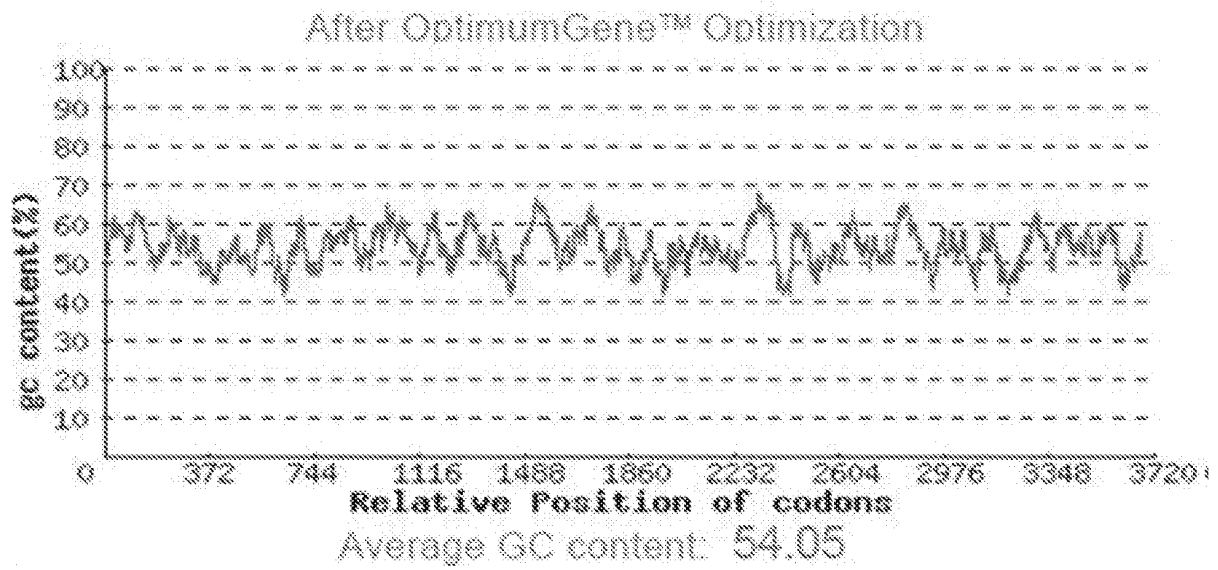


FIG. 66C

Restriction Enzymes	Optimized
* Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.	
BamHI(GGATCC)	1(3778)
EcoRI(GAATTC)	1(3868)
HindIII(AAGCTT)	0
BsmBI(CGTCTC)	0
BsaI(GGTCTC)	0
BbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(3842)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BsrGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0

CIS-Acting Elements	Optimized
Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0

Antiviral Motifs	Optimized
	0

FIG. 66D

After Optimization

Max Direct Repeat: Size:12 Distance:3285 Frequency:2
 Max Inverted Repeat: Size: 12 Tm: 43.5 Start Positions: 1631, 1547
 Max Dyad Repeat: None

FIG. 66E

GGTACC

ATGCTGCTGTATGAGAACTACACCAAGCGGAATCAGATCACAAAGAGCCTGAGGCTGGAGCTGCGCCCTCAAGGC
 AAGACCCCTGAGAAACATCAAGGAGCTGAATCTGCTGGAGCAGGACAAAGGCCATCTACGCCCTGCTGGAGCGGCTG
 AAGCCATGTATCGATGAGCGGCATCAAGGACATCGCCAGAGATAACCTTGAAGAACTGCGAGCTGTCTTTTGAAGAG
 CTGTACCAACCACTTCTCTGAGCGGCACAAAGAGCCCTATGCCCAGGAGTCCGAGCGGCTGAGAGAGGAGATCGTG
 AAAACCCCTGATCAAGAACCTGCGCGAGGGCATCGGCAAGATCTCTGAGATCAACAGCGCCAGTATCTGAATGGC
 GTGCTGTACGACTTCATCGATAAGACACACAAGGACTCTGAGGAGAGGACGAATATCCTGAGCGATATCCTGGAG
 ACAAAGGGCTACCTGGCCCTGTTCTTCCAAGTTTCTGACATCTGGGATCAACACACTGGAGCAGTCCATGCGCTAAG
 AGAGTGATGAGAACTTCTGAGATCTATGCGCGCAATATCCCAAAGATGCAGGACCGCTTGGAGAGGGGAGCGCTG
 TCTTTTSCCATCGAGTACGAGAGCATCTGTTCCGTGGATTACTATAACCAAGATCCTGTCCAGGAGGACATCGAT
 TCTTATAANTCGCCTGATCAGCGGCATCATGACGAGGATGGCGCCAGGAGAGGGGCATCAACGAGACCATCTCC
 GAGAGGAATATCAAGATCAAGTCTGAGCACCTGGAGGAGAGGCCCTTCAGAACTCTGAAGCAGCTGCAAGGAG
 ATCCTGAGAGGAGCGCGAGAGGCCCTTTACAATCGACCAATCGACAGCGATGAGGAGGTGCTGCAAGTGACCAAG
 GAGGCTTTCGAGCAGACAAAGGAGCAGTGGGAGAACATCAGGAGATCAATGGCTTCTACGCCAAGGAGCTCCGCGC
 GATATCAACCTGTTTATCGTGGTGGGCCCTAACCCAGACACACGTCCTGTCTCCAGCTGATCTACGCGGAGCAAGAC
 CGGATCAGACTGCTGCTGGAGGAGTATGAGAGAACACCCCTGGAGGTGCTGCCCCGAGAGAACAAAGTCTGAGGAG
 GCCAGATACGATAAGTTCTGTAATGCGCTGCCCTAAGAGAGGTGGCCAAAGAGAGGCCACACCTTCCAGCGGCTGCGAG
 AAGATGACAGGCGAGCATCGGCTGTTTATCTCTGTATAGAGATGAGGCTGGCCCCGAACTACATGAGGAATCAAGGAG
 GCCTATGGCACCTTTGAGCGGGACATCTCTGAAGAGCAGCGCGCGCATCAAGGGCAATCGGACATGACAGGAGTCT
 CTGCTGAGCTTTTACGATGAGCTGACAAAGTTCAAGGAGCGCCCTGGCGCATCAATTCCTGCAACGAGAGAGAG
 GCGGATCCANTCTTCTATAACACCTTGTATGGCATCTTCTGAGAGGCGCAATAGGACATACAAAGGCGAGAACTG
 TGCCGCAACTATGTGACCAAGTCTCCAGCGGACGATGCGCGCATCATGGCCAGCTGTCTGGGCAACCCAGCAAGG
 CTGCGCACACACTGGTGGAAATGGCGAGAGAACTTCCGCATCAATGACCTGCGCATGATCCCGAGAGGCGATGAG
 TACTATTACTTTGTGCTGACCCAGACGTGAAGCCCGTGGACCTGAAAACCAAGGACGAGACAGATGCCAGATC
 TTGCTGCGCGCAAGGGCGCCAAAGTCCCTTCTGGGCGCTGCCAAGGCGCTGTTCAAGTGACATCCTGGAGCCTTAC
 TTTGAGTCCCGAGAGCACAGAAATGACAAGAACTGCGTGATCGAGGAGTACGTGTCTAAGCCCTGACCATCGAC
 AGGCGCGCCCTATGATATCTTTAAGAACCGGCACCTTCAAGAAACCAATATCGGCATCGAGCGGCTGACCGAAGAG
 AAGTTCAAGGACGATTGCGGCTATCTGATCGACGTGTATAAGGAGTTGATCGCCGTGTATCAAGGTACAGCTGT
 TTTAACATGTCCGCGCTGAAGCGCGCTGACGAGTACATGATATCGGCGAGTTCCTTTCTGACGTGGATAACGCG
 CTGTGCACAATGGAGTGGATTCCCGTGAGCTTCGAGCGCATCAACGACATGGTGGATAAGAGGAGGGCCCTGCTG
 TTTCTGGTGAGGAGCATCTTCTCTGTATAACCGGCGCGAGAGCCCTTATGAGCGCACCTTTATCCAGCTGTTGAGC
 GACTCCAACATGGAGCACACATCTATGCTGCTGAATAGCAGGGCCATGATCCAGTACAGGGCAGCCCTCCCTGCCA
 CGGAGAGTGACCCACAAGAAGGGGAGCATCTGCTGGCGCCCTGCGGGATTCCAACGCGGAGCACATCCCCATGAC
 ATCAGAGAGGGCCATCTACAAGATGAAGAACAATTTTGCATCAGCTCCGAGGATTTCTATCATGCGCAAGGCCTAT
 CTGGCGGAGCACGACGTGGCCATCAAGAAGGCCAAGCGAGGATATCTATCAGGAATAGGCGCTACACCGAGGACAAG
 TTCTTTCTGTCTCTGAGCTATACCAAGAACGCGGATATCAGCGCGCGCACACTGGACTACATCAATGATAAGGTG
 GAGGAGGACACACAGGATTCAGGATGGCCGTGATCTGTGACCCGCAACCTGAAGGACCTGACATAGTGGCCGTG
 GTGGATGAGAGAAACAATGTCTGAGCGAGAGAGCCCTGAACGAGATCGACCGCGTGATTTATCGGAGGCTGCTG
 AAGGAGAGAAACCAAGATCAAGTACCAAGACAGACACCGCTGTGCGCATATGACGTGAGCAGCAAGGCGCTGAG

FIG. 66F

GAGGCCCTACTGTGGAGCTGGCCCTTGACCCAGATCTTCCAAGCTGGGCCACAAAGTATAACGCCCTTGGTGGTGTGGAG
 TCCATGTCTCTACTCTTCAAGGACAACTTCTCTTTTCTGGATGAGCCAGATCTTCAAGGCCCTTTGAGGCCCCCTCTG
 TGCGCCAGAAATGTCCGACCTGTCTTTTAAATACAATCAAGGAGGGCCAGGCCCGGCTCCCTCTCTAAACCCCTTCCAG
 GTGTCCAAACAATAACGGCCAAATTCCTTATCAGGACGGCCGTGATCTACTTCTCTGAATAACGGCCATAACCCGGAACCTG
 TGCCCTGATACCCGCTTTTGTGGACCTGTTCGATAAGAACCCGGCTGATCACBAATGCAGTCTAAGAGACAGTTCCTTT
 GCCAAGATGAAGGACATCAGAAATCGAAGATGGGGAGATGCTGTTCACCTTTAAACCTGGAGGAGTACCTTACAAAG
 AGGCTGCTGGACCCGCAAGGAGTGGACCTTGAAGATGCGCCGGCGATGGCTCCTATTTGGACAAGGATAAGGGCCAG
 TACTTGTACGTGAACGACATCGTGAAGAGAGCAGATCATGCCAGCCCTGCTGGAGGACAAAGGCCCTGTTCGATGGC
 AATATGGCCGAGAAAGTTTCTGGATAAGGACCGCCATCAGCCGGCAAGTCCCTGGAGCTGATCTACAAAGTGGTTGCC
 AACGCCCTTTATGGCATCATCAAAAGAAAGACCGCGGAGAAAGTCTACCGGAGCCCCATCAACCGCACAGAGATC
 GACCTGAGCAAGAACACCCACATACAACTTCGGCAAGAACTTCATGTTCAAGCAGGAGTATAGAGGCCGACCGCGAT
 TTTCTGGACCGCTTCTTGAATTACATGCAGGCCCGAGATATCGCCCTG
 AAAAGGCCCTGGCTTACGAAAGAACTCTTCTAGGACAAAAGGAAAGAGGAGTCTTACTCTTACGATGTTCTTCAAT
 TAGCTTTATGCTTACGACCTGCTTGAATATCGATACCGATATGATGCTCCCGCACTATGCTTAAAGAAATTC

FIG. 66G

MLLYENYTKRNQITKSLRLELRPQGKTLRNIKELNLLEQDKAIYALLERLKPVIDEGIKD
 IARDTLKNCELSFEKLYEHFLSGDKKAYAKESERLKKEIVKTLIKNLPEGIGKISEINSA
 KYLNGVLYDFIDKTHKDSEKQNILSDILETKGYLALFSKFLTSTRITTLEQSMFKRVEN
 FEIYAANIPKMQDALERGAVSFAIEYESICSVDYNNQILSQEDIDSYNRLISGIMDEDGA
 KEKGINQITISEKNIKIKSEHLEEKPFRIKQLHKQILEEREKAFTIDHIDSDEEVVQVTK
 EAFEQTKEQWENIKKINGFYAKDPGDITLFIIVGPNQTHVLSQLIYGEHDIRLLLEEYE
 KNTLEVLPRRTKSEKARYDKFVNAVPKKVAKESHTFDGLQKMTGDDRLFILYRDELARNY
 MRIKEAYGTFERDILKSRGKIGNRDVQESLVSFYDELTKFRSALRIINSGNDEKADPIF
 YNTFDGIFEKANRTYKAENLCENYVTKSPADDARIMASCLGTPARLRTHWWNGEENFAIN
 DVAMIRRGDEYYYFVLTPDVKPVLDLKTDETDQIPVQRKGAKSFLGLPKALFKCILEPY
 FESPEHKNDKNCVIEEYVSKPLTIDRRAYDIFKNGTFKKTNIGIDGLTEEEKPKDDCRYLI
 DVYKEFIAVYTRYSCFNMSGKRADEYNDIGEFFSDVDTRLCTMEWIPVSFERINDMVDK
 KEGLLFLVRSMFLYNRPRKPYERTFIQLFSDSNMEHTSMLLNSRAMIQYRAASLPRRVTH
 KKGSI LVALRDSNGEHI PMHIREAIIYKMKNNFDISSDFIMAKAYLAEHDVAIKKANEDI
 IRNRRYTEDKFFLELSYTKNADISARTLDYINDKVEEDTQDSRMAVIVTRNLKDLTYVAV
 VDEKNNVLEEKSLNEIDGVNYRELLKERTKIKYHDKTRLWQYDVSSKGLKEAYVELAVTQ
 ISKLATKYNNAVVVVESMSSTFKDKFSFLDEQIFKAPEARLCARMSDLSFNTIKEGEAGSI
 SNPIQVSNNGNSYQDGVIIYFLNNAYTRTLCFDTGFVDVFDKTRLITMQSKRQFFAKMKD
 IRIDDGEMLFNLEEYPTKRLLDRKEWTVKIAAGDGSYFDKDKGEYVYVNDIVREQIIPA
 LLEDKAVFDGNMAEKFLDKTAISGKSVELIYKWFANALYGIITKKDGEKIYRSPITGTEI
 DVSKNTTYNFGKKFMPKQEYRGDGDFLDAFLNYMQAQDIIV

FIG. 66H

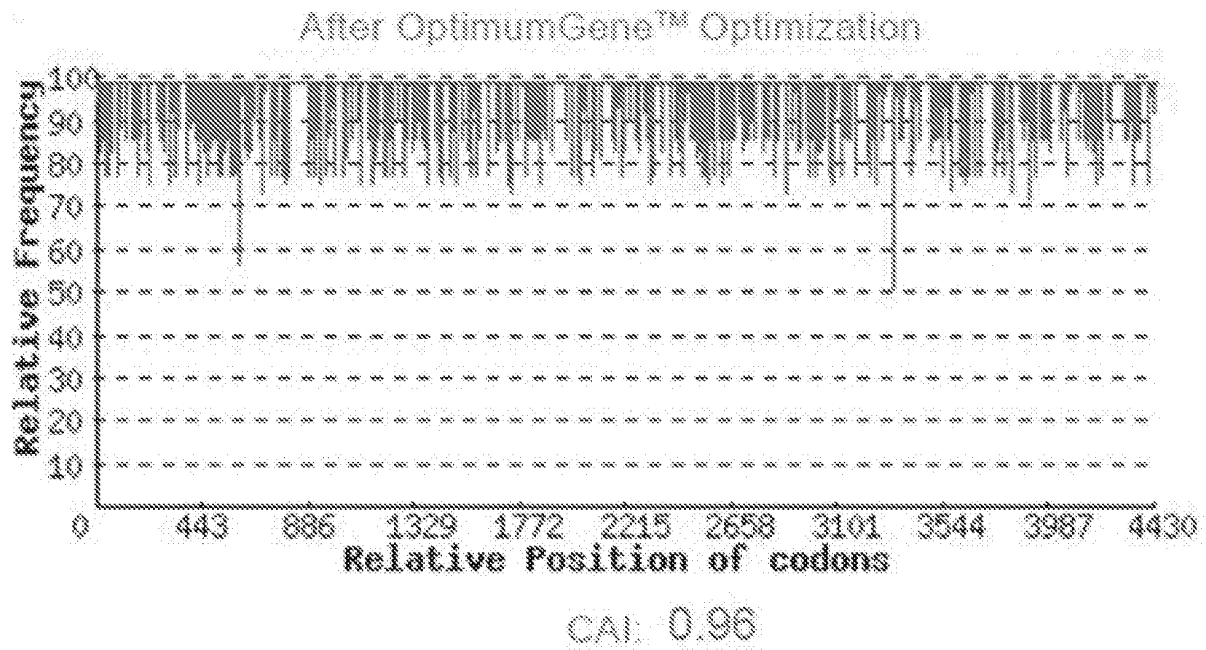


FIG. 67A

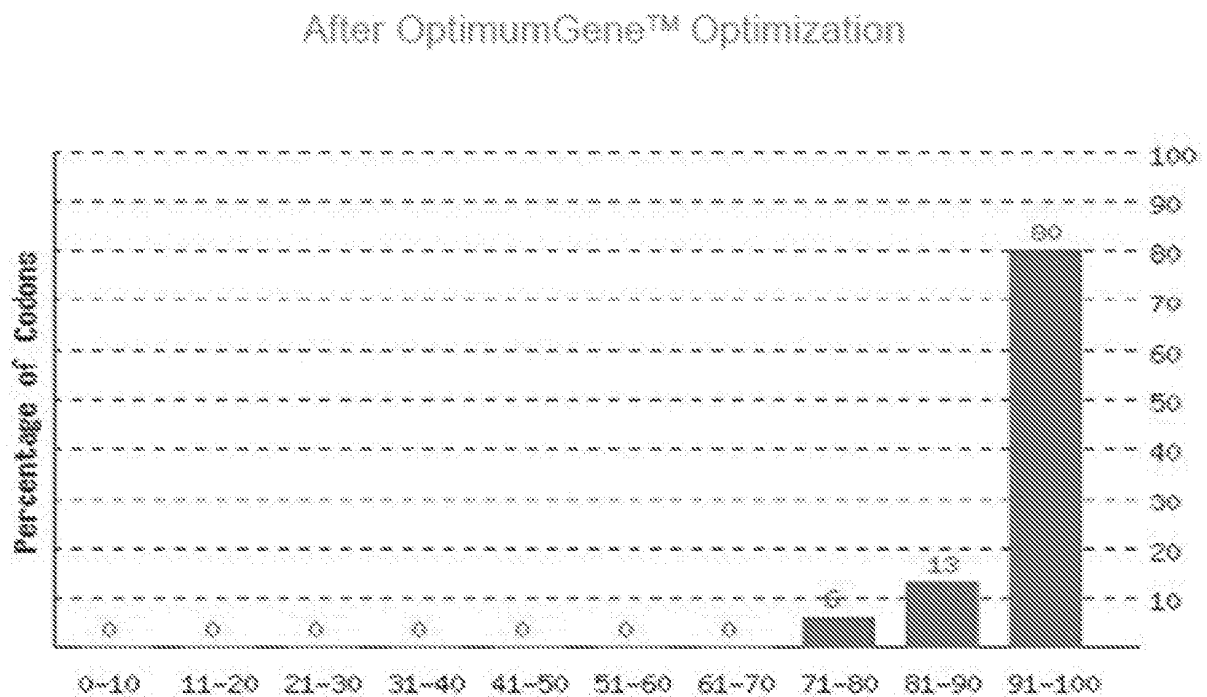


FIG. 67B

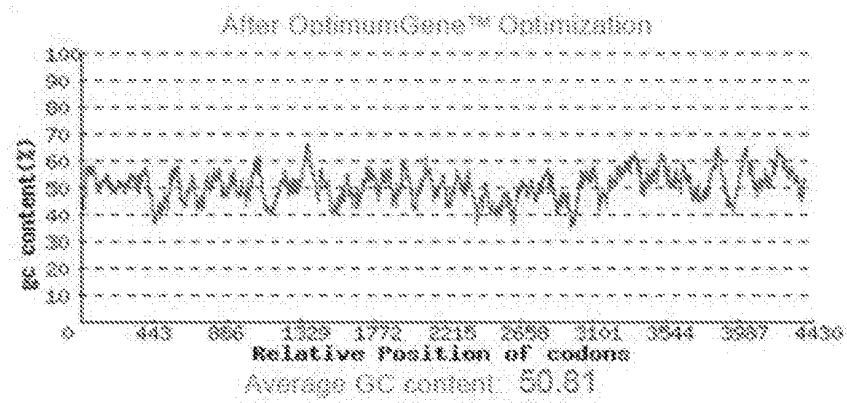


FIG. 67C

Restriction Enzymes	Optimized
* Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.	
BamHI(GGATCC)	1(4486)
EcoRI(GAATTC)	1(4576)
HindIII(AAGCTT)	0
BsmBI(CGTCTC)	0
BsaI(GGTCTC)	0
BbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(4550)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BsrGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0

CIS-Acting Elements	Optimized
Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0

Antiviral Motifs	Optimized
	0

FIG. 67D

After Optimization

Max Direct Repeat: Size:15 Distance:3543 Frequency:2
 Max Inverted Repeat: None
 Max Dyad Repeat: None

FIG. 67E

607A62

ATGTCCAACTTCTTTAAGAAATTTACCAACCTGTATGAGCTGTCCAAGACACTGAGGTTTGGAGCTGAAGCCCTTG
 GGCAGACCCCTGACAAACATGAAGGACCACTGGAGTACGATGAGAGGCTGCAGACCTTCTTGAAGGATCAGAAT
 ATCGACGATGCTTATCAGGCTCTGAAGCCTCAGTTCTAGAGAGATCCAGAGGAGTTTATCAGAGATTCTCTGGAG
 AGCAAGAGGGCCAGGAGATCGACTTCTCCGAGTACCTGGATCTGTTTTCAGGAGAAGAGGAGCTGAACGACTCT
 GAGAAGAAGCTGCGCAACAAGATCGGCGAGACATTCAACAAGGCGGCGAGAGGTGGAGAAGAGAGAGTACCCT
 CAGTATGAGTGAAGAGAGGGCTCCAAGATCGCCAATGGCGCCGACATCCTGTCTTGGCCAGGATATGCTGCAGTTT
 ATCAAGTATAAGAAACCCAGAGGATGAGAGAGATCAAGAAATTACATCGACGATACACTGAAGGGCTTCTTTAAGCTAT
 TTCCGCGGCTTTTAAATCAGAACAGGGCCAACTACTATGAGACAAAGAAGGAGGGCTCCACTCCAGTGGCAACAGG
 ATCCTGCAAGAGAACTCGCCAAAGTTCTGTGACAATGTGTATCCAGTTTAAAGCACATCATCAAGCCGGAAGAAGGAT
 GGCACCTTGAGAAAAACCGAGAGAAAGACCGAGTACCTGAACGCTTACCAGTATCTGAAGAACAAATACAAGATC
 ACACAGATCAAGGACGCGGAGACAGAGAGAGATGATCGAGTCTACACCCATCGCGGAGAGAGATCTTGGACGTGTAC
 TACTTTCAGCAGCTGCTGAGCCAGAGCAGATCGAGGATACAAACCGGATCATCGGCCACTATAATCTGCTGATC
 AACCTGTATAACCAAGGCCAAGAGATCTGAGGGCAAGCACTGAGCGCCAAAGAGAGAGATATAAGGAGCTGCT
 AAGTTCAAGACCTGTATATAAGCAGATCGGCTGCGCAAGAGAGAGGACCTGTTTACACAAATCAAGTGTGATACC
 GAGGAGGAGGCCAATAAGTCCCGGAAAGAGGGCAAGGAGTCCCACTCTGTGGAGGAGATCATCAACAAGGCCGAG
 GAGGCCATCAATAAGTACTTCAAGTCTAATAACGACTGTGAGAGATATCAACACCGTGGCCGAGTTTCATCAACTAT
 ATCCTGACAAAGGAGAAATTAAGAGGGCTGTATTGGAGCAAGGCGGCCATGAACACCATCTCCGACAAGTACTTC
 GCCAATTATCAAGACCTGCAGGATAGACTGAAGGAGGCCAAGGTGTTTCAGAGGCCGAGATAAGAGTCCGAGGAC
 GATATCAAGATCCAGAGGCCATCGAGCTGTCTGSCCTGTTCGGCGTGTCTGACAGGCTGGCCGATTTGGCAGACC
 ACCTGTTTTAAGTCTAGCATCTGAGCAACGAGGACAAAGCTGAAGATCATCACAGATTCCAGAGCCCTCTGAG
 GCTCTCTGAAGATGATCTTCAATGACATCGAGAAGACATGAGTCTTTCTGAAGGAGACAAACGATATCATC
 ACCCTGAAGAGTATAAGGGCAATAAGGAGGGCACCGAGAGAGATCAAGCAGTGGTTGAGCTATACACTGGCCATC
 AACCGGATGCTGAAGTACTTTCTGGTGAAGGAGAAATAGATCAAGGGCAACTCCCTGGATACCAATATCTCTGAG
 GCTCTGAAGAACCTGATCTACAGCGACGATGCGGAGTGGTTCAAGTGGTACGACGCTCTGAGAACTATCTGACC
 CAGAAGCCTCAGGATGAGGGCCAGGAGAAATAAGCTGAAGCTGAATTTGACAACCCATCTCTGGCCGGCGGCTGG
 GATGTGAACAAGGAGTGCAGCAATTTTGGCGTATCTGAAAGGACAAAGAACGAGAGAGGTACCTGGCCATCATG
 AAGAGGGCGAGAAATACCTGTTCCAGAAAGAGTGGACAGAGGGCCGGGGCAAGAACCTGACAAAGAGTCTAAT
 CCACTGTTTCAGATCAATAACTGCGAGATCTGAGCAAGATGGAGTATGACTTTTGGGCCGACGTCAGCAAGATG
 ATCCCCAAGTGTAGCAACCAAGCTGAAGGGCGTGGTGAACCACTTCAAGCAGTCCGACAATGAGTTTCATCTTCTCT
 ATCGGCTACAGGTGACAAGCGGCGAGAGGTTTAGGAGGAGTGCAGATCTCCAAGCAGGACTTCGAGCTGAAT
 AACCAAGGTGTTTTAATAAGAACGAGCTGAGCGTGAACGCCATGCGCTACGATCTGTCTCTACACAGGAGAGGAG
 TATATCAAGGCCCTTCCAGAGGAGTACTGGGAGCTGCTGTTTAAAGCAGGAGAGAGCGGACACCAAGCTGACAAAT
 AACGAGATCTTCAACGAGTGGATCAATTTTTCGAACAAGAGATAGCGAGCTGCTGTCTCTGGAGAGAGAGTAC
 AAGGATGCCCTGACCAATTTGATCAACTTCTGTAAGTACTTTCTGAGCAAGTATCCCAAGGACCACTGTTCAAC
 TACTCTTTTAAAGGAGAGCGAGAAATTATAACTCCCTGGACGAGTCTTACCCGACGTTGGATATCTGTTCTTACAG
 CTGAATATCAACACCAATCAATAAGAGCATCTGTGATAGACTGGTGGAGAGGGCAAGCTGTACCTGTTTGAG
 ATCAAGAATCAGGACAGCAACGATGGCAAGTCCATCGGCCACAAGAAATAACCTGCACACCATCTACTGGAAAGCC
 ATCTTCGAGAAATTTTGACAACAGGCCCTAAGCTGAATGGCGAGGCGGAGATCTTCTATCGCAAGGCCATCTCCAAG

FIG. 67F

GATAAGCTTGGGCATCTTTGAAGGGCAAGAAAAACCAAGAAACGGCACCTGAGATCATCAAGAAATTACAGATTTCAGCCAAG
 GAGAAGTTTTATCTTGCACGTGGCCAATCAACCTTGAACCTTCTGCTCCAAATAACGAGTATGTTGAATGACATCTGTGAAC
 ACBAAGTTTCTACAAATTTTTCCBACCTGCACTTTTCTGGGCATCGATAGGGGCCGAGAAACACCTGGCCTACTATTCT
 CTGCTGAATAAGAAACGGGAGATCTTGGACCAGGGCACTGTAACCTTGCCTTTTCACTGACAAAGGATGGCAATCAG
 CGCAGCATCAAGAAAGGAGAACTACTTTTATAACAAAGCAGGAGGACAAAGTGGGAGGCCAAGGAGGTTGGATTGTTGG
 AATTATTAACCACTTGTGCTGATGCCATGGCCCTCTAAACGGGACATGGCCAGAAAGAAATTGGCAGAGATTCGGCACC
 ATCAAGGAGGCCAAGAAACGGCTACGTGAGCTTGGTCATCAAGGAAGATCGCCGATCTTGGCCCTGAATAACGAGGCC
 CCCCCTTTCAATCTGTCTGGAGGACCTGAATAACAGGCTTTTAAGCGCTCCAGACAGAAAGATCGATAAGAGCGCTGTAC
 CAGAAGTTTGGAGCTGGCCCTGGCCCAAGAAAGCTGAACCTTTCTGCTGGACAGAAATGCCAAGCGCGATGAGATCGGC
 TCCCCCTACAAAGGCCCTGTGCAGCTGACCCCCCTGTGGAATAACTACGGCGACATTGAGAAACAAAGAGCAGGCCGGC
 ATCATGCTGTATACCCGGGCCCAATTATACCTCTCAGACAGATGCCAGCCACAGGCTGGAGAAAGACCATCTATCTG
 AAGGCCCGGCCCGAAGGAGACAAACNTACAGAGGGAACGCCAAGATCAAGAACAAAGAGGCTGAAGGACCCAGATCATC
 GAGACATTCACCGATATCGGCTTTGACGGCAAGGATTACTATTCTTGGAGTACGACAAAGGCGAGTTTGTGGATGAG
 AAAACCGGCGAGATCAAGCCCCAAGAACTGGCGGCTGTACTCCGGCGAGAAATGGCAAGTCCCTGGACAGTTTCCGC
 GGAGAGAGGGGAGAAAGGATAAGTATGAGTGGAAAGATCGACAAGATCGATATCGTGAAGATCCTGGACGATCTGTTT
 GTGAATTTTGAACAAGACATCAGCCTGCTGAAGCAGCTGAAAGGAGGCGGTGGAGCTGACCCCGGAATAACGAGCAC
 GGCACAGGCGAGTCCCTGAGATTCGCGCATCAAACTTGATCCAGCAGATCCCGAATAACCGGCAATAACGAGAGAGAC
 AACGATTTTCATCCTGTCTCCCTCAGTGAGGGACAGAAATGGCAAGCACTTTGACTCTCGCGAGTACTGGGATAAGGAG
 ACAAGCGGCGAGAAAGATCAGCATGCCCCAGCTCCCGGAGTGCCCAATGGCGGCTTCAACATCGCCCGGAAGGCGATC
 ATCATGAACGCCCACATCTTGGCCAAATAGCGACTCCAAAGGATCTGTCCCTGTTCTGTCTGACGAGGAGTGGGAT
 CTGCACCTGAAATAACAAGACCGGAGTGGAAAGCAGCCTGAACATCTTTTCTAGCAGGAAGGCCATGGCCAGCGC
 AAGAAG
 AAGAGGCGGCGGCGCACGAAAGAGGCGGCGCACGAAAGAGAAAGAGAAAGAGGCTTCTACCGCTACGATGTTTCCAGAT
 TACGCTTATCTCCCTACGAGCTTGCTTGATTATGCACTACCGCATATGATGTTCCCGGACTATGGCTGAGAAATTC

FIG.67G

MSNFFKNPTNLYELSKTLRFELKPVGDTLTNMKDHLEYDEKLQTFLLKDQNIDDAYQALKP
 QFDEIHHEEFITDSLESKKAKEIDFSEYLDLFQEKKEKELNDSEKKLENKIGETFNKAGEKWK
 KEKYPQYEWKKGSKIANGADILSCQDMLQFIKYKNPEDEKIKNYIDDTLKGFFTYFGGFN
 QNRANYYYETKKEASTAVATRIVHENLPKPCDNVIQFKHIIKRKKDGTVEKTERKTEYLNA
 YQYLKNNNKITQIKDAETEKMIESTPIAEKIFDVVYFSSCLSQKQIBEYNRIIGHYNLLI
 NLYNQAKRSEGGHLSANEKKYKDLPKPKTLYKQIGCGKKKDLPYTIKCDTEEBANKSRNE
 GKESHVSVEEIIINKAQEAINKYFKSNNDNENINTVPDFINYILTKENYEGVYWSKAAMNTI
 SDKYFANYHDLQDRLKEAKVFQKADKKSEDDIKIPEAIELSGLFGVLDLADWQTTLFKS
 SILSNEDKLKIITDSQTPSEALLKMI FNDIEKNMESFLKETNDIITLKKYKGNKEGTEKI
 KQWFDYTLAINRMLKYFLVKENKIKGNSLDTNISEALKTLIYSDDAEWFKWYDALRNYLT
 QKPQDEAKENKLKLNFDNPSLAGGWDVNKECSNFCVILKDKNEKKYLAIMKKGENTLQK
 EWTEGRGKNLTKKSNPLFEINNCEILSKMEYDFWADVSKMIPKCSTQLKAVVNHFKQSDN
 EPIFPPIGYKVTSGEKFRECKISKQDPFLNNKVFNKNELSVTAMRYDLSSTQEKQYIKAF
 QKEYWELLFKQEKRDTKLTNNBEIFNEWINFNCNKYSELLSWERKYKDALTNNWINFCKYPL
 SKYPKTTTLFNYSFKESENYNLSLDEFYRDVDICSYKLNINTTINKSILDRIVEEGKLYLFE
 IENQDSNDGKSIGHKNNLHTIYWNAIFENFDNRPKLNGEAEIFYRKAISKDKLGIVKGKK
 TKNGTETIKNYRPSKEKFILHVPITLNFCSNNEYVNDIVNTKFYNFSLHFLGIDRGEKH
 LAYYSLVNKNGEIVDQGTNLNLPFTDKDGNQRSIKKEKYFYNKQEDKWEAKEVDCWNYNDL
 LDAMASNDRMARKNWQRIGTIKEAKNGYVSLVIRKIADLAVNNERPAFIVLEDLNTGPKR
 SRQKIDKSVYQKPELALAKKLNPLVDKNKRDEIGSPTKALQLTPPVNNYGDIENTKKQAG
 IMLYTRANYSQTDPATGWRKTIYKAGPEETTYKKDGKIKNKSVDQIIETPTDIGFDG
 KDYYFEYDKGEFVDEKTGEIKPKKWRLYSGENGKSLDRFRGEREKDKYEWKIDKIDIVKI
 LDDLFVNFDKNISLLKQLKEGVELTRNNEHGTGESLRFAINLIQQIRNTGNNERDNDFIL
 SPVRDENGKHFDSEYWDKETKGEKISMPSSGDANGAFNIARKGIIMNAHILANSKDL
 SLFVSDEEWDLHLNNKTEWKKOLNIFSSRKAMAKRKK

FIG. 67H

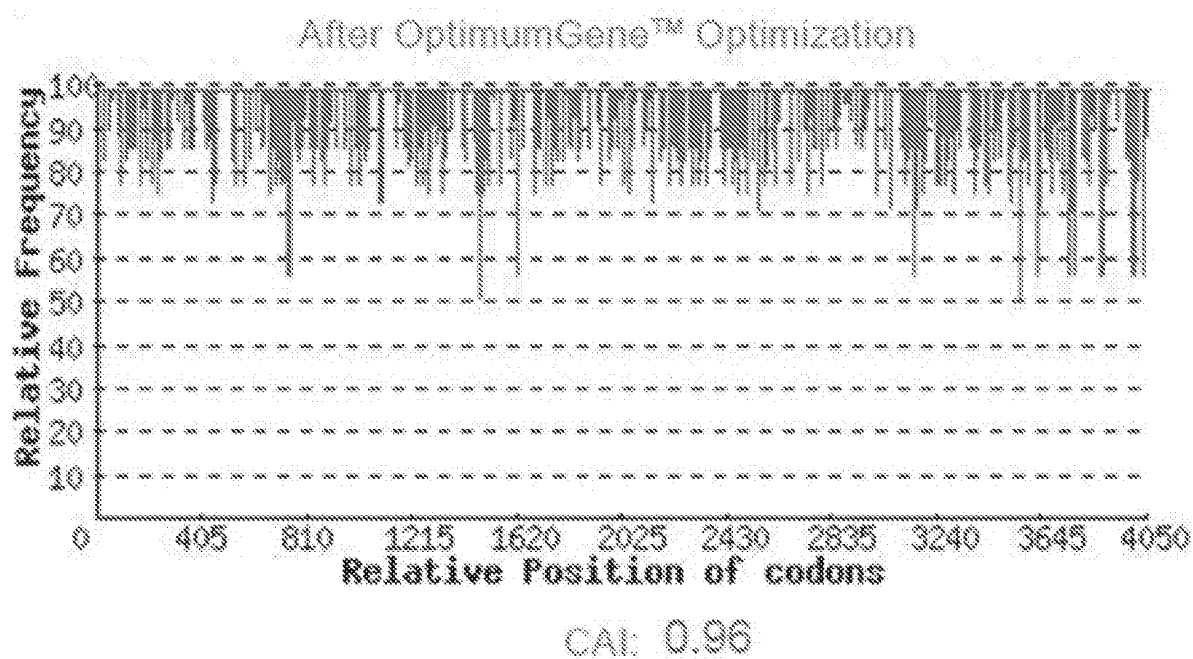


FIG. 68A

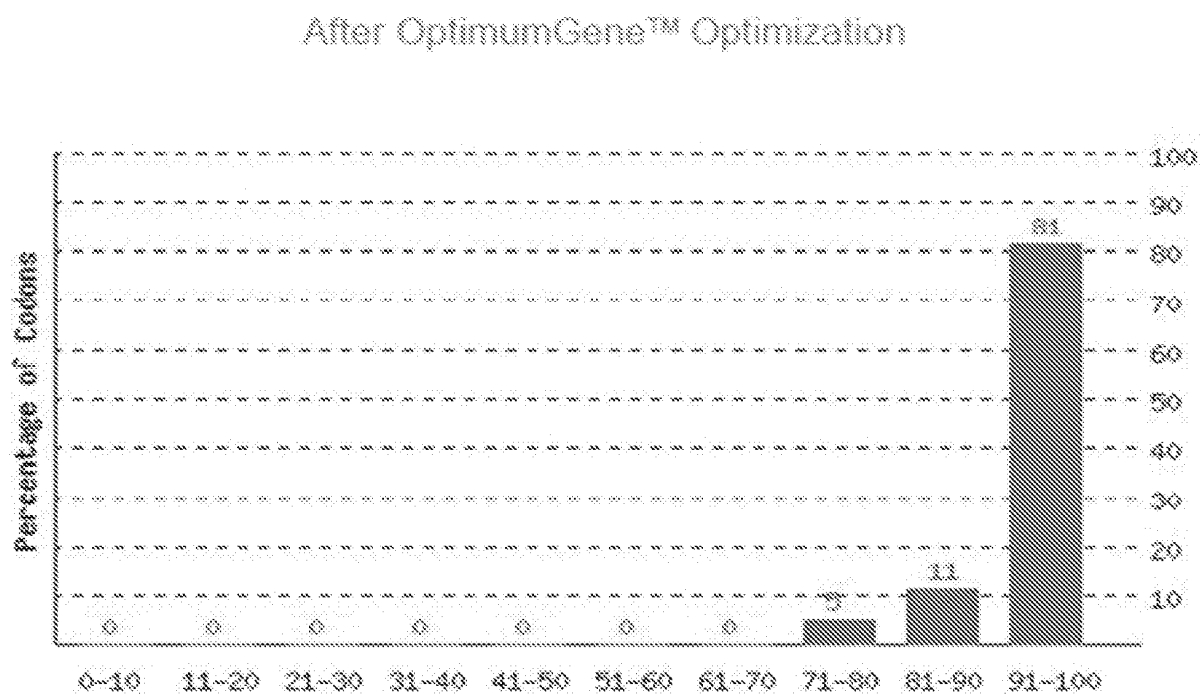


FIG. 68B

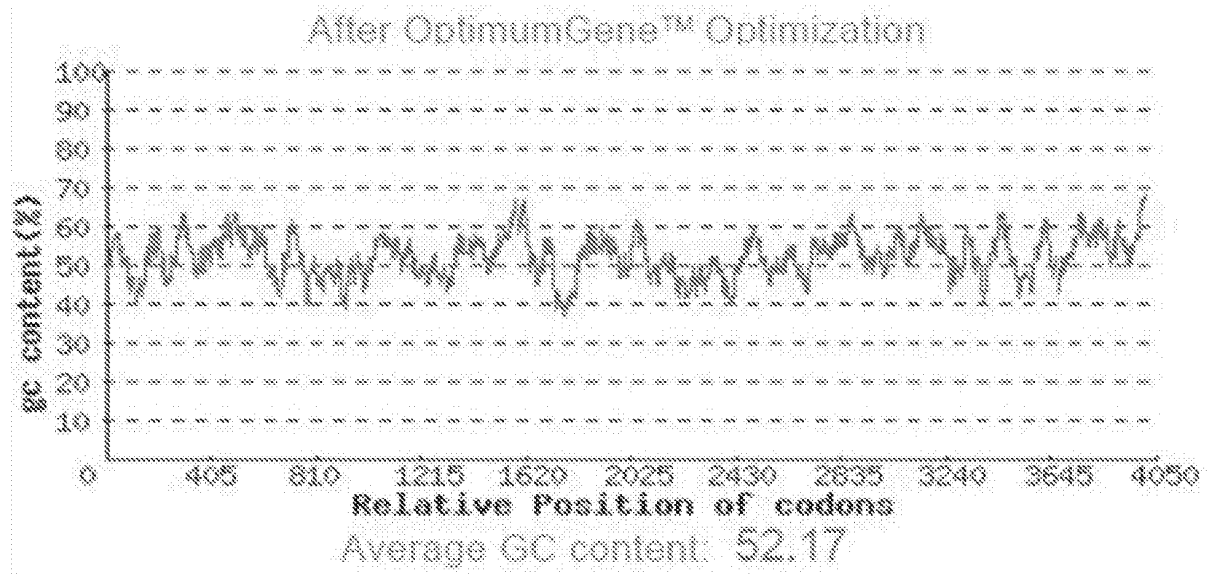


FIG. 68C

Restriction Enzymes	Optimized
* Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.	
BamHI(GGATCC)	1(4111)
EcoRI(GAATTC)	1(4201)
HindIII(AAGCTT)	0
BsmBI(CGTCTC)	0
Bsal(GGTCTC)	0
BbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(4175)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BsrGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0
CIS-Acting Elements	Optimized
Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0
Antiviral Motifs	Optimized
	0

FIG. 68D

After Optimization

Max Direct Repeat: Size:14 Distance:2001 Frequency:2
 Max Inverted Repeat: None
 Max Dyad Repeat: None

FIG. 68E

CGTACT

ATGGAGAACATCTTCGACCCAGTTTATGSSCAAGTACASDCTGTCCAAAGACDCTGAGATTTCGAGCTGAAAGCCCGTG
 GGCAGACAGAGGACTTCTGAAGATCAACAAGGTGTTTGAGAAAGATCAGACCATCGACGATAGCTACATCAG
 GCCAAGTTCTATTTTGAATTCCTGCAACAGAAAGTTTATGCAACCCCGCCCTGACCCTCGATAAGACATTCGAGCTG
 TCTTTCCAGAACTTTGCCCACTGCTGGAGAAAGCAGAAATAAGATCATCTTGATTAAGAAGAGAGAGATGGGGCCC
 CTGAGGAAACGCGACAGGAAACCCCTGGGCATCGATAAGCTTCAGAAAGGAGATCAATGACCGCGAGGATATCATC
 CAGAAGGAGAAAGAGAGATCTACAAGGACGTGCGCACCTCTGTCGATAACGAGGCTGAGTCTTGAGAAACCTAC
 TATCAGGAGCGCGAGGCTGACCGGCAAGAGATCACCTTCAGGCAAGGCCGACCTGAAGCAGAAAGGGCCCGGATTTT
 CTGACAGCCCGCCGATCTCGAGGCTGCTGAAGTATGAGTTCCCGGAGGAGAAAGGAGAAAGGTTTCAAGCCAG
 AACCCAGCCCTCCCTGTTCTGAGGAGAGAGAGAAATCTGAGCAGAACAGGTACATCTTCGACTCTTTTGATAAG
 TTGCCCCTTATCTGACCAAGTTTCAGCGAGACAAAGAGAAATCTGTACCGAGCAGACCGCACCAGGCACAGCAGTG
 GCCACCCGATCTGCGATAACTTTATCATCTTCCACGAGAAATACCAAGGTGTTCCGAGGACAGGATACAGAACAAT
 CACACAGACCTGGGCTTCGATGAGGAGAACATCTTTGAGATCGAGAGGTATAAGAAATGCTGCTGCGACGCGAG
 ATCGAGCACATCAAGAATGAGAAATAGCTACAACAAGATCATCGGCCGATCAATAAGAGATCAAGGATATCGG
 GACCAGAGGSCCAAGGATACCAAGCTGACAAAGTCCGACTTCCCTTTCTTTAGAACTTGAGATAAGCAGATCCTG
 GCGAGGCTGAGAAAGGAGAGCAGCTGATCGAGAAACCCCGGAGAAACCCGAGGAGGAGCTGCTGATCGAGCCG
 TTCAAGGAGTTTATCGAGAACATGAGGAGAGGTTCAACCCCGCCAGAAAGCTGATGAATGCCCTTCTGTAAAGCC
 GAGTTTGAGTCCGAGTACGAGCGCATCTATCTGAAGAAATAGGCGCATCAACACAATCTCCCGAGATGCTTCGTG
 TGTGACAGAGATTTTGAGCTGAAGCTGCCCTCAGCAGAGTCCAAGACAACTCTGAGAAAGATGAGCCAAAGGTG
 AAGAAAGTTCTCTCATCGCCGAGTCAAGAAAGCCGTGAGGAGGCTGACCGGATATCTTTAAGGCCGTGTT
 TACGACAAAGAGATCATCGCCCGAGGCGGCTCTAAGCTGAGAGGATTCCTGATCATCTGGAAGTACGAGTTTGAG
 TATCTGTTCCGGACATCGAGAGAGAGAACCGGCGAGAAAGCTGCTGGGCTATGATAGCTGCTGAAGATGCGCAAG
 CAGCTGGGCTCTTCCCAAGAGAGAGAGGCCCCCGAGAGGCAACCCCGCTGATCAAGAAATACCGCGAGCC
 GGCCTGGGCTCTTCCAGATGATGAAGTATTTTTCTCTGAGGATTAAGGATCGGAAGACACCCCGGCGAGCTG
 AGCACAAGTTTCTACGCGAGTATGACCGCTACTACAAGGATTTGAGTTTATCAAGTACTACAACGATTTTAGG
 AACTTCTCTACCAAGAGGCTTTGAGCGAGGATAAGATCAAGCTGAACCTTGAGAAATGGGCCCTCTGAGAGGCC
 TGGAGCGAGAACAAAGGATACCATTTCTAGGCGTGAATCTGAAGAAAGAGGCGGCTGTATCTGGGCTCATG
 CACAAGAACCAAGGAGCTGTTTCAGTCCATGGGCAATGCCAAGGGCGACACGCGCAATAGATACAGAGATG
 ATCTATAAGCGATCGCCGAGGCTCTATAGGATGTGCCCCGCTGCTGCTGACCGAGCAGAAAGGCTATGAGAAAG
 TTCAAGCCTTCCAGGAGATCTGAGATCAAGAAAGAGAAACCTTCAAGCGGAGAGCAAGAAATTTTCCCTG
 AGAGATCTGCAAGCCCTGATCGAGTACTATAGGAACTGCTCCCTCAGTACAGCAATTTGTTCTTTTATGACTTC
 CAGTTTCAGGATACCGGCAAGTACAGAAATATCAAGGATTTCAAGAGGATGTGCGAAGTACGGCTATAAGATC
 TCCTTTCCGACATCGAGGATGAGTATATCAATCAGGCGCTGAAAGAGGCAAGATGTAATCTGTTGAGGCTGAG
 AACAGGATATCTATAACCAAGAAATGGCTCCAAGGATCTGCACACACTGTACTTTGAGCAGATCTGTCTGCC
 GAGAACTCAATGACCCAGTGTTCAGGCTGTCTGGCATGCCCCAGATCTTTCAGCGGCGAGCCAGGCTGAAAGAA
 AGAGAGAAAGATCAACACAGAGAAATCAGTGTATCTGAGCAAGGGCGATAGAGCTACAAGTATAGGCGCTAC
 ACCGAGAAAGAGATCATGTTCCACATGAGCCTGCTGCTGAAACACAGGCAAGGGCGAGATCAAGCAGTTCAGTTT
 AATAAGATCATCAACAGAGGATCAGCTCTCTGACAAAGGATGAGGCTGAATGTGATCGGCATCATCGCGG
 GAGAGGAACTGCTGTACTATAGCTGCTGAGGAGAGATGAGGAGATCATCGAGCAGGCTCTCTGAGCAGGATC

FIG. 68F

MENIFDQFIGKYSLSKTLRFELKPVGKTEDFLKINKVFEKDQTIDDSYNQAKFYFDSLHQ
 KFIDAALASDKTSELSPQNFADVLEKQNKIILDKKREMGALRKEDKNAVGIDRLQKEIND
 AEDIIQKEKEKIYKDVRTLFDNEAESWKTYYYQEREVDGKKITFSKADLKQKGADFLTAAG
 ILKVLKYEFFEKEKEFQAKNQPSLFEVEKENPGQKRYIFDSFDKFAGYLTKFQQTCKNL
 YAADGTSTAVATRIADNFIIFHQNTKVFRDKYKNNHTDLGFDEENIFEIERYKNCLLORE
 IEHIKNENSYNKIIGRINKKIKEYRDQKAKDTKLTKSDPPFFKNLDKQILGEVEKEKQLI
 EKTREKTEEDVLIERFKEFIENNEERFTAACKLMNAFCNGEFESEYEGIYLNKKAINTIS
 RRWFVSDRDPFELKLPQQKSKNKSEKNKPKVKKFISIAEIKNAVEELDGDIFKAVFYDKKI
 IAQGGSKLEQFLVIWKYEFYLFREDIERENGEXLLGYDSCLKIAKQLGIFPQKEKEAREKA
 TAVIKNYADAGLGIFQMMKYFSLDDKDRKNTPGQLSTNFYAEYDGYKDFEPIKYNEFR
 NFITKKPFDEDKIKLNFENGALLKGWDENKEYDFMGVILKKEGRLYLGIMHKNHFKLFQS
 MGNAGKDNANRYQKMIYKQIADASKDVPRLLLTSEKKAMEKFKPSQEILRIKKEKTFKRES
 KNFSLRDLHALIEYYRNCIPQYSNWSFYDFQFQDTGKYQNIKEFTDDVQKYGYKISFRDI
 DDEYINQALNEGKMYLFEVVNKDIYNTKNGSKNLHTLYFEHILSAENLNDPVFKLSGMAE
 IFQRQPSVNEREKITTQKNQCILDKGDRAYKYRRYTEKKIMFHMSLVLTGKGEIKQVQF
 NKIINQRISSESDNEMRVNVIGIDRGEKNLLYYSVVKQNGEIIIEQASLNEINGVNVYRDKLI
 EREKERLKNRQSWKPVVKIKDLKKEGYISHVTHKICQLIEKYSAIVVLEDLNMRFKQIRGG
 IERSVYQQFEKALIDKLGVLVFKDNRDLRAPGGVLNGYQLSAPFVSFEKMRKQTGILFYT
 QAEYTSKTDPI TGFRKNVYISNSASLDKIKEAVKKFDAIGWDGKEQSYFFKYNPYNLADE
 KYKNSTVSKEWAI FASAPRIRRQKGEDGYWKYDRVKVNNEFEKLLKVWNFVNPKATDIKQ
 EIIKKEKACDLQGEKELDGRLRNFWHSFIYLFNLVLELRNSFSLQIKIKAGEVIAVDEGV
 DFIASPVKPPFTTPNPYIPSNLCWLAVENADANGAYNIARKGVMILKKIREHAKKDPEFK
 KLPNLFISNAEWDEAARDWGKYAGTTALNLDH

FIG. 68H

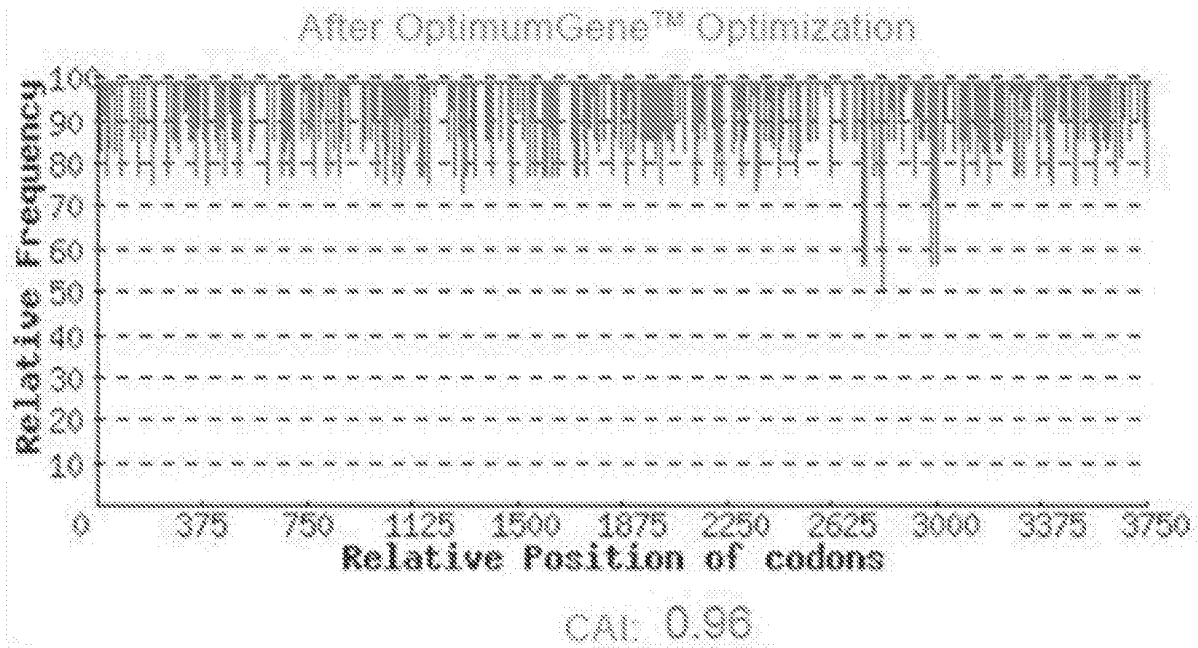


FIG. 69A

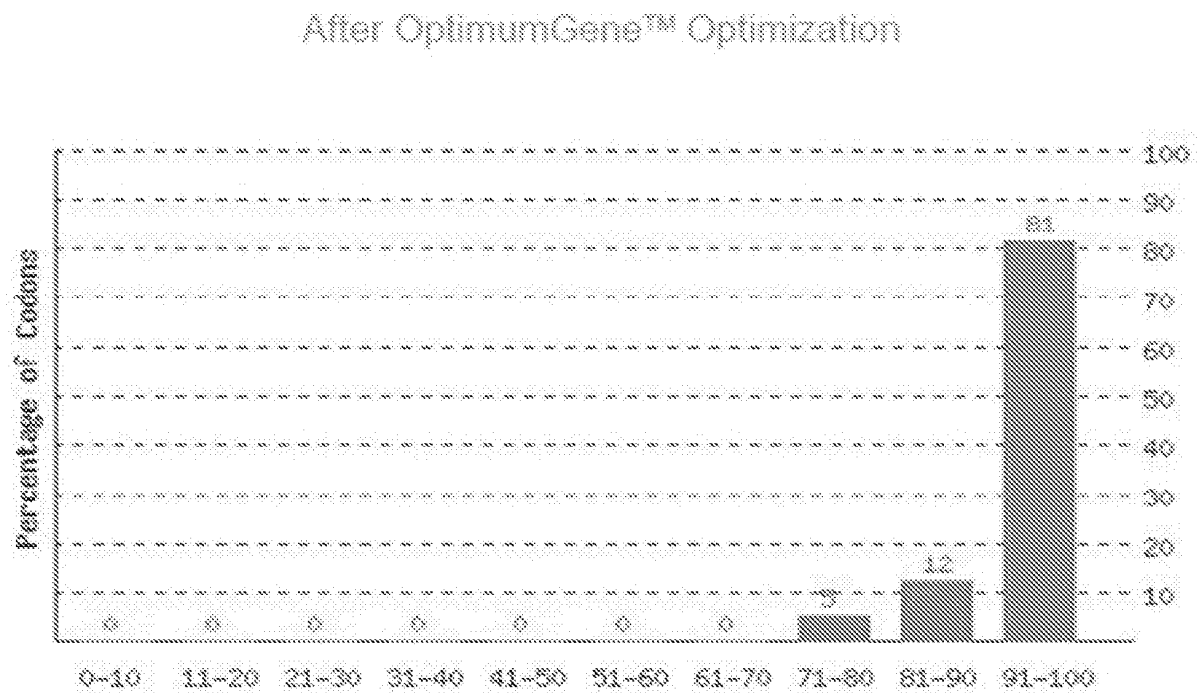


FIG. 69B

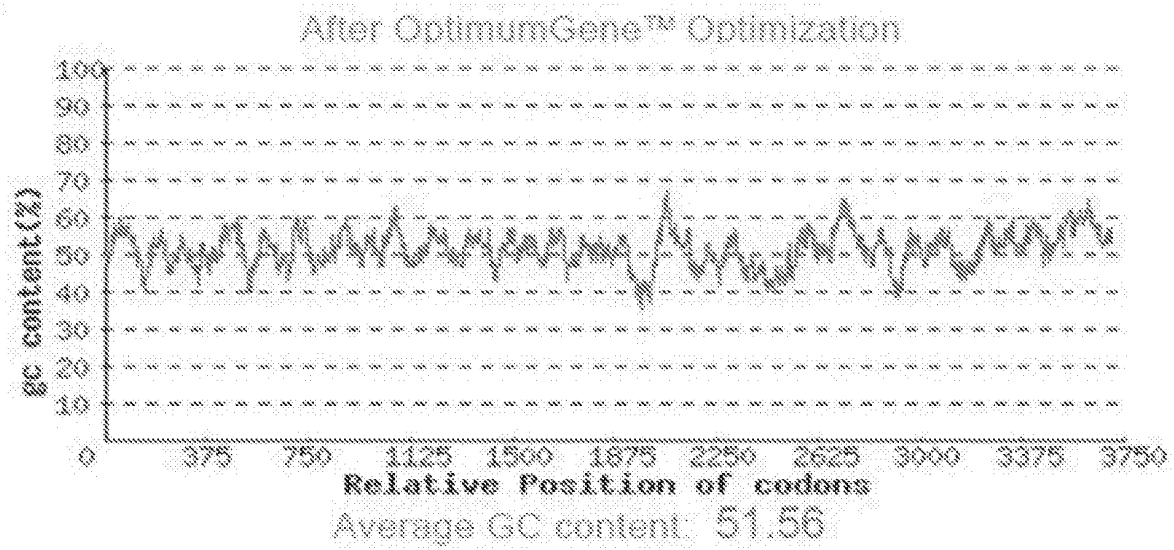


FIG. 69C

Restriction Enzymes	Optimized
* Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites	
BamHI(GGATCC)	1(3805)
EcoRI(GAATTC)	1(3895)
HindIII(AAGCTT)	0
BsmBI(CGTCTC)	0
BsaI(GGTCTC)	0
BbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(3869)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BsrGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0

CIS-Acting Elements	Optimized
Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0

Antiviral Motifs	Optimized
	0

FIG. 69D

After Optimization

Max Direct Repeat: Size:15 Distance:93 Frequency:2
 Max Inverted Repeat: Size: 12 Tm: 39.3 Start Positions: 842, 2867
 Max Dyad Repeat: None

FIG. 69E

697962

ATGCGAGACCCCTGTTTGGAGAACTTCACAAATCAGTACCCAGTGTCCAGAACCCCTCCGCTTTGAGCTKATCCCCCAAG
 GGCAAGACAANGKACTTCATCGAGCAGAAAGGCCCTGCTGAGAGAGGATGAGGACCCGGGCCCTAGAACTATAAGAAAG
 GTGAAGAACATCATCGATGAGTACCACAAGGACTTCATCGAGAAAGTCTCTGAATGAGCTGAGAGCTGGACGGGCTTG
 GAGAACTACAGACCCCTGTATCTGAAGCAGGAGAGGRCUATAAGGATAAGAAAGCCCTTTGACAAAGGAGAAAGGAG
 AACCTGCCCCAAGCAGATCCGCAATGCCCTTCGGAACCAATGAGAAAGTTTAAAGACACTGTTTCCGCAAGGAGCTGATC
 AAGAACGATCTGATGTCTTTCCGCTGCGAGGAGGACAGAGAGAAATGTGAAGGAGTTTGAGGGCTTCACCAATAC
 TTCACCGGCTTCCACCCAGAACCGCGGCCAATATGTACGTCGCGGATGAGAAAGGACAACAGCCCATCCGCGAGCAGGCTG
 ATCCACGAGAACCTGCCCCAAGTTTATCGACAATATCAAGTCTTCGAGAAAGATGAGAGAAAGGAGGCCCCCGAGCTG
 CTGTCTCTCTTTCAACCAGACCCCTGAAGGATATGAAGGAAGTGATCAAGGGCACCACACTGGAGCGAGATCTTTAGC
 CTGATTATTTCAACAAGACCCCTGACACAGAGCGGCATCGACATCTACAAATTCGCTGATCGGCGGCGAGAACCCCT
 GAGGGGGGCAAGACAAAGATCAAGGGGCTGAAACGAGTACATCAATACCGACTTCAACCCAGAAAGCAGACAGACAAAG
 AAGAGCGCGCAGCCCAAAGTTCAAGCAGCTGTATTAAGCAGATCCTGAGCGATAGGCAGAGGCTGTCTTTTATCGCC
 GAGGCTTCAAGAAACGACACCGAGATCTCGAGGGCCATCGAGAAAGTTTACGTAATGAGCTGCTGCACTTCAGC
 AATGAGGGCAAGTCCACAAACGTCCTGGAAGCCCATCAAGAAATGCCGTGTCTAACTCGAGAGCTTTAACTTGACC
 AAGATGTATTTCCGCTCCGGCGGCTCTCTGACAGACCTGAGCCCGAAGGTGTTTGGCGAGTGGAGCATCATCAAT
 AGAGCTCTGGACAACTACTATGECACCCATATCCAACTAGAGCCCGAGAGAGAGTCTGAGAACTACGAGGAGAGG
 AAGGAGAAAGTGGCTGAAGCAGGACTTCAAGCTGAGGCTGATCCAGACCGCCCATCGATGAGTACGACAACGAGGACA
 GTGAAGGGCAAGAACAGCGGCAAGTGTCTGCGATTATTTTCCCAAGTTCTGCGAAGATTAAGGAGACAGACCTG
 ATCCAGAAAGTGAACGAGGGCTACATCGCGTGAAGGATCTGCTCAATACACCCCTGTCTCTGAGAACGAGAGCTG
 GGCAGCAATANKKACGAGTGAAGCAGATCAGGGCCCTTTATGGATTCTATCATGACATCATGCGCTTCTGTGGCC
 CCCCCTGAGGCTGAAGGATACCGACAAGGAGAAAGGATEAGACATTTCTACTCCCTGTTTACACCTCTGTACGAGCCAC
 CTGACCCAGACAAATCGCCCTGTATAACAAAGGTGCGGAACTATCTGACCCAGAAAGCCTTACAGCACAGAGAGAGTC
 AAGCTGAACTTCGAGAACAGCACCCCTGCTGGGCGGCTGGGATCTGAATAAGGAGACAGACAAACAGGCCATCATC
 CTGAGGAGAGGATAAAGTGTACTATCTGGGCATCATGGAAGAGGACACAANTCGCATCTTTTGGAAAGTGGCCCAAG
 GCCGATAAGAAAGGACTTCTGCTACAGAGAGATGGTGTATAAGCTGCTGCTGGCGCCAACAGATGCTGCCAAAG
 GTGTTCTTTTCTCAGAGCAGAAATCCAGAGTTTACCCCTTCGCGCAAGCTGCTGGAGAACTACGCCAATGAGACA
 CACAAGAAAGGGCGATAATTTCAACCTGAATCACTGTACAAAGCTGATCGATTTCTTTAAGGACTCTATCAACAAG
 CACGAGGATTTGAAGAAATTTGACTTTAGGTTCAAGGCCACCTCCACCTACGCGGACCTGAGCGGCTTTTACCCAC
 GAGGTGGAGCCACAGGGCTACAAGATCTCTTTTCAAGAGTGGGCGGATTCCTTCATCGACGATCTGTTGAAGGAG
 GGCAGGCTGTACTGTTCCAGATCTATTAATAAGGACTTTTCCCCATTCTCTAAGGGCAAGGCCAACCTGCACACC
 CTGTACTGGAGGATGCTGTTTGTATGAGAACAAATCTGAAGGAGCTGTTGTATTAAGCTGAATGGCGAGGCGGAGGTG
 TTCTACCGCAAGAGAGGATTTGCGAGAGAGAACACCACAAATCCACAGGCGCAATGAGTCCATCATCAACAAGAAAT
 CCTGATAACCCAAAGGCCACAGCACCTTCAACTATGATATCGTGAAGGACAAAGAGATACACCATGACAAAGTTT
 CAGTTCCACATCCCAATCACAATGAACTTTAAAGGCCAGGGGCATCTTCAACATGAATCAGAGGGTGAATCAGTTC
 CTGAAGGCCAATCCGATATCAACATCATCGGATGACAGAGGCGAGAGGCCAAGCTGCTGTACTATGCGCTGATC
 AACCAGAAAGGCCAAGATCTGAAGCAGGATACCTTGAATGTGATCGCCAAAGAGAAAGCAGAAAGGTGGACTACCAC
 AATCTGCTGENTAGAAAGGAGGGCGAGCCGCAACCGCAAGGCAAGAGTGGGCGTGTATCGAGACAATCAAGGAG
 CTGAAGGAGGGCTATCTGTCCAGGTCATCCACAAGCTGACCGATCTGATGATCGAGAACAAATGCCATCATCTG

FIG. 69F

FIG. 89G

MQTLFENFTNQYPVSKTLRFELIPQGKTFDPIEQKGLLKKDEDRAEKYKKVKNI IDEYHK
 DFIEKSLNGLKLDGLEKYKTLYLKQEKDDKDKKAFDKEKENLRKQIANAFRNNEKFKTLF
 AKELIKNDLMSFACEEDKKNVKEPEAFTTYFTGFHQNRANMYVADEKRTAIASRLIHENL
 PKFIDNIKIFEKMKKEAPELLSPFNQTLKDMKDVIKGTTLLEEIPSLDYFNKTLTQSGIDI
 YNSVIGGRTPEEGKTKIKGLNEYINTDFNQKQTDKKKRQPKFKQLYKQILSDRQSLSFIA
 EAFKNDTEILEAIEKPYVNELLHFSNECKSTNVLDIAKNAVSNLESFNLTKMYFESGASL
 TDVSRKVPGEWSI INRALDNYIATTPYIKPREKSEKYEERKEKWLKQDFNVSLIQTIDE
 YDNETVKGKNSGKVIADYFAKFCDDKETDLIQKVNIEGYIAVKDLLNTPCPENIEKLGSKND
 QVKQIKAFMDSIMDIMHFVRPLSLKDTDKEKDETFYSLFTFLYDHLTQTIALYNKVRNYL
 TQKPYSTEKIKLNFENSTLLGGWDLNKETDNTAII LRKDNLYYLGIMDKRHNRI FRNVPK
 ADKKDFCYEKMVYKLLPGANKMLPKVPFSSQSRIQEFTPSAKLLENYANETHKKGDNFNLN
 HCHKLIDFPKDSINKHEDWKNFDFRFSATSTYADLSGFYHEVEHQGYKISFQSVADSFID
 DLVNEGKLYLPQIYNKDFSPFSKKGKPNLHTLYWKMLFDENNLKDVVYKLNGEAEVFYRKK
 SIAEKNTTIHKANESI INKNPDNPKATSTFNYDIVKDKRYTIDKFQFHIPITMNFKAEGI
 FNMNQRVQNQFLKANPDINIIGIDRGERHLLYALINQKGKILKQDTLNVIANEKQKVDYH
 NLLDKKEGDRATARQEWGVETIKELKEGYLSQVIHKLTDLMIENNAIIVMEDLNFEGFKR
 GRQKVEKQVYQKFEKMLIDKLNLYLDKNKKANELGGLLNAFQLANKPESFQKMGKQNGFI
 FYVPAWNTSKTDPATGFIDFLKPRYENLNQAKDFFFEKFDSIRLNSKADYFEPAPDFKNFT
 EKADGGRTKWTVCCTTNEEDRYAWNRLNNNRGSQEKYDITAEKSLFDGKVDYKSGKDLKQ
 QIASQESADFFKALMKNLSITLSLRHNNGEKGDNEQDYIILSPVADSKGRFFDSRKADDDM
 PKNADANGAYHIALKGLWCLEQISKTDLKEVKLAISNKEWLEFVQTLKG

FIG. 69H

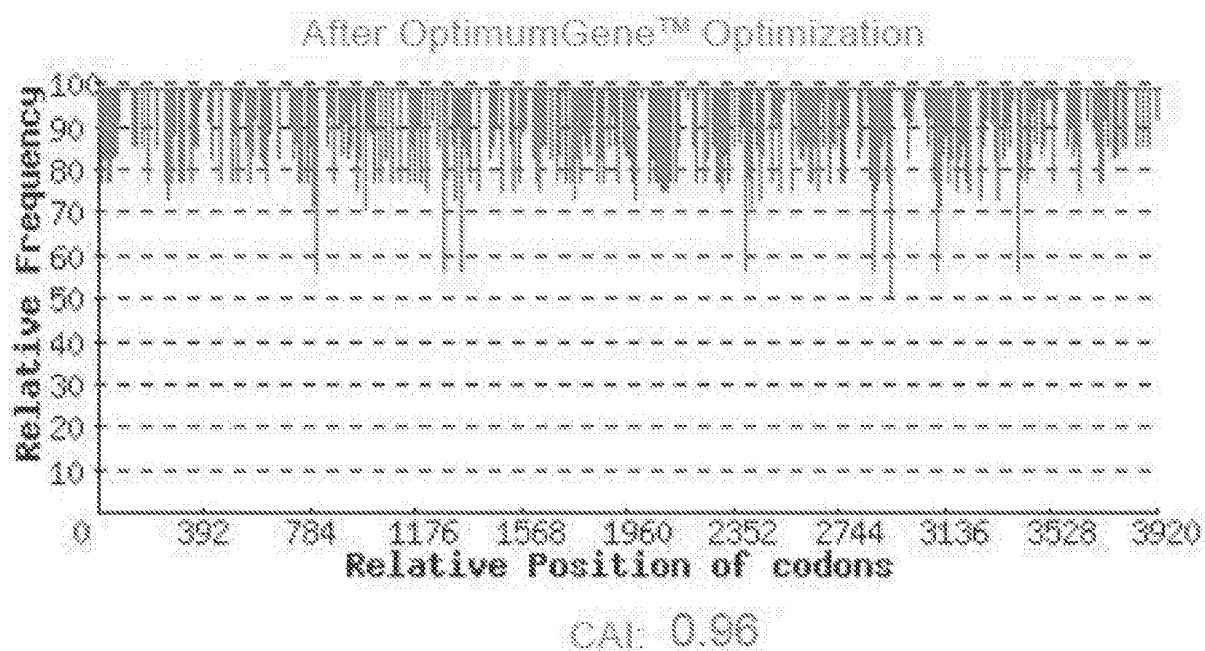


FIG. 70A

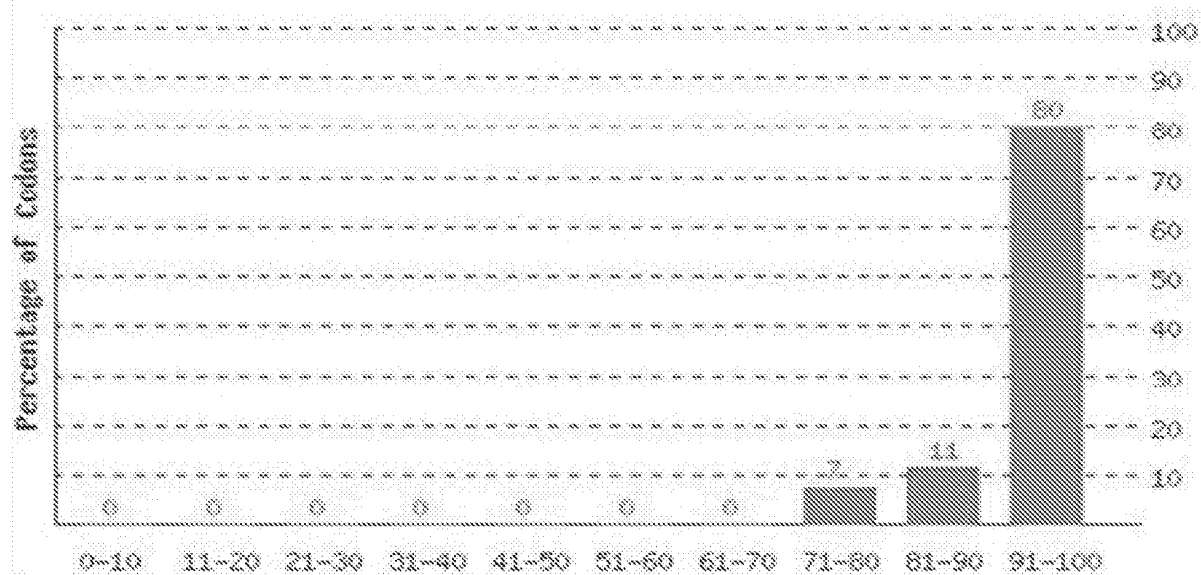


FIG. 70B

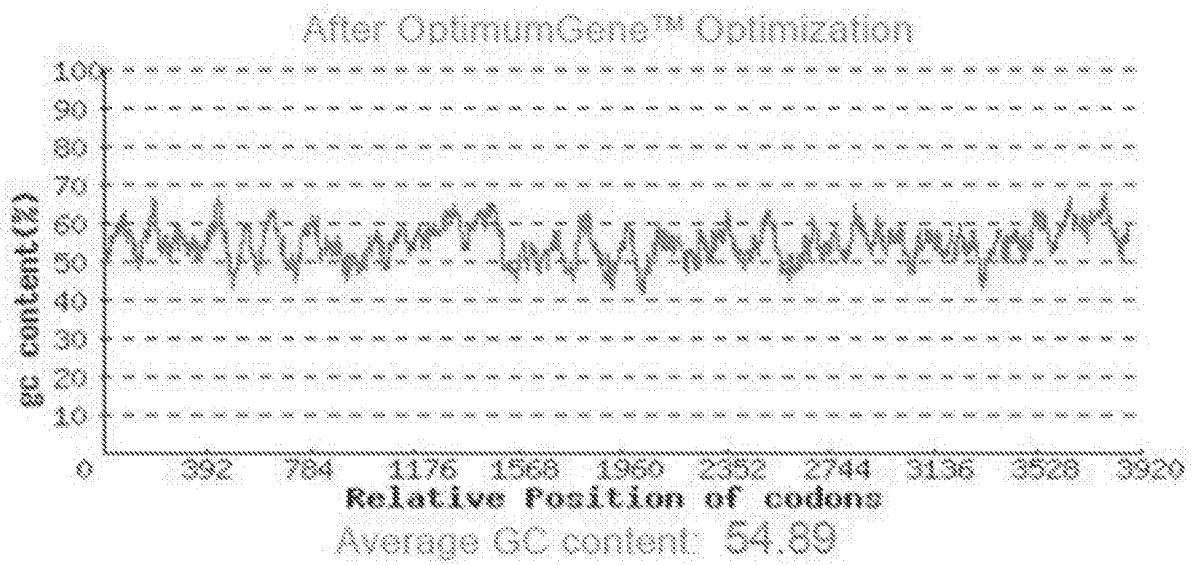


FIG. 70C

Restriction Enzymes Optimized

* Green, filtered sites; Blue, checked sites (not filtered); Red, kept sites.

BamHI(GGATCC)	1(3976)
EcoRI(GAATTC)	1(4066)
HindIII(AAGCTT)	0
BsmBI(CGTCTC)	0
BsaI(GGTCTC)	0
BbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(4040)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BsrGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0

CIS-Acting Elements Optimized

Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0

Antiviral Motifs Optimized

0

FIG. 70D

After Optimization

Max Direct Repeat: Size:15 Distance:36 Frequency:2
 Max Inverted Repeat: None
 Max Dyad Repeat: None

FIG. 70E

GenBank:

ATGACACAGTTTCGAGGCTTTTACCAACCTGTATCAGGTTGAGCAAGACACTGCGGTTTGGAGCTGATCCACACAGGAC
 AAGACCCGTGAAGCAGCTCCAGGAGCAGGCTTCATCGAGGAGGACAGGCCCCCAATGATCACTACAGGAGCTG
 AAGCCCATCATCGATCGATCTACAAGACCTATGCCCCAAGTGGCTGCGAGCTGGTGCAGCTGGATTGGGAGAAC
 CTGAGCGCCCGCATCGACTCCTATAGAAAGGAGAAACCGAGGAGACAGGAGACGCCCCGATCGAGGAGCAGGCTC
 ACATATCGCAATGCCATCCACGACTACTTCATCGGCGCGACAGACAACTGACCGATGCCATCAATAAGAGACAC
 GCCGAGATCTACAGGGGCTGTTCAGGCGGAGCTGTTAATGGCAGGCTGCTGAAGCAGCTGGGCACCGTGACC
 ACAGCCGAGCAGGAGGAGCGCCCTGCTGCGGAGCTTGGACAAAGTTTACAACTTACTTCTCCGCTTTTATGAGAAC
 AGGAGAAACGTGTTCAAGCCGAGGATATCAGCAGACGCTATCCACACCCCATCGTGCAGGACACTTCCCTCAAG
 TTTAAGGAGATTTGTCACTCTTTCACAGCGCTGATCAGCGCCGTGCCCCAGCCCTGCGGGAGCACTTTGAGAACCTG
 AAGAAAGCCATCGGCATCTTGTGAGCAGCTCCATCGAGGAGGCTGTTTCTCTTCCCTTTTATAACAGCTGCTG
 ACACGAGCCGAGATCGACTGTATAGCCAGCTGCTGAGGAGATCTCTCGAGGAGGAGGACCGAGAGAGATCGAG
 GGCTTGAAAGAGGCTGTAATCTGGCCATCCAGAGGATGATGAGACAGCCACATCATCGCTTCCCTGCCACAC
 AGATTCTATCCCCCTGTTTAAAGCAGATCCTGTCCGATAGGAAACACCTGTCTTTCTATCTGAGGAGTTTAAAGGC
 GACGAGGAAGTGATCCAGTCCCTCTGCAAGTACAGACACTGCTGAGAAAAGAGAACGCTGCTGGAGACAGCCGAG
 GCGCTGTTTAAAGAGCTGAAACAGCATCGACCTGACACACATCTTCATCAAGCCACAAGAGCTGGAGACAAATCAGC
 AGCGCCCTGTGCGGCCACTGGGATACACTGAGGAATGCGCTGTATGAGCGGAGAACTCTCGAGCTGACAGGCAAG
 ATCACCAGTCTGCGCAAGGAGAGGCTGAGCGCAGGCTGAGGACAGGAGATATCAAGCTGAGGAGATCATCTCT
 GCCGCAAGCAGGAGCTGAGCGAGGCTTCAAGCAGAAACCGAGCGAGATCCTGTCCACCCACACGCGCGCCCTG
 GATCGCCACTGCTTACAGCCCTGAAAGAGCAGGAGGAGAGAGAGATCTGAACTCTCAGCTGGACAGCGCTGCTG
 GGCTGTACCACTGCTGAGCTGCTTTGCGGTGGATGAGTCCAAAGAGGTGGAACCCGAGTTCTCTGCCCCGCTG
 ACGGCGATCAGGCTGGAGATGGAGGCTTCTCTGAGCTTCTACAACAGGCGCAGAAATTATGCCACCAAGAGGCGC
 TACTTGGTGGAGAGTTCAAGCTGAACTTTGAGATGCTTACACTGGCTCTGCTGGGACCTGAATAGGAGAGAG
 AACATAGGCCCATCTGTGTTGTAAGAGAGGCGCTGTACTATCTGGGCTCATGCCAAAGCAGAGGCGCAGGTAT
 AAGGCGCTGAGCTTGGAGCCACAGAGAAAACCGAGGAGGCTTTGATAGATGTACTATGACTACTTCCCTGAT
 GCCGCCAAGATGATCCCAAAGTGACGCAACCGAGCTGAAGGCGGTGACAGCCCACTTTGAGACCCACACAGCCCC
 ATCTGCTGTCCAGCAATTTTCATCGAGCTCTGGAGATCAAAAGGAGATCTACGACCTGAACATCTCTGAGAG
 GAGCCAAAGAGTTTCAGACAGCTTACGCCAGGAAACCGGCGACAGAGGCGCTACAGAGAGGCGCTGTGCGAG
 TGGATGAGCTTCACAGGAGATTTTCTGTCCAGATATACCAAGACAGCTCTATGATCTGTCTAGGCTGCGGCA
 TCTCTCTAGTATAAGGAGCTGGGCGAGTACTATGCCGAGCTGAATCCCTGCTGTACACATCAGCTTCCAGAG
 ATCGCCGAGAGGAGATCATGATGCGCTGGAGACAGGCAAGCTGTACTGTTCCAGATCTATAACAGGACTTT
 GCCAAGGCGCCACACCGGCAAGCTTAATCTGCACACACTGTATTGGACCGGCTGTTTTCTCCAGAGAACCTGCGC
 AAGACAGCATCAAGCTGAATGGCCAGGCGGAGCTGTTCTACCGCTTAAGTCCAGGATGAAGAGGATGGCACAC
 CGGCTGGGAGAGAGATGCTCAACAAGAGCTGAAGGATCAGAAAACCTCAATCCCGACACCTGTACAGGAG
 CTGTACGACTATGTGAATCACAGACTGTCCACAGACTGTCTGATGAGGCGAGGCGCTGCTGCCAACCTGATC
 ACCAAGGAGGTGTCTCAGAGATCATCAAGGATAGGCGCTTTACAGCGACAGTTCTTTTTCCAGCTGCTATC
 ACGTGAAGTATCAGGCGCCCAATTCCTCATAGTTCAACAGAGGTTGAATGCCATCTGAAGGAGCACCCC
 GAGACACTATCATCGGATCATCGGCGGAGAGAAACCTGATCTATATCACAGTGATCGACTCCACCGGCGAG
 ATCTTGGAGCAGCGGAGCTGACACCATCCAGCAGTTTGATTACAGAGAGAGCTGGACACAGGAGAGAGAG

FIG. 70F

FIG. 70G

MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQQFIEEDKARNDHYKELKPIIDRIYKT
 YADQCLQLVQLDWNLSAAIDSYRKEKTEETRNALIEEQATYRNAIHDFYIGRTDNLTD
 INKRHAETIYKGLFKAELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYENRKNVF
 SAEDISTAIPHRIVQDNFPPKFKENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSIEEV
 PSFPFYNQLLTQTQIDLYNQLLGGISREAGTEKIKGLNEVLNLAIQKNDETAHI IASLPH
 RFIPLFKQILSDRNTLSPILEEFKSDEEVIQSFCKYKTLRLNENVLETAEALFNELNSID
 LTHIFISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSAKEKVQRSLKHEDINL
 QEII SAAGKELSEAFKQKTSEILSHAHAAALDQPLPTTLKQEEKEILKSQLDSSLGLYHL
 LDWFAVDESNEVDPEFSARLTGKLEMEPSLSFYNKARNYATKKPYSVEKFKLNFQMFTL
 ASGWDVNKEKNNGAILFVKNGLYLGLIMPKQKGRYKALSFEPTKTSSEGFDMYYDYFPD
 AAKMIFKCSTQLKAVTAHFQTHHTPILLSNNPIEPLITKEIYDLNNPEKEPKKFQYAYA
 KKTGDQKGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYH
 ISFQRIAEKEIMDAVETGKLYLPQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKTSIK
 LNGQAELFYRPFKSRMKRMAHRLGEKMLNKKLKDQKTPIDOTLYQELYDYVNHRLSHDLSD
 EARALLPNVITKEVSHEI IKDRRFTSDKFFPHVPITLNYQAANSPSKFNQRVNAYLKEHP
 ETPIIGIDRGERNLIYITVIDSTGKILEQRSLNTIQQFDYQKKLDNREEKERVEARQAWSV
 VGTIKDLKQGYLSQVIHEIVDLMIHYQAVVVLENLNFQFKSKRTGIAEKAVYQQFEKMLI
 DKLNCLVLKDYPAEKVGGVLNPHYQLTDQFTSFAKMGTQSGFLFYVPAPYTSKIDPLTGFV
 DPFVWKTIKNHESRKHFLEGFDLHYDVKTGDFILHFKMNRNLSFQRGLPQGFMPAWDIVF
 EKNETQFPAKGTFFIAGKRIVPVIENTHRFTGRYRDLYPANELIALLEEKGIVFRDGSNIL
 PKLLENDDESHAITMVALIRSVLQMRNSNAATGEDYINSPVRDLNGVCFDSRFQNPWPM
 DADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYIQELRN

FIG. 70H

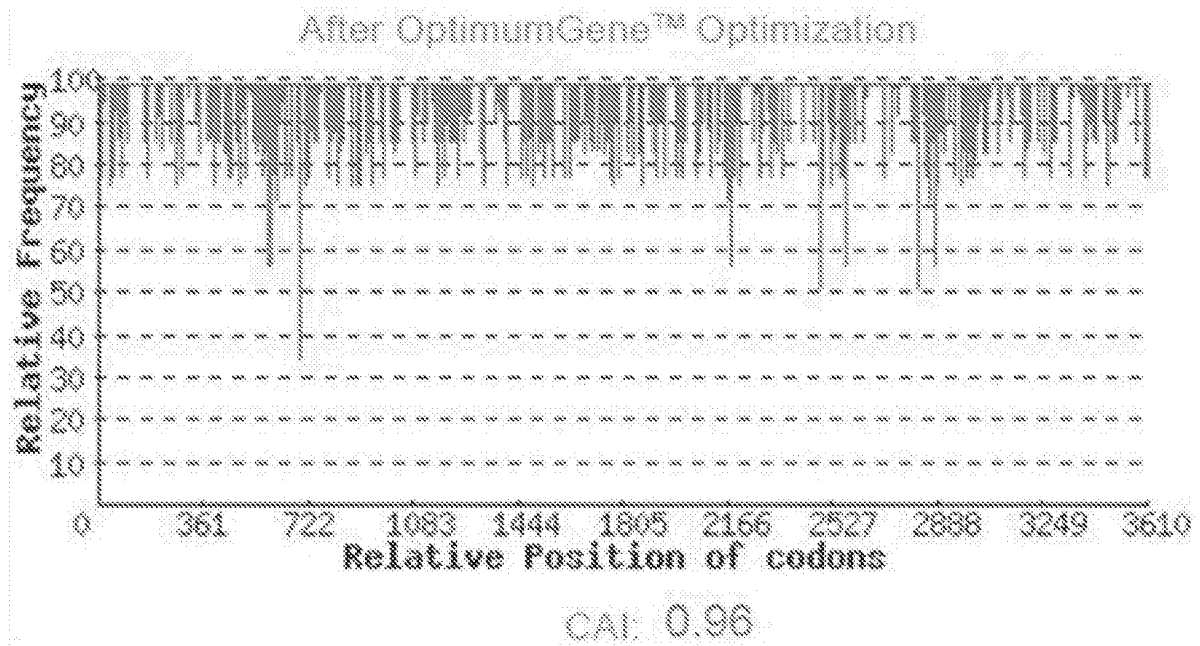


FIG. 71A

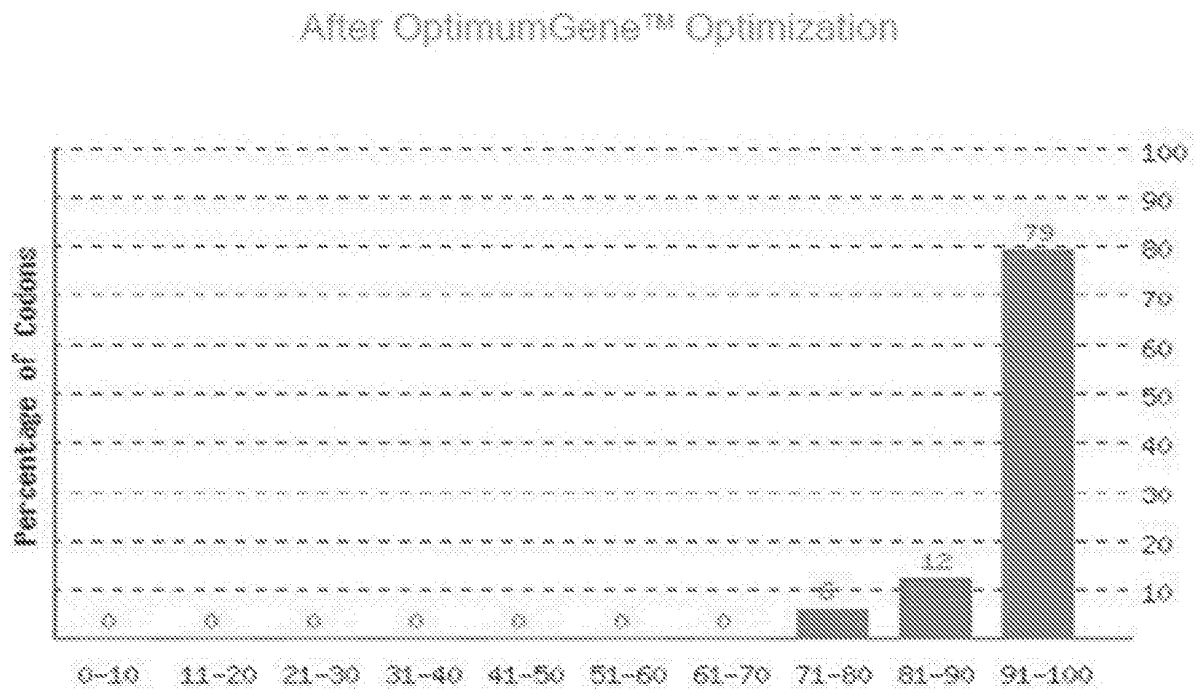


FIG. 71B

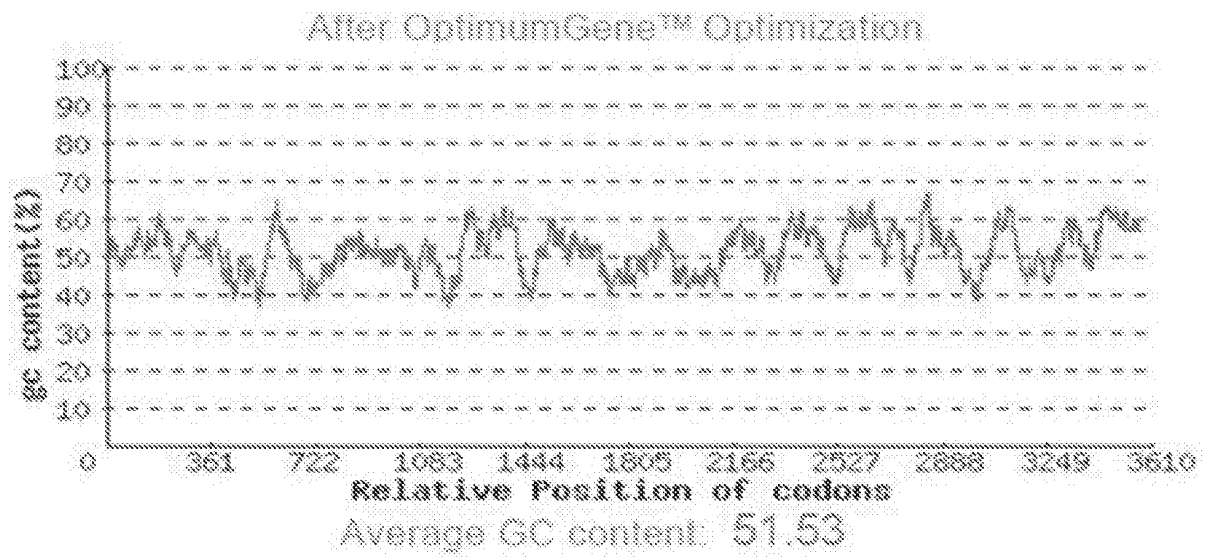


FIG. 71C

Restriction Enzymes	Optimized
<i>* Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.</i>	
BamHI(GGATCC)	1(3673)
EcoRI(GAATTC)	1(3763)
HindIII(AAGCTT)	0
BsmBI(CGTCTC)	0
Bsal(GGTCTC)	0
EbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(3737)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BsrGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0

CIS-Acting Elements	Optimized
Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0

Antiviral Motifs	Optimized
	0

FIG. 71D

After Optimization

Max Direct Repeat:	Size:17 Distance:1800 Frequency:2
Max Inverted Repeat:	None
Max Dyad Repeat:	None

FIG. 71E

FIG. 71F

ATGTACTATGAGTCCCTGACCAAGCAGTACCCCGTGTCTAAGACAAATCCGGAATGAGCTGATCCCTATCCGCAAG
 ACACTGGATAACATCCGCTAGAACATATCTCTGGAGAGGACCTGGAAGCGGAAGCAGAACTACGAGCACCTGAAG
 GGCATCTCTGGATGASTATCAACAAGCAGCTGATCAACGAGGCCCCCTGGACAATTCACCCCTGCCATCCCTGAAGATC
 GCCGCGGAGATCTACCTGAAGAATCAGAAGGAGGTGTCTGACAGAGAGGATTTCAACAAGACACGAGACCTGCTG
 AGGAAGGAGGTGCTGGAGAGGCTGAAGGCCACAGAGAACTTTACCAAGATCGGCAAGAAGGACATCCTGGATCTG
 CTGGAGAAAGCTGCCCTTCCATCTCTGAGACGATTACAAATCCCTGGAGAGCTTCCGCACTTTTACACCTATTTT
 ACATCCTACAACAAGGTGCGGAGAAATCTGTATTCTGTATAAGGAGAAGAGCTCCACAGTGGCTTACAGACTGATC
 AACGAGAATTTCCCAAAGTTTCTGGACAATGTGAAGAGCTATAGGTTTGTGAAAACCGCAGGCAATCCTGGCAGAT
 GGCCTGGAGAGGAGGAGGACAGGACTCCCTGTTTCTGCTGTAGACATTTCAACAAGACCCCTGACACGAGACCGCATC
 GATACCTACAATTTCTCAAGTGGCAAGATCAACTCTAGCATCAATCTGTATAACAGAGAAGATCAGAAGGCCAAT
 GGCCTTCAGAAAGATCCCAAGATGAAGATGCTGTATAAGCAGATCCTGTCCGATAGGAGAGAGTCTTTCTATCGAC
 GAGTTTCAGAGCGATGAGGTGCTGATCGACAACCGTGGAGTCTTATGGCAGCGGTGCTGATCGAGTCTCTGAAGTCC
 TCTAAGCTGAGCGCTCTTTTCTGATGCCCCCTGAGAGAGTCTAAGGGCAAGAACGTGTACGTGAAGAATGACTTGGCC
 AAGACAGCCATGAGCAACATCGTGTTCGAGAATTGGAGGAGCCTTTGACGATCTGCTGAACGAGAGTACGAGCTG
 GGCACCGAGAACAGAAGAAGGAGAGATTAAGTATTTGAGAGAGCCGACAGAGAGCTGAGAGAGAAATAAGAGCTAC
 TCCCTGGAGCACCTGTGCAACCTGTCCGAGGATTCTTGTAACTGATCGAGAATTATATCCACCAGATCTCCGAC
 GATATCGAGAATATCATCATCAACAATGAGACATTTCTGCCCATGCTGATCAATCAGGACAGACAGGTCCCGCAAG
 CTGGCCAGAAACCGGAAGGCCGTGAAGGCCATCAAGGACTTTCTGGATTCTATCAAGGTGCTCGAGCGGAGCTG
 AAGCTGATCAACAGCTCCGCCAGGAGCTGAGAGAGGATCTGATCGTGTACTCTGCCACGAGAGGCTGCTGGTG
 GAGCTGAAGCAGGTGAGACAGCTGTATAACATGACCAGAAATTATCTGACAAAGAAGCCTTTCTCTACCGAGAAG
 GTGAAGCTGAACCTTTAATCCAGGACACTGCTGAACCGCTGGGATCGGAATAAGGAGACAGACAACCTGGGCGTG
 CTGCTGCTGAAGGACCGCAAGTACTATCTGGGCATCATGALCACAAGCCCAATAAGGCCCTTCTGGAATCCCTCT
 GTGGCCAGGACCGAGAAGGTGTTTAAGAAAGGTGATTAAGAGCTGCTGCCAGTGCCCAACCGAGATGCTGCCAAG
 GTGTTCTTTGCCAAGAGCAATATCGACTTCTATAACCCCTCTAGCGAGATCTACTCCAATTATAAGAAGGCCACC
 CACAAGAAAGGCCAATATGTTTTCCCTGGAGGATTGTCAACCTGATCGACTTCTTTAAGGAGTCTATCGGCAAG
 CACGAGGACTGGAGCAAGTTCCGGCTTTAAGTTGAGCGATACAGGCTCCTACAAAGACATCTCCGAGTTCTATCGC
 GAGGTGGAGAGGAGAGGCTACAAGCTGACCTATACAGACATCGATGAGACATACATCAATGATCTGATCGAGCGG
 AACGAGCTGTACCTGTTCCAGATCTATAATAAGGACTTTAGCATGTACTCCAGGGCAAGCTGAACCTGACACA
 CTGTATTTTCATGATGCTGTTTGTATCAGCGCAATATCGAAGACCTGTTGTATAAGCTGAACCGGAGAGGAGAGGTG
 TTCTATAGGCCAGGCTCCATCTCTGAGGACGAGCTGATCATCCACAAGGCCGCGAGGAGATCAAGAACAGGAAT
 CCTAACCGGGCCAGAACCAAGGAGACAAACACCTTCAGCTACGACATCTGGAAGGATAAGCGGATATAGCAGGAT
 AAGTTTACCTGACATCCCATCAAGATGAAGTTCGCGGTGGATGAGGTGAAGCGGTTCAACGAGCGCGTGAAC
 AGCGCCATCCGATCGATGAGAATGTGAACGTGATCGGCATCGACCGGGCGAGAGAAATCTGCTGTACGTGGTG
 GTCATCGACTCTAAGGGCAACATCTCTGAGGAGATCTCCCTGAAGCTTATCATCAATAAGGAGTACGACATCGAG
 ACAGATTATCACGCACTGCTGGATGAGAGGGAGGGCGGAGAGATAAGGCCCGGAAGGACTGGAACACCGTGGAG
 AATATCGAGGAGCTGAAGGCCGCTACCTGAGCCAGGTTGTGAACGTGTTGCGCAAGCTGCTGAGATATAAT
 GCCATCATCTGCTGGAGGAGCTGAACCTTGGCTTCAGAGAGGGCCGCCAGAGGTGGAGAGGCGGTTTACAG
 AAGTTTCAGAGAGATGCTGATCGATAAGCTGAATTACTGTTTCATCGACAAGAGCCGCGAGGAGACATCCCTAAG

FIG. 71F

FIG. 71G

MYYESLTKQYFVSKTIRNELIPIGKTLDNIRQNNILESDVKKQNYEHVKGILDEYHKQL
 INEALDNCTLPFLKIAAEIYLKNQKEYSDREDFNKTQDLLKEVVEKLLKAHENFTKIGKK
 DILDLLLEKLPSISEDDYNALESFRNFYTYFTSYNKKVRENLYSDKEKSSTVAYRLINENFP
 KFLDNVKSYSRFVKTAGILADGLGEEEEQDSLFIIVETFNKTLTQDGDITYNSSQVGKINSSIN
 LYNQKNQKANGFRKIPKMKMLYKQILSDREESFIDEFQSDDEVLIIDNVESYGSVLIESLKS
 SKVSAPFDALRESKGNVYVKNDLAKTAMSNIVFENWRTFDDLLNQEYDLANENKKKDDK
 YFEKRQKELKKNKSYSLEHLNLSSEDSNLIENYIHQISDDIENIIINNETFLRIVINEH
 DRSRKLAKNRKAVKAIKDFLDSIKVLERELKLINSSGQELEKDLIVYSAHSELLVELKQV
 DSYLNMTRNYLTCKPFPSTEKVKLNPNRSTLLNGWDENKETDNLGVLLLKDGKYYLGIMNT
 SANKAFVNPPFAKTEKVFKKVDYKLLPVFNQMLPKVFFAKSNIDFYNPSSEIYSNYKKGT
 HKKGNMFSLEDCHNLIDFFKESISKHEDWSKFGFKFSDTASYNDISEFYREVEKQGYKLT
 YTDIDETYINDLIERNELYLFQIYNKDPSMYSKGLNLHTLYFMMLFDQRNIDDVVYKLN
 GEAEVFYRPAISISEDELIHKEAGEEIKNKNPNRARTKETSTFSYDIVKDKRYSKDKFTLH
 IPITMNFVGVDEVKRFNDVNSAIRIDENVNVIGIDRGERNLLYVVVIDSKGNILEQISLN
 SIINKEYDIETDYHALLDEREGGRDKARKDWNTVENIRDLLKAGYLSQVNVVAKLVLYKYN
 AIIICLEDLNFQFKRGRQKVEKQVYQKFEKMLIDKLNLYLVIDKSREQTSPKELGGALNALQ
 LTSKFKSPKELGKQSGVIYYVPAYLTSKIDPTTGPFANLFYMKCENVEKSKRPFDFGDFIR
 FNALENVFEFGFDYRSFTQACGINSKWTVCNTERI IKYRNFDKNNMFDEKVVVVVTDEM
 KNLFEQYKIPYEDGRNVKDMIISNEEAEFYRRLYRLLQOTLQMRNSTSDGTRDYIISPVK
 NKREAYFNSELSDGSPKADANGAYNIARKGLWVLEQIRQKSEGEKINLAMTNAEWLEY
 AOTHLL

FIG. 71H

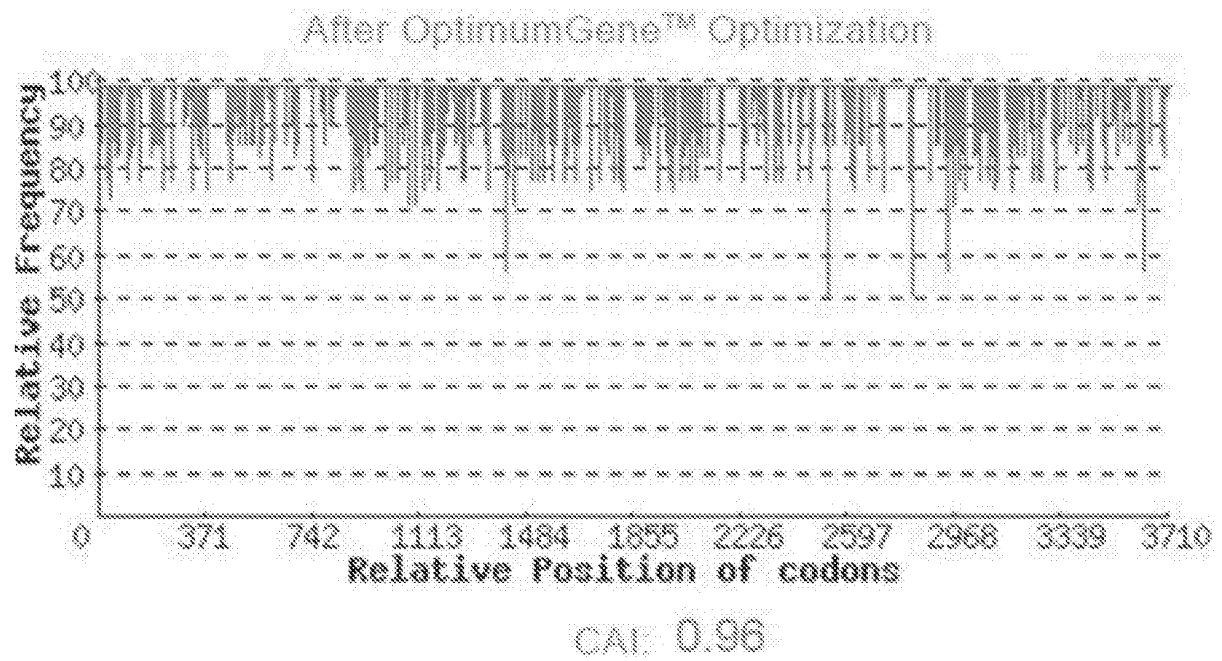


FIG. 72A

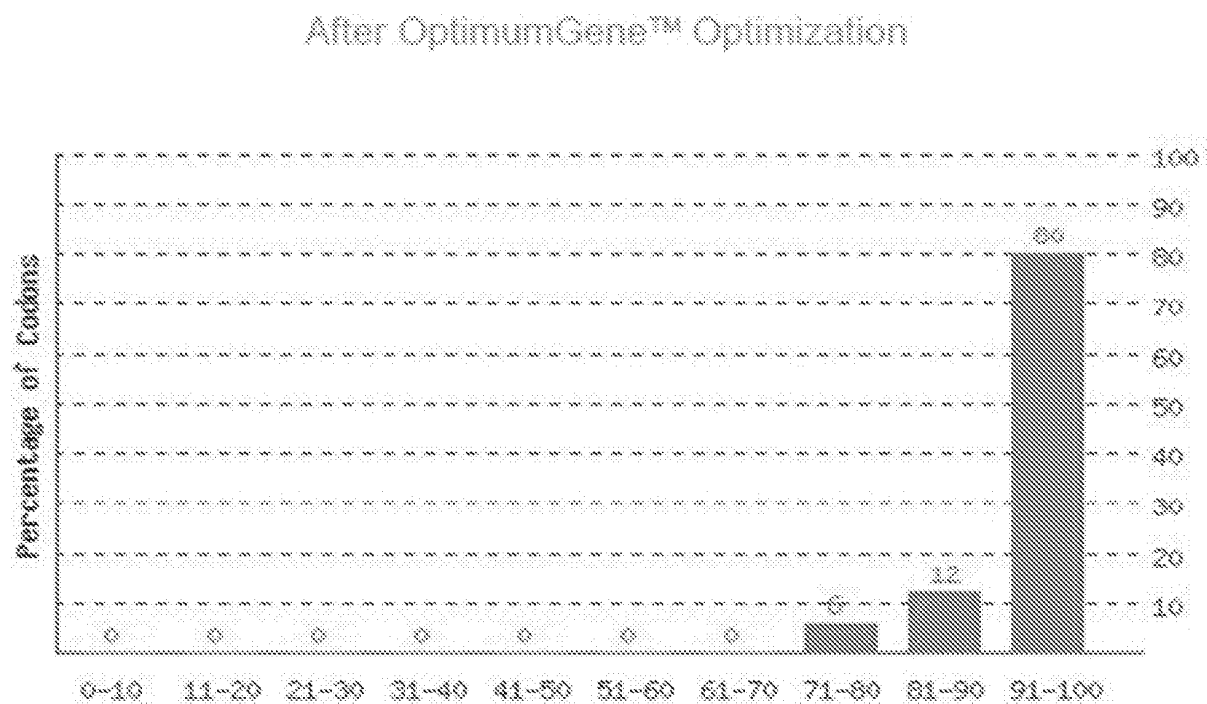


FIG. 72B

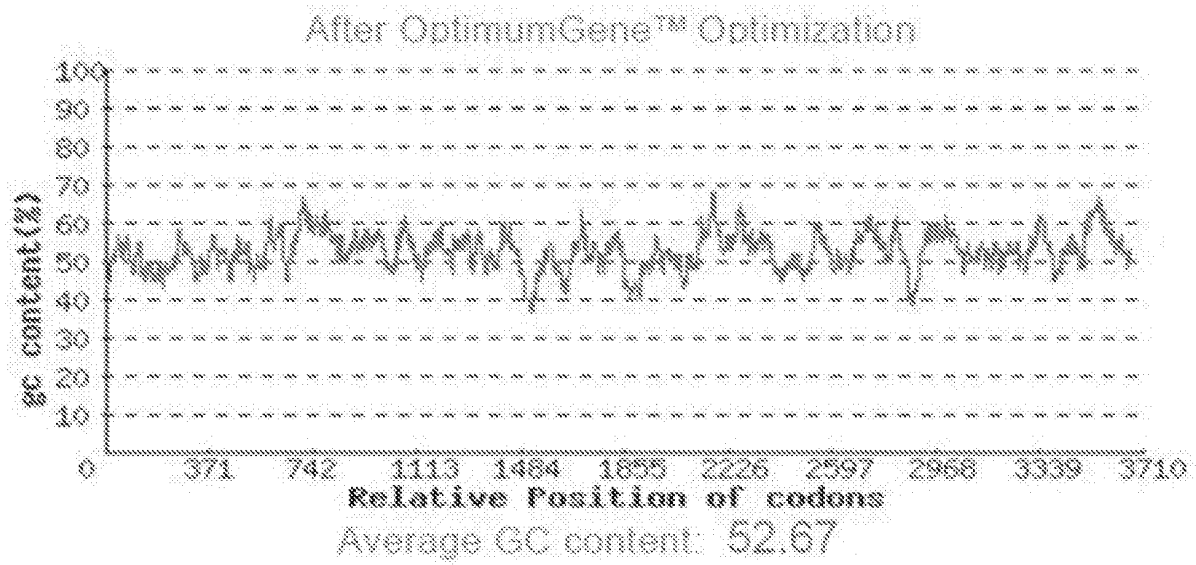


FIG. 72C

Restriction Enzymes Optimized

** Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.*

BamHI(GGATCC)	1(3769)
EcoRI(GAATTC)	1(3859)
HindIII(AAGCTT)	0
BsmBI(CGTCTC)	0
Bsal(GGTCTC)	0
BbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(3833)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BsrGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0

CIS-Acting Elements Optimized

Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0

Antiviral Motifs Optimized

0

FIG. 72D

After Optimization

Max Direct Repeat: Size:13 Distance:1506 Frequency:2
 Max Inverted Repeat: None
 Max Dyad Repeat: None

FIG. 72E

GenBank

ATGAACAATTACGACGAGTTTCAACCAAGCTGTATCTCTATCCAGAAAAACCATCCGTTTTGAGCTGAAGCCACAGGTC
 AGAACCATGGAGCACTCTGGAGACATTCAACTTCTTTGAGGAGGACCGGGATAGGCCCCGAGAGATATAGATCTCT
 AAGGAGGCCATCGACGAGTACCACAGGAGTTTATCTATGAGCACCTGACCAATATGTCTCTGGATTGGAACTCT
 CTGAAGCAGATCAGCCAGAGTACTATAAGAGCAGGGAGGAGGAGGACAGAGAGGTGTTCTCTCTCCAGCAGAGAG
 AGGATCCGCCAGGAGATCTGTCTGAGTTTAAAGAGGACGATCTCTTCAAGGACCTGTTTTCCAAGAGCTGTTTC
 TCTGAGCTGCTGAAGGAGGAGATCTACAAGAGGGCCACACCACAGGAGATCGACGCCCTGAGAGAGCTTGGATAAG
 TTTTCCGGCTATTTTCATCGGCCCTGACAGAGAAATAGGAGCAACATGTACTCCGACCGGCCGATGAGATCACCGCCATC
 TCCATCCGATCTGTGAATGAGAACTTCCCCAAGTTTCTGGATAACCTGCGAGAGTACCAGGAGGCCACGAGAGAGG
 TATCTCTGAGTGGATCATCAAGGCCGAGAGGCCCTGTTGGGCCACAATATCAAGATGGACGAGGTGTTCTCTCTG
 GAGTACTTTAATAAGGTGCTGAACCAAGGAGGGCATCCAGCGGTACACCTGGGCCCTGGGCCGCTATGTGACCAAG
 AGGCCCCAGAGAGATGATGGGCTGAACTGATGCCCCGAACTTGGGCCACCCAGTCCGAGAGAGGCTCCAAGGGCAGA
 ATCCACATGACCCCCCTGTTCAAGCAGATCTCTGTCGAGAGAGGAGTCTTCTCTTACATCCCCGACGTGTTTACA
 GAGGATTCTCAGCTGCTGCTAGCATGGGCCCTTCTTTTCCGAGATCGAGAGATGACAAGGATGGCAACATCTTC
 GACTGGGCCCTGGAGCTGATCTCTAGCTACGCCGAGTATGATAACCGAGCGGATCTATATCAGACAGGCCGACATC
 AATAGAGTGTCCAACGTGATCTTTGGAGAGTGGGGCACCCTGGAGAGGCCCTGATGAGGGAGTACAGAGGCCGACTCT
 ATCAATGATATCAACCTGGAGCGCAGATGCAAGAGGTTGGACAGGTGGCTGGATTCTAGGAGTCTTGGCCCTGAGC
 GATGTGCTGAGAGCCATCAAGAGGAGCCGCAACAATGACGCCCTTCAACGAGTATATCTCCAGAGTCCGACACGCC
 AGAGAGAGATCGATGCCGCCGCCAGGAGATGAAGTTCATCAGCGAGAGAGATCTCCGGCGATGAGAGATCTATC
 CACATCATCAGGACCTCTGCTGGACAGCTGACGAGTCTCTGCACTTCTTTAATCTGTTTAAAGGCAAGGCAGGAC
 ATCCCACTGGATGGAGGCTTCTACGCCGAGTTTGAACGAGTGCACAGCAGCTGTTTGGCATCTGCCCCCTGTAT
 AACCAAGGTGCGGAACTATCTGACCAAGAACAACTCTGAACACAAAGAGAGTCAAGCTGAAATTTCAAGAACCCTACA
 CTGGCCAATGGCTGGGACCAAGAACAGGTGTACGATTTAGCCCTCCTGTATCTTTCTGGGGACCGCAATTACTAT
 CTGGGCATCATCAATCCTAAGAGAAAGAGAGCATCAAGTTCTAGCAGGGCTCTGGCAACCGGCCCTTCTACCGG
 AAGATGGTGTATAAGCAGATCCCGGCCCTAATAAGGAACTTGGCCAAGAGTGTCTCTGACCTCCACAAAGGGCAGG
 AAGGAGTATAAGCCCTCTAAGGAGATCATCGAGGGCTACGAGGCCGACAGCACATCAGGGGCCGATAAGTTCCAGC
 CTGGATTTTTGTACAGAGCTGATCGATTTCTTTAAGGAGTCCATCGAGAGGCACAGGAGCTGGTCTAAGTTCAAC
 TTCTACTTCAGCCCCAACCGAGAGCTATGCCGACATCTCTGAGTTCTACCTGGATGTGGAGAGCCAGGGCTATCCG
 ATGCACCTTTGAGATATCAGCGCCGAGACAACTGACGAGTATGTGGAGAGGGCGATCTGTTTCTGTTCCAGATC
 TACAACAGGATTTTGTCAAGGCCGCCACCGGCCAGAGAGGACATGACACATCTACTGGAAATGCCGCCCTTCAGC
 CCGGAGAACCTGCGAGACCTGGTGGTGAAGCTGAAACCGGAGGCCGAGCTGTTTATAGGAGCAAGTCCGATATC
 AAGGAGATCGTGACACCGCGAGGGCGAGATCTGGTGAATAGGACCTACAAACCGGCCACACCAAGTGGCCGACAG
 ATCCACAGAGAGCTGACCGATTATCACAATGGCCGACCAAGGACCTGGCCGAGGCCAGGAGTACCTGGATAAG
 GTGAGATACTTCAAGGCCCACTATGACATCACCAAGGATCGGAGATACTGAAACGACAAGATCTATTTCCAGGTG
 CTTCTGACCTTGAACCTTCAAGGCCAAGCGGCAAGAGAGATCTGAACAAGATGGTCATCGAGAGGTTCTGTCTGGAT
 GAGAAAGGCCACATCATCGGCATCGACAGGGGCCGAGGCCAATCTGCTGTACTATTCCATCATCGACAGGTCTGGC
 AAGATCATCGATCGAGAGAGCTGAATGTGATCGACGGCTTTGATTATCGGGAGAGCTGAACCAAGAGAGAGATC
 GAGATGAAGGATGCCGCCAGTCTTGGAAACCCATCGGCAAGATCAAGGACCTGAAGGAGGGCTACCTGAGCAAG
 GCCGTGACAGAGATCAACAGATGCCCCATCCAGTATATGCCCATCTGTGTCATGGAGGAGCTGAACCTACGGCTTC

FIG 72F

MNNDYDEPTKLYPIQKTIRFELKPKQGRTEHLETFNFFFEEDRDRAEKYKILKEAIDEYHKK
 FIDEHLTNMSLDWNSLKQISEKYYKSREEKDKKVFLSEQKEMRQEIVSEFKKDDRFKDLF
 SKKLFSELLKBEIYKKGNHQEIDALKSFDKPSGYFIGLHENRKNMYSYGDEITAISNRIV
 NENFPKFLDNLQKYQEARKKYPEWIIKAESALVAHNIKMDDEVFSLEYFNKVLNQEGIQRY
 NLALGGYVTKSGEKMMGLNDALNLAHQSEKSSKGRIHMTPLFKQILSEKESFSYIPDVFT
 EDSQLLPSIGGFPAQIENDKDGNI FDRAL ELISSYAEYDTERIYIRQADINRVSNVIFGE
 WGTLGGLMREYKADSINDINLERTCKKVDKWLDSEKFA LSDVLEAIKRTGNND AFNEYIS
 KMRTAREKIDAARKEMKFISEKISGDEESIHI IKTLLDSVQQFLHFFNLFKARQDIPLDG
 AFYAEFDEVHSKLFAI VPLYNKVRNYLT KNNLNTKKIKLNFKNPTLANGWDQNKVYDYAS
 LIPLRDGNYYLGI INPKRKKNIKFEQGGSGNGPFYRKVMYKQIPGPKNKLPFRVFLTSTK GK
 KEYKPSKEIIEGYEADKHIRGDKFDLDFCHKLIDFFKESIEKHKDWSKFNPFYFSPTESYG
 DISEFYLDVEKQGYRMHFENISAETIDEYVEKGD LFLFQIYNKDFVKAATGKKDMHTIYW
 NAAFSPENLQDVVVKLNGEAELFYRDKSDIKEIVHREGEILVNRTYNGRTPVPDKIHKKL
 TDYHNGRTKDLGEAKEYLDKVRYFKAHYDITKDRRYLNDKIYFHVPLTLNFKANGKKNLN
 KMVIEKFLSDEKAHIIGIDRGERNLLYYSIIDRSGKIIDQQSLNVIDGFDYREKLNQREI
 EMKDARQSWNAIGKIKDLKEGYLSKAVHEITKMAIQYNAIVVMEELNYGFKRGRPKVEKQ
 IYQKFENMLIDKMNYLVFKDAPDESPGGVLNAYQLTNPLESFAKL GKQTGILFYVPAAYT
 SKIDFTTG FVNLFNTSSKTNAQERKEFLQKPE SISISYSAKDGGIFAFAFDYRKFGTSKTDH
 KNVWTAYTNGERMRYIKEKKRNELFDPSKEIKEALTSSGIKYDGGQNILFDILRSNNNGL
 IYTMYS SFIAAIQMRVYDGKEDYIISPIKNSKGEFFRTDPKRREL PIDADANGAYNIALR
 GELTMRAIAEKFDPDSEKMAKLELKHKDWFEFMQTRGD

FIG. 72H

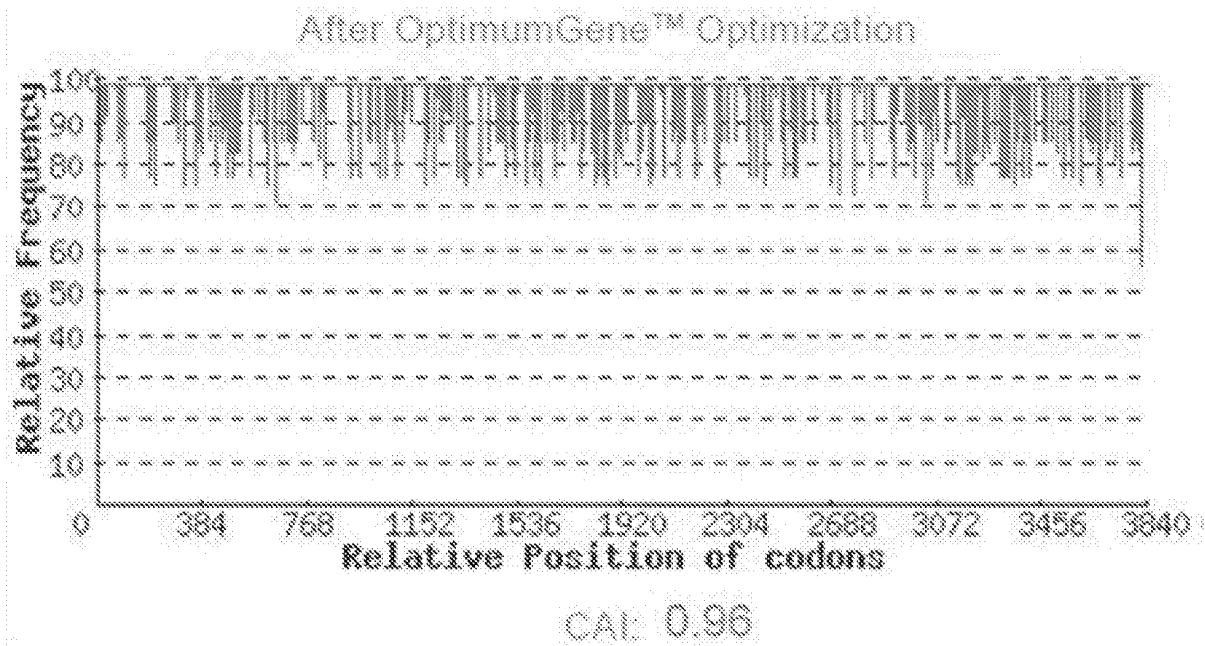


FIG. 73A

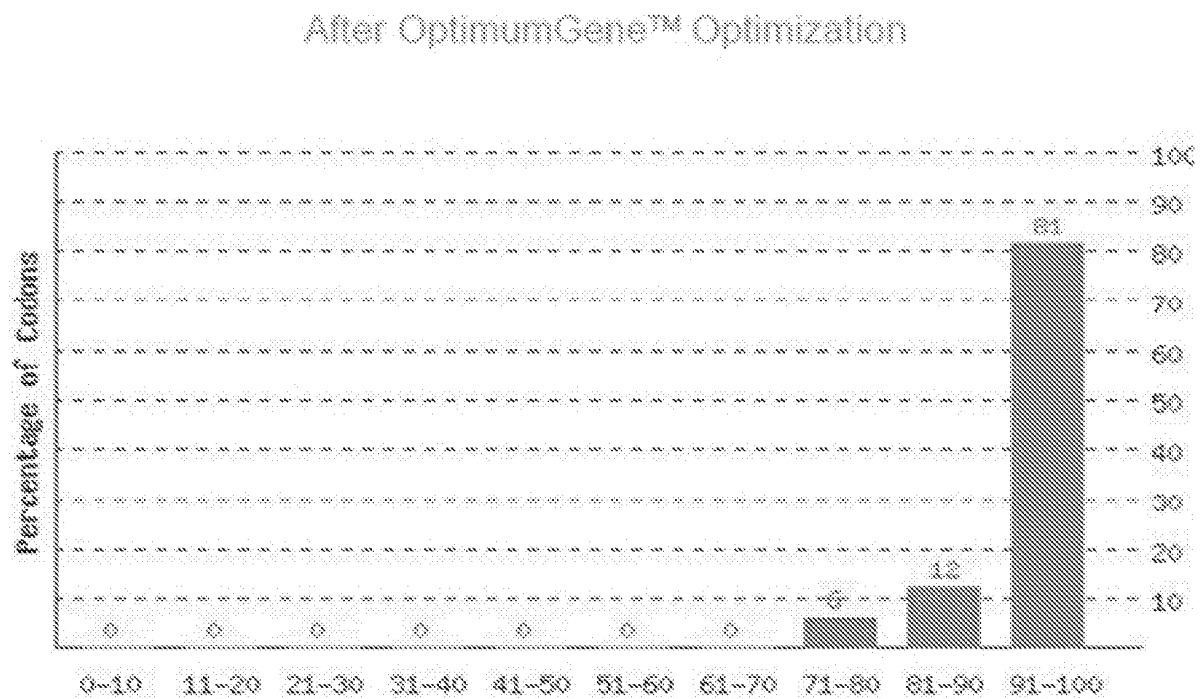


FIG. 73B

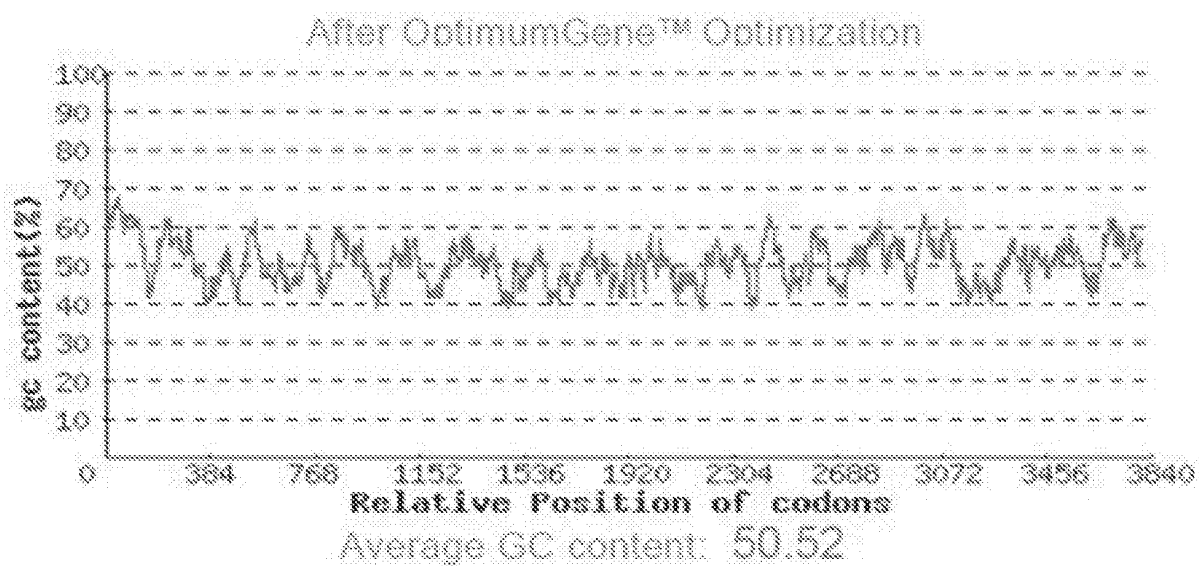


FIG. 73C

Restriction Enzymes	Optimized
<i>* Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.</i>	
BamHI(GGATCC)	1(3901)
EcoRI(GAATTC)	1(3991)
HindIII(AAGCTT)	0
BamBI(CGTCTC)	0
BsaI(GGTCTC)	0
BbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(3965)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BsrGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0

CIS-Acting Elements	Optimized
Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0

Antiviral Motifs	Optimized
	0

FIG. 73D

After Optimization

Max Direct Repeat: Size:14 Distance:894 Frequency:2
 Max Inverted Repeat: None
 Max Dyad Repeat: Size: 13 Tm: 40.1 Start Positions: 2322, 3831

FIG. 73E

GenBank

ATGAACGGCAATAGGTCCATCGTGTACCGGAGTTCTGTGGCGTGTATCCCGTGGCCAGACCCCTGAGGAATGAG
 ETGGGCCCCGTGTGGGACACACACAGGAGCACATCATCCAGAACGGCCTGATCCAGGAGGACGAGCTGCGGCGAGG
 AAGAGCACCCGAGCTGAGAGACATCATGTGACGATTACTATACAGAGTACATCCATAAATCTCTGAGCGGCGTACCC
 GACCTGGACTTTCACCTTGTCTTTCGAGCTGATGAACTGTGTGAGAGCTCCCCCTCCAAAGCAATAAGAAAGGCC
 CTGGAGAAAGGAGCAGTCTAAGATGAGGAGGAGGATCTGTCACCCAGCTGCACTCCGACTCTAACTACAAGATATC
 TTTAACGCCAAGCTGCTGAAGGAGATCCTGCGCTGATTTCTCAAGAACTACAACTCAGTATGAGCTGAAGGATAAG
 GCGGCGAAGCTGGAGACACTGGGCCCCCTTTAATGGCTTCAGGCACATACTTTACCGACTTCTTTGAGAGAGGAGG
 AACGTGTTTCACTAAGGAGGCGGTGAGCACATCCATCGCCTACCGCATCGTGCACGAGAACTCCCTGATCTTCTCTG
 GCCAATATGACCTCTTATAAGAAGATCAGCGAGAGAGGCCCTGGATGAGATCGAAGTGTTCGAGAGAGAACAAATCAG
 GACAAGATGGGCGAATTGGGAGCTGATCAGATCTTTAACCCCTGACTTCTACAATATGCTGCTGATCCAGTCCGCGC
 ATCGACTTCTACAGCAGATCTGCGGCGTGTGTAATGCCACATGAACCTGTACTGTGAGCAGACCAAGAACAAAT
 TATAACCTGTTCAGATGCGGAAGCTGCACAAGCAGATCCTGGCTTACACCAGCACCAGCTTTCAGGTTGCGCAAG
 ATGTTTCAGGACGATATGAGCGTGTATAACGCCGCTGACGCTTTCATCGACGAGACAGAGAAAGGCGAACATCATC
 GGCAGGCTGAAGGATATCGTGAATTAAGTACGACGAGCTGGATGAGAGAGAGATCTATATCAGCAAGGACTTTTAC
 GAGACACTGAGCTGCTTCATGTCCGCGAAGCTGGAACTGTGATCACAGGCTGCTGTGAGAACTTCTACGATGAGAAC
 ATCCACGCCAAGGCGAAGTCCAGGAGGAGAGAGGTTGAAGAGGCGGCTGAAGGAGGACAGTACAAGTCTATCAAT
 GAGGTGAAAGATCTGCTGGAGAGATATATCGATGAGAGAGAGAGGAAATGAGTTCAAGAACAGCAATGCCAAGCAG
 TACATCCGCGAGATCTCCACATCATCACCGACACAGAGACAGCCGCTGAGGTATGAGGATCACATCTCTCTG
 ATCGAGAGCGAGGAGGAGGCGCGAGAGATGAGAGAGCGGCTGGATATGTATATGAAACATGTACTCTGGGCGAAG
 GCTTTATCGTGGAGGAGGCTGCGACAGAGATGAGATGTTCTACAGCGATATGAGCGATATCTATAATATCTCTG
 GAGAACATCGTCCGACTGTATAATCGGGTGAGAAACTACGTGACCCGAGAGCCCTACAACTCTAAGAAATCAAG
 CTGAATTTCCAGAGCCCTACACTGCGCAATGGCTGCTCCGCTCTAAGGAGTTGACACAAATGCCATCATCTCTG
 ATCAGAGATAACAAGTACTATCTGCGCATCTTCAATGCCAAGAACAGCCAGACAAAGAGATCATCCAGGCGAAC
 TCCGATAAGAGAGAACGACAAAGATTAAGAAGAGTGGTGTATAACCTGCTGCCAGGCGCGAACAGATGCTGCGCC
 AAGGTGTTTCTGTCTAAGAGAGGCGATCGAGACATTCAGAGCCCTCGAGCTATATCATCTCTGCGCTACAGCGCGAC
 AAGCACATCAAGACAGCGAGAAATTTGATATCTCTTCTGTGCGGAGCTGATCGATTACTTCAAGAACAGCATC
 GAGAGGCAAGCGGAGTGGAGAAAGTATGAGTTCAAGTTTTCGCGCACCGACAGCTACTCCGATATCTCTGAGTTG
 TATCGGAGGCTGAGATGCGAGGCTACAGAACTGACTGAGATATATCAGCGAGGCGGACATCAACAAGCTGGAT
 GAGGAGGCGAAGATCTATCTGTTTCAGATCTACAATAGGATTTGCGCGAGACAGGACCGCGAAGGAGAAATCTG
 CACACAATGTACTTTAAGAACATCTTCTCCGAGGAGAACTGGAAGGACATCATCATCAAGCTGAACGGCCAGGCC
 GAGCTGTTTTATCGGAGAGCCTCTGTGAAGAATCCCGTGAAGCACAGAGAGGATAGCGTGTGCTGTTGAACAAGACC
 TACAAGAATCAGCTGACAAACCGCGAAGTGTGTGAGAAATCCCATCCCTGACGATATCTATAACGAGATCTACAG
 ATGTATAANTGCTACATCAAGAGCTCCGACCTGTCTGAGGCGCGCGAGAGTACCTGAGATAAGGTGGAGGTGAGG
 ACCGCGGAGAGAGACATCTGAAGGATTACCGCTATACAGTGGACAGGTACTTCACTCACACACCTATCACCATC
 AACTATAAGGTGACCGCGCGCAACAAATGTGAATGATATGCTGCTGAAGTACATCGCGGAGAACAGAGATATCCAC
 GTGATCGCATCGACCGGCGGAGAGGAACTGATCTACATCTCCGTGATCGATTCTCAAGCGCAACATCTGTGAAG
 CAGAAATCTTACACATCTTGAACAACTACGACTACAGAGAGAGCTGCTGAGAGAGGAGAAACCGCGGAGTAC
 GCCAGAGAGAACTGGAAGAGCATCGCAATATCAAGGAGCTGAAGGAGGCGCTATATCTCCGCGCTGCTGACAGAG

FIG. 73F

FIG. 73G

MNGNRSIVYREFVGVIPVAKTLENELRPVGHTEQHIIQNGLIQEDELQEKSTELKNIMD
 DYYREYIDKSLSGVTDLDFTLLFELMNLVQSSPSKDNKKALEKEQSKMREQICTHLQSDS
 NYKNIFNAKLLKEILPDFIKNYNQYDVKDKAGKLETALFNGFSTYFTDFPEKRKNVPTK
 EAVSTSIAYRIVHENSILIFLANMTSYKKISEKALDEIEVIEKNNQDEMGDWELNQIFNPD
 FYNMVLIQSGIDFYNEICGVVNAHMNLVYCQQTKNYNLFPKMRKLHKQILAYTSTSFEVPK
 MPEDDMSVYNAVNAPIDETEKGNIIGKCLKDIVNKYDELDEKRIYISKDFYETLSCFMSGN
 WNLITGCVENFYDENIHAKGKSKEEKVKKAVKEDKYKSINDVNDLVEKYIDEKERNEFKN
 SNAKQYIREISNIITDTETAHLEYDDHISLIESEEKADMKRLDMYMNMYHWAKAFIVD
 EVLDRDEMFSYSDIDDIYNILENIVPLYNEVRNYVTQKPYNSKKIKLNFQSPTLANGWSQS
 KEFDNNAIILIRDNKYYLAIFNAKNKPKDKKIIGNSDKKNDNDYKKMVYNLLPGANKMLP
 KVFLSKKGIETFKPSDYIISGYNAHKHIKTSENFDISFCRDLIDYFKNSIEKHAEWRYE
 FKFSATDSYSDISEFYREVEMQGYRIDWTYISEADINKLDEEGKIYLFQIYNKDFAENST
 GKENLHTMYFKNIFSEENLKDIIIKLNCQAELFYRRASVKNPVKHKKDSVLVNKTYKNQL
 DNGDVVRIPIPDDIYNEIYKMYNGYIKESDLSEAAKEYLDKVEVRTAQKDIVKDYRYTVD
 KYFIHTPITINYKVTARNNVNDMVVKYIAQNDDIHVIGIDRGERNLIYISVIDSHGNIVK
 QKSYNILLNNYDYKKKLVEKEKTREYAPKNWKSIGNIKELKEGYISGVVHEIAMLIVEYNA
 I IAMEDLNYGFKRGRFKVERQVYQKFESMLINKLNYFASKEKSVDEPGGLLKGYQLTYVP
 DNIKNLQKQCGVIFYVPAAFSTKIDPSTGPFISAFNFKSISTNASRKQPFMQFDEIRYCAE
 KDMFSFGFDYNNFDTYNITMGKTQWTVYTNGERLQSEFNNARRTGKTKSINLTETIKLLL
 EDNEINYADGHDIRIDMEKMDDEKKSEFFAQLLSLYKLTVMRNSYTEAEQENGISYDK
 IISPVINDEGEFFDSNYKESDDKECKMPKADANGAYCIAKGLYEVLKIKSEWTEGDF
 DRNCLKLPHAEWLDFIQNKRYE

FIG. 73H

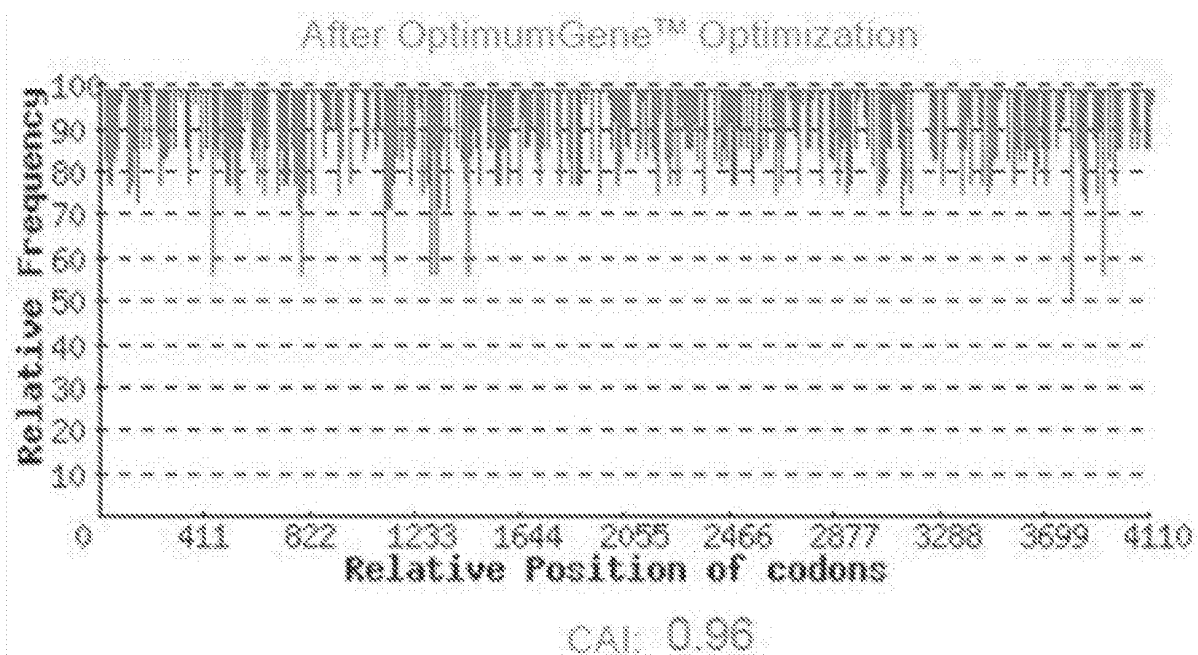


FIG. 74A

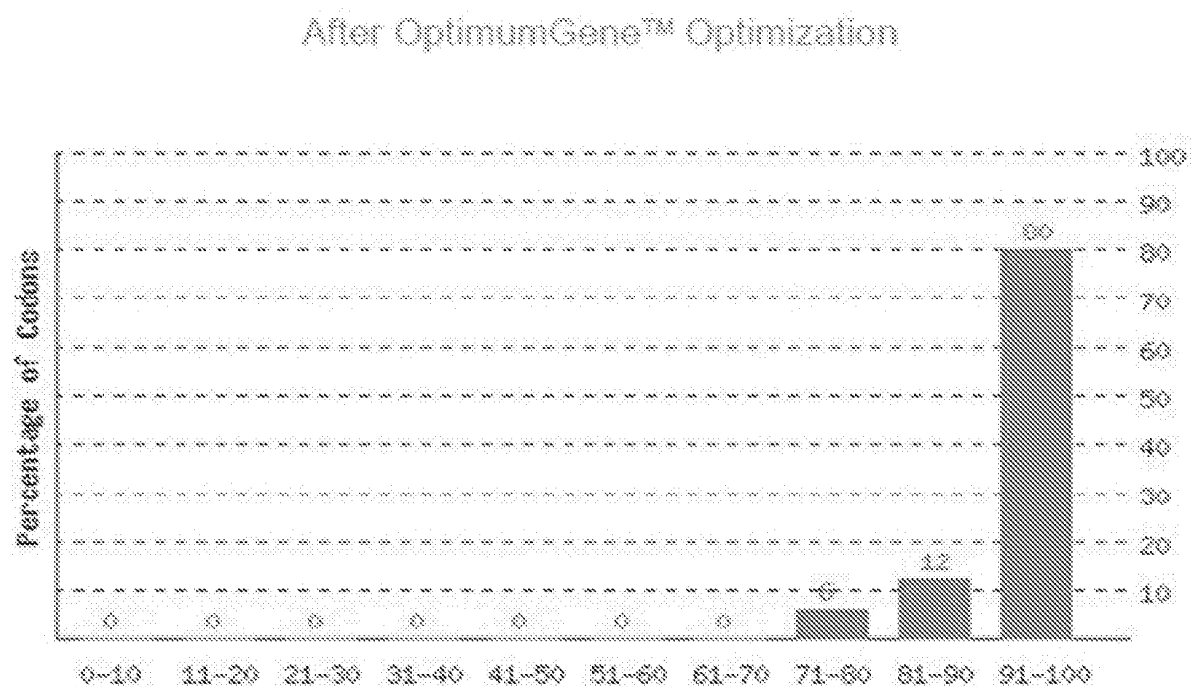


FIG. 74B

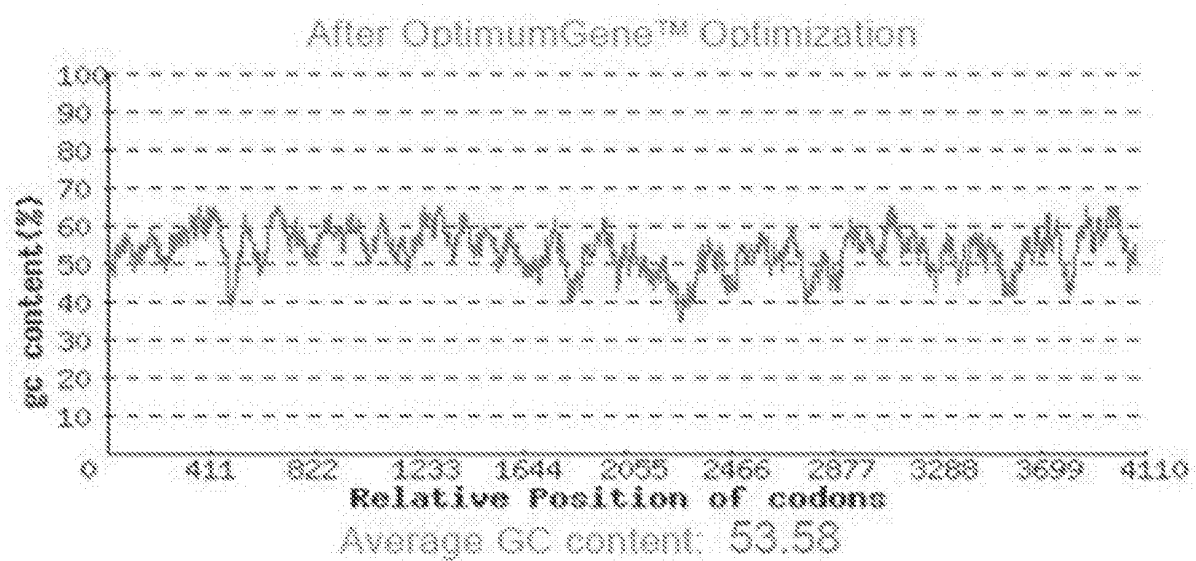


FIG. 74C

Restriction Enzymes		Optimized
* Green, filtered sites; Blue, checked sites (not filtered); Red, kept sites.		
BamHI(GGATCC)		1(4174)
EcoRI(GAATTC)		1(4264)
HindIII(AAGCTT)		0
BsmBI(CGTCTC)		0
BsaI(GGTCTC)		0
BbsI(GAAGAC)		0
AgeI(ACCGGT)		0
XhoI(CTCGAG)		0
NdeI(CATATG)		1(4238)
NotI(GCGGCCGC)		0
KpnI(GGTACC)		1(1)
BsrGI(TGTACA)		0
SpeI(ACTAGT)		0
XbaI(TCTAGA)		0
NheI(GCTAGC)		0
ARE		0

CIS-Acting Elements		Optimized
Splice(GGTAAG)		0
Splice(GGTGAT)		0
PolyA(AATAAA)		0
PolyA(ATTAAA)		0
Destabilizing(ATTTA)		0
PolyT(TTTTTT)		0
PolyA(AAAAAAA)		0

Antiviral Motifs	Optimized
	0

FIG. 74D

After Optimization

Max Direct Repeat: Size:14 Distance:525 Frequency:2
 Max Inverted Repeat: None
 Max Dyad Repeat: None

FIG. 74E

ATTACG

ATBCTGTTCCN93ACTTTTACCCACCTGTATCCACTGTCCNAGACAGTGAONTTTSAGCTGAAGCCCCATCGATAGG
 ACCCTGGAGCACATCCACGCCAAGAACCTTCTGTCTCAGGACGAGACAATGCGCGATATGACCCAGAAAGGTGAAA
 GTGATCCTGAGCGATTACCACCGCGACTTCATGCGCGATATGATGGGCGAGGTGAAGCTGACCAAGCTGGCCGAG
 TTCTATGACCTGTACCTGAACTTTCCGAGAGAACCCAAAGACGATGAGCTCCAGAGCCAGCTGAAAGATCTGCAG
 GCGGTGCTGAGAAAGGAGATCGTGAAGCCCATCGGCGATGCGCGCAAGTATAAGGCCGGCTACGACAGGCTGTTT
 GCGGCCAAGCTGTTTAAAGGACCGGCAAGGAGCTGGGCGATCTGGCCAACTTCGTGATCGACAGGAGGAGAGAGC
 TCCCCAAAGCTGKCCCACTGGGCCACTTCGAGAACTTTTCCACTATTTTCACAGGCTTTTCACGATAACCCGGAAG
 AATATGTATTTCTGACGAGGATAAGCACACCGCCATCGCCCTACCGCTGATCCACGAGAACCTGCCCCGCTTTATC
 GACAATCTGCAGATCTGACCAATCAAGCAGAAACCTCTGCGCTGTACGATCAGATCATCAAGAGCTGAGC
 GCGAGCGGCTTGACCTGTCTCTGGCCAGGCCACTGATGCGCTATCACAGCTGCTGACACAGGAGGCGATCACC
 GCCTACAAATACACTGCTGGAGGAGATCTCCGAGAGGCGAGGCTCTCTTAAGATCCAGGCGATCAAGAGCTGATC
 AATTCTCACCAAAACCAAGCACTGCCACAAGAGCGAGAGAAATGCGCAAGCTGAGGCCACTGCAAGAGCAGATCTCTG
 TCCGAGCGCATGAGCGTGTCTCTCTGCGCTCTAAGTTTTCGCGAGGATAGCGAGATGTGCCAGGCCGTGAACGAG
 TTCTATCGCCACTACGCCGACGTGTTCCGCAAGGTCGAGAGCCTGTTCCAGCGCTTTGACGATCACCGAAGGAT
 GGCATCTACCTGGAGCACAGGAACCTGAATGAGCTGTCCAAAGCAGGCCUTTCGGCGACTTTGCACTGCTGGAGCGC
 GTGCTGGACCGATACTATGTGGATGTGTGAATCCAGAGTTCAACGAGCGGTTTTCGCAAGGCCAAGAGCCGACAT
 GCCAAGGCCAAGCTGACAAAGGAGAGGATAAGTTTCATCAAGGCGTGCACTCCCTGGGCTCTCTGAGCGAGGCC
 ATCGAGCACTATACCGCAAGGCACGACGATGAGAGCTTCAGGCGAGGCCAAGCTGGGACACTACTTCAGCGACCGC
 CTGGCCGGAGTGGACAACCCCATCCAGAGGATCCACAACAATCACAGCACCATCAAGGCCCTTCTGAGAGGGAG
 CGCCCTGCGAGAGAGAGAGCCCTGCCAAGCATCAAGTCGCGCAAGATCTGAGATGACACAGCTGAGCGAGCTG
 AAGGAGCTGCTGGATAACGCCCTGAATGTGGCCCACTTCGCCAAGCTGCTGACCAAGAGCCACACTGGACAT
 CAGGATGGCACTTCTATGGCGAGTTTGGCGTGTCTGTACGAGAGCTGGCCAAGATCCCCACCTGTATTAACAAG
 GTGAGAGATTACCTGAGCCAGAAGCCTTTCTCCACCGAGAGTACAAGCTGAACCTTTGCAATCCAACACTGCTG
 AATGGCTGGGACCTGAACAAGGAGAGGATAATTTCCGCGTGATCTTCGAGAGGACCGGCTGCTACTATCTGGCC
 CTGCTGGACAAGGCCCCACAAGAAGGTGTTTGATAAGGCCCTTAATACAGGCCAAGAGCATCTATCAGAAGATGATC
 TATAAGTACCTGAGGTGAGGAAGCAGTTCCCAAGGTGTTCTTTTCCAGGAGGCCATCGCCCTCAACTACCAC
 CCTTCTAAGGAGCTGGTGGAGATCAAGGACAAGGCCCGCAGAGATCCGACGATGAGCGCTGAAGCTGTATCGG
 TTTATCCTGGAGTGTCTGAAGATCCACCTAAGTACGATAAGAACTTCGAGGGCGCCATCGGCGACATCCAGCTG
 TTTAAGAAGGATAAGAAGGGCAGAGAGGTGCCAATCAGCGAGAGGAGCCTGTTTCGATAAGATCAAGGCGATCTTT
 TCTAGCAAGCCTAAGCTGGAGATGGAGACTTCTTTATCGGCGAGTTCAAGAGGTATAACCCAGGCCAGGACCTG
 GTGGATCAGTATAATATCTACAGAGAGATCGACTCCAAAGGATAATCCCAAGAGGAGAGATTTCTACAACATCAC
 GCGAAGTTTAAGGAAGATCTGGTCCGCTACTATTACGAGTCTATGTGCAAGCACAGGAGTGGGAGGAGAGCTTC
 GAGTTTTCCAAAGAGCTGCAAGACATCGGCTGTACGTGATGTGAACGAGCTGTTTACGAGATCGAGACACCG
 AGACTGAATTATAAGATCTCTTCTGCAACATCAATGCCGACTACATCGATGAGCTGCTGGAGGAGGCGCGAGCTG
 TATCTGTTCCAGATCTACAACAGGAGCTTTTCCCAAGGCGCCACGCGCAAGGCCAATCTGCAACCCCTGTACTTC
 AAGGCCCTGTTTCTGAGGACAACTGGCCGATCCTATCTATAAGCTGAATGGCGAGGCCAGATCTTCTACAGA
 AAGGCTCTCTGACATGAAAGAGACAACATCCACAGGCCCGGCGAGGCTGCTGGAGAGCAAGAATCCCGATAT
 CCTAAGAGAGAGAGTTCTGTACGACATCATCAAGGATAAGAGGTACACACAGGACAGGTTCTGCTGACGCTG

FIG. 74F

CCAATCAACATGAACTTTTGGCTTCAGGGCATGACAAATCAAGGATTTCAATTAAGAAAGTTGAACCAATCTATCCAG
 CAGTATGACCGAGCTGAACTGTGATCGGCATCGATCGGGSCAGAGACACCTGCTGTACCTGACCGTGAATCAATAGC
 AAGGGCCAGATCCTGAGCAGTGTFTTCCCTGAAACGACATCAACACAGCCTCTGACCAATGGCCACACAGATGACCACA
 CCTTACCCACAAGATCCTGGAATAGGAGGGGAGATCGAGCGCCTGAAACGCCCGCTGGGATGGGAGAGATCGAGACA
 ATCAAGGAGCTGMACTCTGGCTATCTGAGCCACTTGGTGCACCCAGATCAGCCAGCTGATGCTGAAGTACAAAGCC
 ATCGTGGTGTCTGAGGACCTGAATTTCCGCTTTTAAGAGGGGGCCCTTTAAAGGTGAGAGAGCAGATCTATCAGAAC
 TTCCAGAAATGCCCTGATCAAGAAAGCTGAACCCACTTGGTGTCTGAAGGACAAAGGCCAGATGAGATCGGCTCTTTAC
 AAGAATGCCCTGCGASCTGAACCAACANTTTCAAGATCTGAAGGAGCATCGGCAAGCAGACCGGCTTCTGTGTTTAT
 GTGCCCCGCTTGGAAACACCTCTAAGATCGACCCCTGAGACAGGCTTTGTGTGATCTGCTGAAGCCAGATACGAGAAC
 ATCGGCTCAGAGCCAGGCTTCTTTGGCAAGTTCCGACAGATCTGCTATAATGKCGACAAAGGATTTACTTGAAGTTT
 CACATGGACTACGCCAAGTTTACTCGATAAGGCCAAGGAATAGCCGCCAGATCTGGACAAATCTTTTCCACGGCGAC
 AAGCGGTACGTTTACGATAAGACAGCCAAACCAAGAAATAGGAGCGGCTCGCAAGGKCATCAACTTGAATGATGAGCTG
 AAGTCCCTGTGTGGCCCGCCACCAATCAACGAGAAAGCAGCCCAACCTGGTCATGGACATCTGGCAGAACAAATGAT
 AAGGAGTTTCAAGATCTCTGATGTACTTGGCTGAAAGCCCTGCTGGCCCTGCGGTACAGCAACGGCTCTCTTGAC
 GAGGATTTCACTCTGTCTCCCGCTGGCAAGCAGAGGCGCTGTCTTTAATAGCCGCCCTGGCCGACGATACACAG
 CCTCAGAAATGCCATGECCAACGGCCCTTACCACATGCGCCCTGAAGGSCCTTTGGCTGCTGAATGAGCTGAAGAAC
 TCCCGACGATCTGAAACAGCTGAAGCTGGGCATCGACAAATCAAGACCTGGCTGAATTTTCCGCCAGAACAGG
 AAAAGGCTGGCTGGTCACTGAAAGGCTCGCTGGCTGAAAGGAAAGGAGATCTTACTGATGGCTTTTCTAGAT
 TAGCTTATCTGCTAGACCTGGCTGATTAAGCATAGCCTATGAAATGCTGGGACTATGGCTTGAAGATG

FIG. 74G

MLFQDPHTLYPLSKTVRFELKPIDPTLEHIHAKNPLSQDETMA DMHQVKVILDDYHRDF
 IADMMGEVKLT KLAEFYDVYLKFRKNPKDDELQKQLKDLQAVLRKEIVKFIGNGGKYKAG
 YDRLFGAKLFPKGKELGDLAKFVIAQEGESSPKLAHLAHPEKFPSTYPTGFHDNRKNMYS
 EDKHTAIAYRLIHENLPRFIDNLQILTTIKQKHSALYDQIINELTASGLDVSLASHLDGY
 HKLLTQEGITAYNTLLGGISGEAGSPKIQGINELINSHHNQHCHKSERIAKLRPLHKQIL
 SDGMSVSFLPSKFADDSEMCQAVNEFYRHYADVFAKVQSLFDGFDDHQKDGIIYVEHKNLN
 ELSKQAFGDFALLGRVLDGYVVDVNVPEFNERFAKAKTDNAKAKLTKEKDKFIKGVHSLA
 SLEQAIENHYTARHDDSVQAGKLGQYFKHGLAGVDNPIQKIHNHSTIKGFLE RERPAGE
 RALPKIKSGKNPEMTQLRQLKELLDNALNVAHFAKLLTTKTTLDNQDGNFYGEFGVLYDE
 LAKIPTLYNKVRDYL SQKPFSTEKYKLNFGNPTLLNGWDLNKEKDNFGVILQKDGCCYLLA
 LLDKAHKKVFDNAPNTGKSIYQKMIYKYLEVRKQFPKVPFSKEAIAINYHPSKELVEIKD
 KGRQRSDDERLKLRYRFILECLKIHPKYDKKFEGAIGDIQLFKKDKKGREVPIS EKDLFDK
 INGIFSSKPKLEMEDPFIGEFKRYNPSQDLVDQYNIYKKIDSNDNRKKENFYNNHPKFKK
 DLVRYYYESMCKHEEWEESEFEFSKKLQDIGCYVDVNELFTEIETRLN YKISFCNINADY
 IDELVEQQQLYLFQIYNKDFSPKAHGKPNLHTLYFKALFSEDNLADPIYKLNGEAQIFYR
 KASLDMNETTIHEAGEVLENKNPDNPKKRQFVYDI IKDKRYTQDKFMLHVPITMNFVGVOG
 MTIKEFNKKVNQSIQQYDEVNVIGIDRGERHLLYLTVIN SKGEILEQCSLNDITTASANG
 TQMTTPYHKILDKREIERLNARVGWGBIETIKELKSGYLSHV VHQISQLMLKYNAIVVLE
 DLNFGFKRGRFKVEKQIYQNFENALIKKLNHLVLKDKADDEIGSYKNALQLTNNFTDLKS
 IGKQTGFLFYVPAWNTSKIDPETGFVDLLKPRYENIAQSQAFFGKFDKICYNADKDYFEF
 HIDYAKFTDKAKNSRQIWTICSHGDKRYVYDKTANQNKGAAKGINVNDELKSLFARHHIN
 EKQPNLVMDICQNNDEKFHKS LMYLLKTLALRYSNASSDEDFILSPVANDEGVFFNSAL
 ADDTQFQNA DANGAYHIALKGLWLLNELKNSDDLNVKVLAI DNQTWLNFAQNE

FIG. 74H

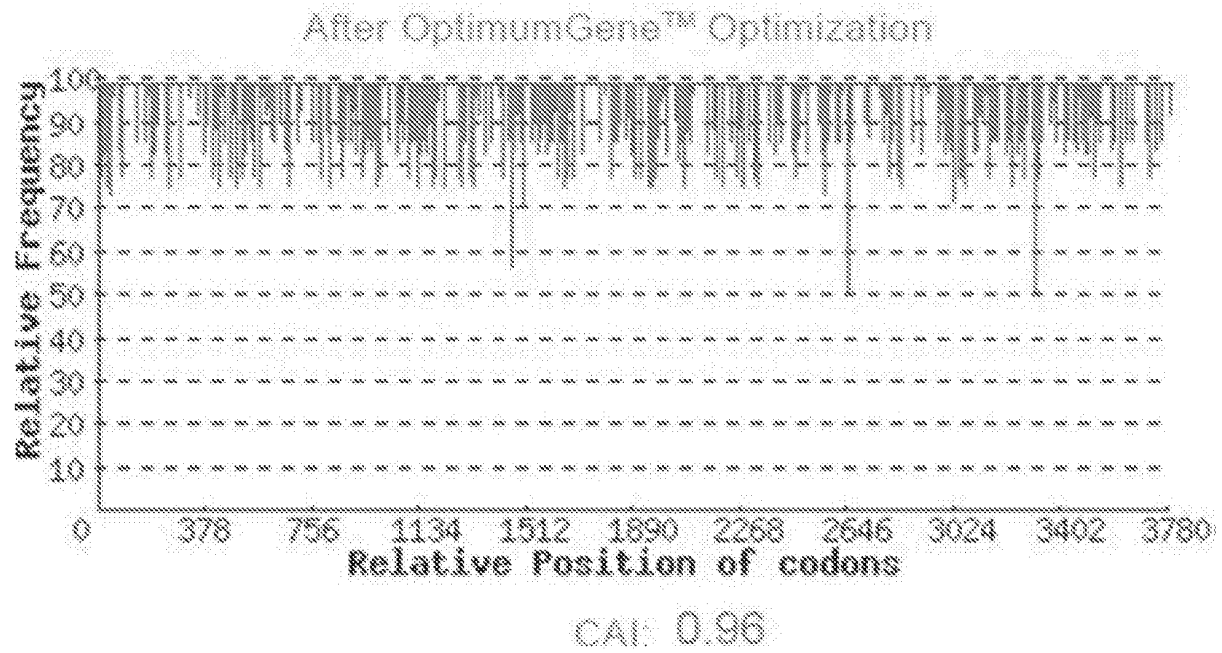


FIG. 75A

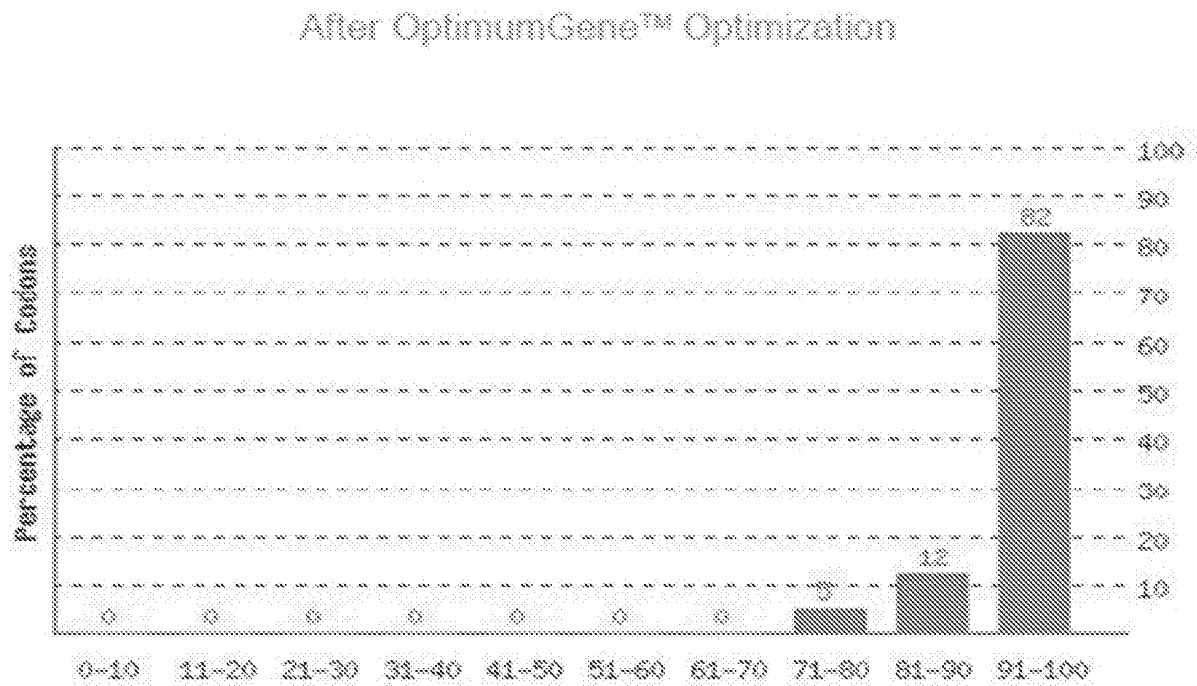


FIG. 75B

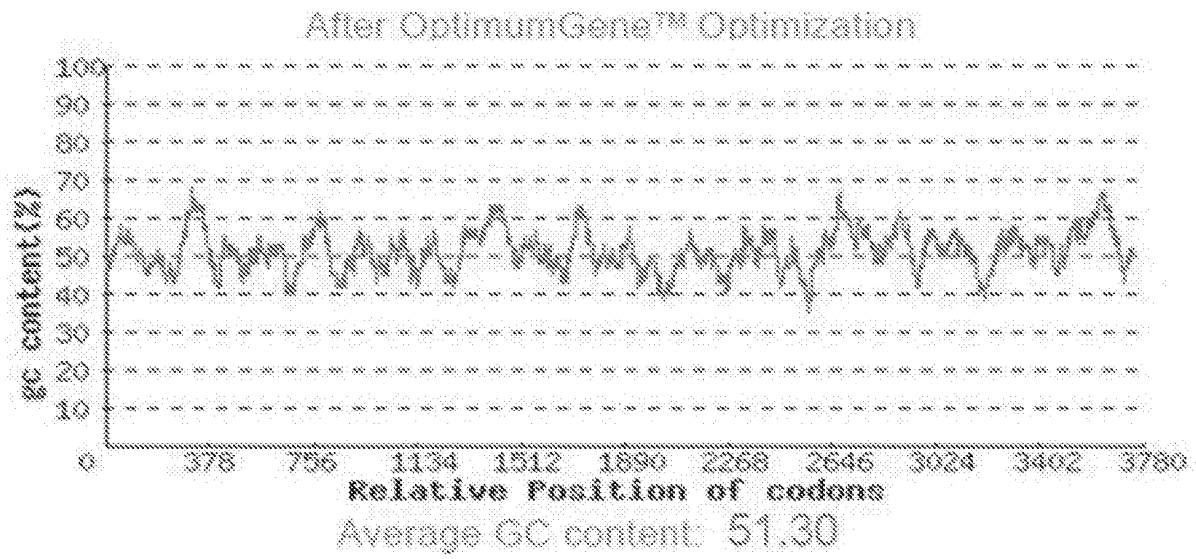


FIG. 75C

Restriction Enzymes	Optimized
* Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.	
BamHI(GGATCC)	1(3844)
EcoRI(GAATTC)	1(3934)
HindIII(AAGCTT)	0
BamBI(CGTCTC)	0
Bsal(GGTCTC)	0
BbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(3908)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BsrGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0

CIS-Acting Elements	Optimized
Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0

Antiviral Motifs	Optimized
	0

FIG. 75D

After Optimization

Max Direct Repeat: Size:14 Distance:111 Frequency:2
 Max Inverted Repeat: None
 Max Dyad Repeat: Size: 13 Tm: 35.4 Start Positions: 1027, 221

FIG. 75E

FIG. 75F

ATGGAGGACTATTTCGGGCTTTGGTGAACATCTACTCTATTCAGAAAACCCCTGAGGTTTGGAGCTGAAGCCACTGAGG
 AAGACACTGGAGCACATCGAGAAGAAGGGCTTCTTGAAGAAGGACAGATCCGGGCCGAGGATTACAAGGCCCTG
 AAGAGATCATCGATACTACCAAGAGCCCTATATCGAGAGAGGTTGTTTGATTCCGTGCTGACCAGAGAAGAAG
 AAGGACAAGACCCGCTTTTCTACACAGTTTCATCAAGGAGATCAAGGAGTTTCAGCGAGCTGTACTATAAGACCGAG
 AAGAACATCCCGACAAAGGAGAGGCTGGAGGGCCCTGAGCCAGAAGCTGCCCAAGATGCTGGTGGGCGCCTTTAAG
 GGCGAGTTCTCCGAGGAGGTGGCCGAGAAGTATAAGAACCTGTTTTCTAAGGAGCTGATCAGGAATGAGATCGAG
 AAGTTCTGCGAGACAGACGAGGAGGCGCAAGCAGGTGTTCTAACTTCAAGAGCTTCAACACATACTTTAACCCTTCTC
 CACTCCAAACAGGCAGAAATATCTATTCGACGAGAGAGTCTACAGCCATCGGCTACCGCATCATCCACAGAAC
 CTGCTTAAGTTCTCTGATAATCTGAAGATCATCGAGTCCATCCAGCGCCGGTTCAAGGACTTCCCATGCTCTGAT
 CTGAAGAGAAGACCTGAAGAGATCGATAAGAAATATCAAGCTGACCGAGTACTTCAGCATGACCGCTTCTGTAAC
 GTGCTGAATCAGAAGGGCATCGATGCTTCAACACAAATCTCTGGGCGGCAAGTCCGAGGAGTCTGGGAGAGAGATC
 CAGGGCTGAACGAGTACATCAATCTGTATCGGAGAGAAACATATCGACAGAAAGAACCTGCCCCAATGTGAGG
 ATCTGTTTTAAGCAGATCTCTGGCGATAGGAGACAAAGAGCTTTATCCCTGAGGCGCTTCCAGAGCATCAGTCC
 GTGCTGAACCTCTATCACAGAGTTGCGCAAGTACCTGAAGCTGAGTAAGAAGAGAGAGAGCATCATCGCCGAGCTG
 AAGAAGCTTTCTGAGCTCTTCAATCGCTACGAGCTGACAGCGCATCTATCTGGCCAACTATAATAGCCCTGGCTCT
 ATCAGCACCTTCTCTGTTTGAGCATTTGCTCTTTATCAAGAGTCCGTGCTTTCAAGTATGACGAGTCCCTGCGC
 GAGCCCCAAGAGAGAGATCAAGTCTCCCTGAGAGTACGAGAGGAGAGAGAGAGAGTGGCTGAAGCAGAACTACTAT
 ACAATCTCTTTCTGAGAGATGCCATCGAGAGCTATTCCAAGTCTCAGGACGAGAGAGAGAGTGAAGATCCGCTG
 GAGGCTTACTTTGCGGAGTTCAAGAGCAGGACCATGCCAAGAGCCAGTTCCAGCTGCTGAGAGAGATCGAGGAG
 GCCTATGCCATCTGAGAGCTCTGCTGGGAGCAGAGTACCCCAAGGAGCCCAACCTGAAGGCCGATAAGAAAGAA
 GTGGGCAAGATCAAGGACTTCTGAGATGAGATCAAGTCCCTGCACTTCTTTCTGAAGCTCTGCTGCTGCTGCTGCTG
 ATCTTTGAGCAGAAAGATCTGGCTTCTACAATCAGCTCGAGGGCTACTATGAGGAGATCGATTCTATCGGCCAC
 CTGTATAACAAGGTGCGGATTTATCTGACCGGCAAGATCTACAGCAAGGAGAGATTTAAGCTGAGCTTCCAGAAC
 AGCACCTGCTGAAGGGCTGGGAGCAGAGAACCGGAGAGTGGCCAATCTGTGCTGATCTTCAGAGAGGACCAAGAG
 TACTATCTGGGCTGTATGAGTAAGGAGAACATATACCATCTGTGCGACATCCCCAAGGTGAAGCCTAAGAGCTG
 TTTTACGAGAGAGTGGTGTATAAGCTGATCCCCACACCTCACATGCACTGCCCCGGATCATCTTCTCTAGCGAC
 AACCTGTCTATCTATATCTAGCAAGTCCATCTGAGAGATCAGAGAGGCGCAAGAGCTTTAAGGAGGCAAGAAC
 TTCAAGCTGAAGGACTGTCAAGTTTATCGATTTCTACAAGGAGTCTATCAGCAAGAAATCAGGACTGGAGCGGA
 TTGAGCTTCAAGTTCAAGCAGACGAGCAGCTACGAGAACATCGAGGAGTTTTTACCGGAGGTTGAGAGAGCAGGGC
 TATAACCTGAGCTTCAAGAGAGTGTCTAAGTTCTACATCGACAGGCTGCTGAGGAGTGGCAAGCTGTACCTGTTT
 CAGATCTATACAGGAGCTTTTCTATCTTCAGCAAGGCGCAAGCCCAATCTGACACCATCTATTTTGGTCCCTG
 TTCTCTAAGGAGAACCTGAAGGAGCTGTGCTGAGAGCTGAATGGCGAGGCCAGATGTTCTTTGGAAGAGATCC
 ATCAACTACGATGAGAGAGAGAGCGGAGGGGCCCAACCCCGAGCTGTTTGAGAGAGCTGAAGTATCTATCTCTG
 AAGGACAAAGAGTACAGCGAGGATAGGTTTCAGTTCCACCTGCCATCAGCCTGAACCTTCAAGTCCAGGAGCGG
 CTGAGCTTTAATCTGAAGTGAATGAGTTCTGAGAGAGAACAAAGGACATCAATATCATCGGCATCGATCGGGC
 GAGAGAAACCTGCTGTACCTGGTCATGATCAATCAGAAGGCGGAGATCTGGAAGCAGACCTGCTGGACAGCATG
 CAGTCCGCGAGGGCCCGCTGAGATCAACTACAAGGAGAGGCTGCGAGAGAGAGATCGAGAGGAGTAAAGCC
 CGCAGAGCTGGGGCACAGTGAAGAAATATCAAGGAGCTGAAGGAGGCTATCTGTCTATCGTGTATCCACGAGATC

FIG. 75F

FIG. 75G

MEDYSGFVNIYSIQKTLRFELKPVGKTLHEIEKKGFLKKDKIRAEDYKAVKKIIDKYHRA
 YIEEVFDSVLHQKKKKDKTRFSTQFIKEIKEFSELYYKTEKNIPDKERLEALSEKLRKML
 VGAFKGEFSEEVAEKYKNLFSKELIRNEIEKFCETDEERKQVSNFKSFTTYFTGFHSNRQ
 NIYSDEKKSTAIGYRIIHQNLPKFLDNLKIIIESIQRRFKDFPWSDLKKNLKKIDKNIKLT
 EYPSIDGFVNVLNQKGIDAYNTILGGKSEESGEKIQGLNEYINLYRQKNNIDRKNLPNVK
 ILFKQILGDRETKSPFIPRAPDDQSVLNSITEFAKYLKLDKSKKSIIAELKKFLSSFNRY
 ELDGIYLANDNSLASISTFLFDDWSFIKKSVSFKYDESVGDPKKKIKSPLKYEKEKEKWL
 KQKYTTISFLNDAIESYSKSQDEKRVKIRLEAYFAEFKSKDDAKKQFDLLERIEEAYAIV
 EPLLGAEYPRDRNLKADKKEVGKIKDFLDSIKSLQFFLKPLLSEAEIFDEKDLGFYNQLEG
 YEEIDSIGHLYNKVRNYLTGKIYSKEKFKLNFENSTLLKGWDENREVANLCVIFREDQK
 YYLGVMKDKENNTILSDIPKVKPNELFYEKMVYKLIPTPHMQLPRIIFSSDNLSIYNPSKS
 ILKIREAKSFKEGKNFKLKDCHKFIDFYKESISKNEEDWSRFDKFKSKTSSYENISEFYRE
 VERQGYNLDFKKVSKFYIDSLVEDGKLYLFQIYNKDFSIFSKGKPNLHTIYFRSLFSKEN
 LKDVCLKLNGEAEMFFRKKSINYEKKKKREGHHPELFEKLYPILKDKRYSEDKFQPHLP
 ISLNPKSKERLNFNLKVNEFLKRKNDINIIGIDRGERNLLYLVMINQKGEILKQTLTLDSE
 QSGKGRPEINYKEKLQEKEIERDKARKSWGTVENIKELKEGYLSIVIHQISKLMVENNAI
 VVLEDNLNIGFKRGRQKVERQVYQKFEKMLIDKLNFLVFKENKPTTEPGGVLKAYQLTDEFQ
 SFEKLSKQTGFLFYVPSWNSTSKIDPRTGFIIDFLHPAYENIEKAKQWINKFDSIRFNSKMD
 WFEFTADTRKFSENMLGKNRVWVICTTNVERYFTSKTANSSIQYNSIQITEKLEKELFVD
 IPFSNGQDLKPEILRKNDVFFKSLLFYIKTTLSLRQNNNGKKGEEKDFILSPVVDKGR
 FFNSLEASDDEPKDADANGAYHIALKGLMNLLVLNETKEENLSRPKWKIKNKDWLEFVWE
 RNR

FIG. 75H

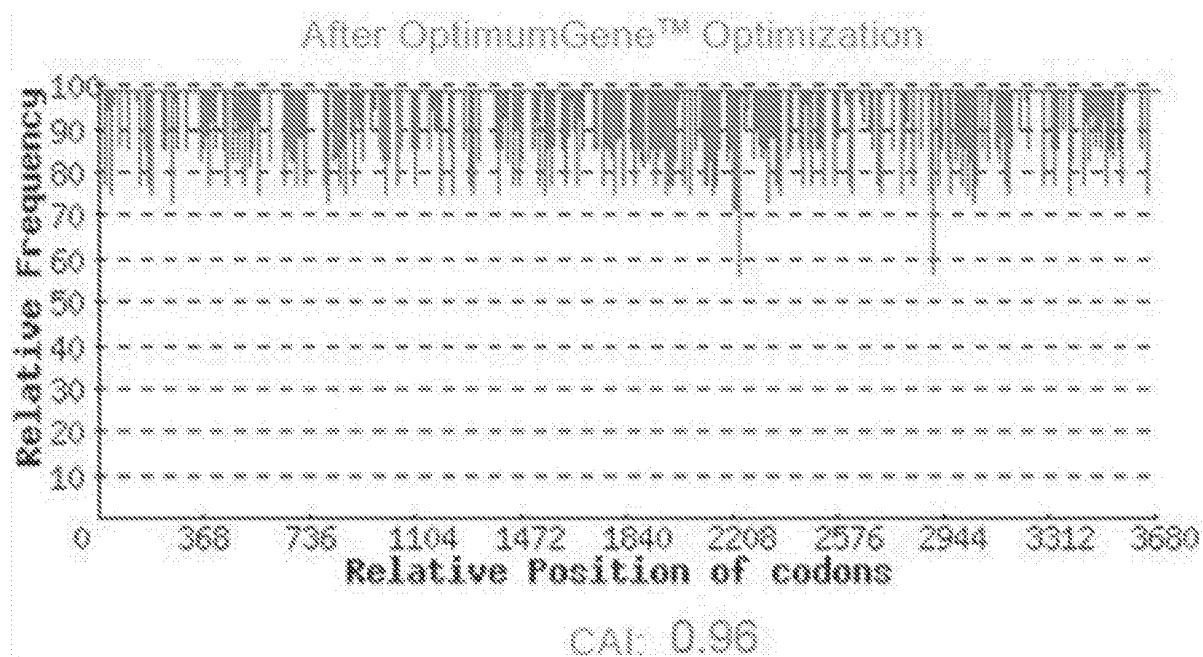


FIG. 76A

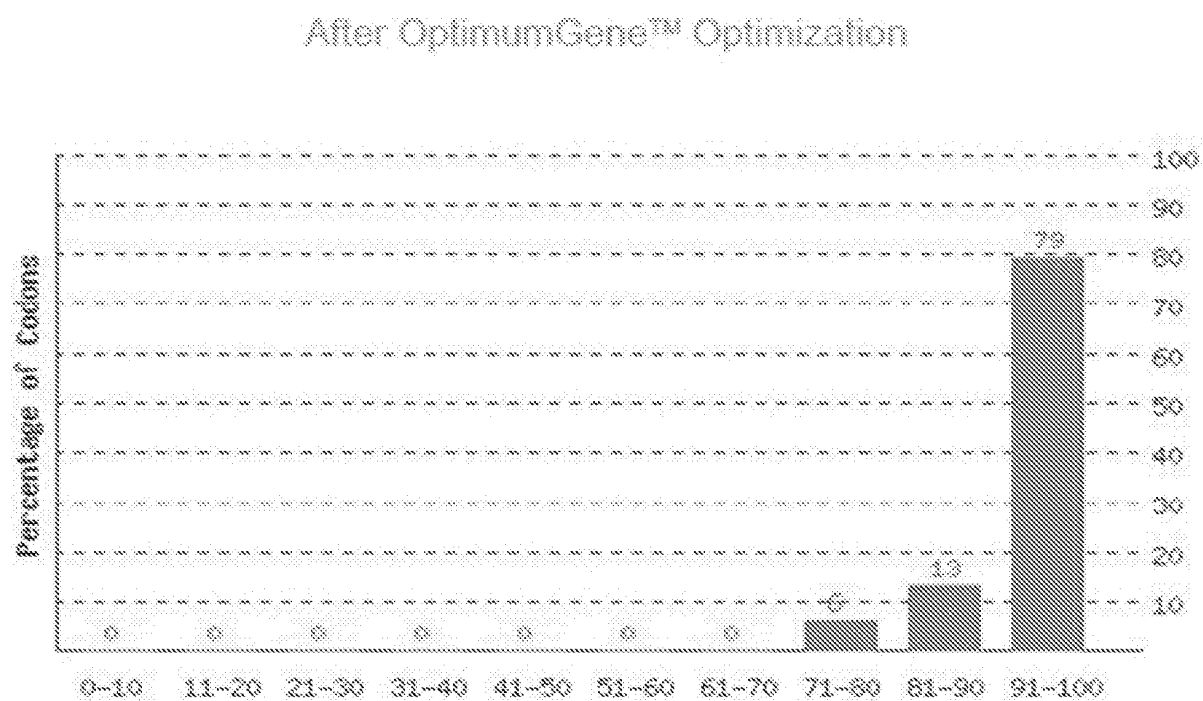


FIG. 76B

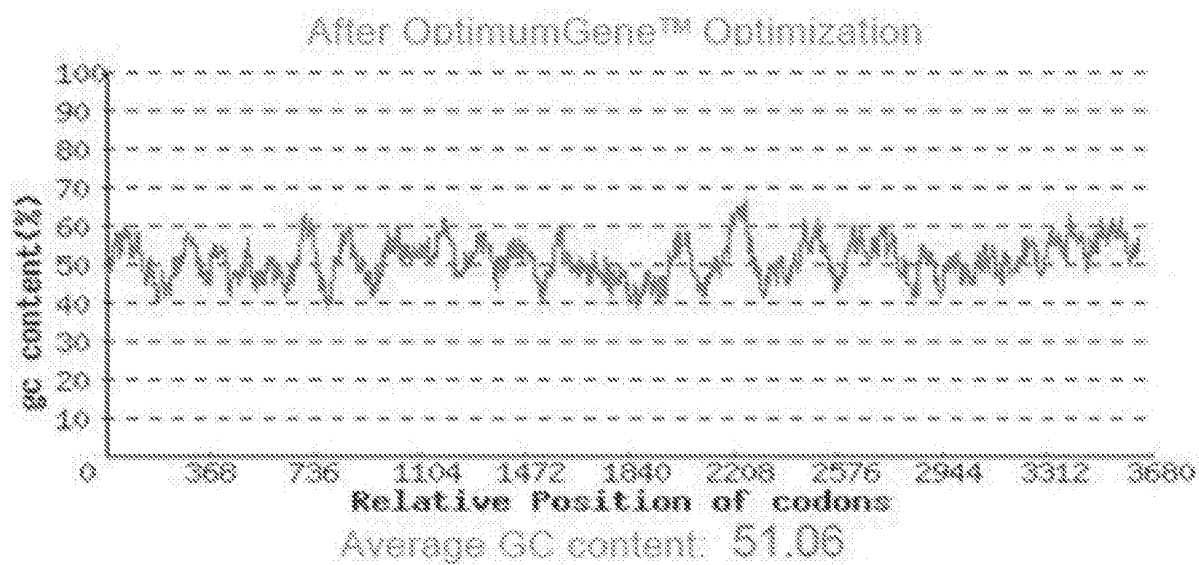


FIG. 76C

Restriction Enzymes	Optimized
<i>* Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.</i>	
BamHI(GGATCC)	1(3739)
EcoRI(GAATTC)	1(3829)
HindIII(AAGCTT)	0
BsmBI(CGTCTC)	0
BsaI(GGTCTC)	0
BbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(3803)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BsrGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0

CIS-Acting Elements	Optimized
Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0

Antiviral Motifs	Optimized
	0

FIG. 76D

After Optimization

Max Direct Repeat: Size:14 Distance:150 Frequency:2
 Max Inverted Repeat: None
 Max Dyad Repeat: None

FIG. 76E

FIG. 76F

ATGAGCAAGCTGGAGAAAGTTTACAAACTGCTACTCCCTGTCTAAGACCCCTGAGGTTTCAAGGCCATCCCTGTGGGC
 AAGACCCAGGAGAACATCGACAAATAAGCGGCTGCTGCTGGAGGACGAGAACGAGAGCCGAGGTTTATAAGGGCGTGG
 AAGAGGCTGCTGGATCGCTACTATCTGTCTTTTATCAACGAGCTGCTGCACAGCATCAAGCTGAGAGATCTGAAC
 AATTACATCAGGCTGTTCGGGAGAGAAAACCAAGACCGAGAGGAGAGTAAGGAGCTGGAGAACCTGGAGATCAAT
 CTGCGGAGGAGATCGCCAAAGGCCCTTCAAGGGCAACGAGGGCTACAAGTCCCTGTTTAAGAGGATATCATCGAG
 ACAATCCTGCCAGAGTTCTTGGACGATAAGGACGAGATCGCCCTGGTGAACAGCTTCAATGGCTTTACCAAGGCC
 TTCCCGGCTTCTTTGATAACAGAGAGAGATATGTTTTCCGAGGAGGCCAAGAGCAGATCCATCGGCTTCAGGTGT
 ATCAACGAGAGATCTGACCCGCTACATCTCTAATATGAGCATCTTCGAGAAAGTTGGAGGCCATCTTTGATAGGAC
 GAGGTCACGAGATCAAGGAGAGATCTGAAACAGGCACTATGATGTGGAGGATTTCTTTGAGGGCGAGTTCTTT
 AACTTTGTGCTGACACAGGAGGGCATCGACGTGTATAGCGCCATCATCGGCGGCTTCTGACGAGAGCGGGCGAG
 AAGATCAAGGGGCTGAACGAGTACATCAAAGCTGTATATTCAGAAAAACCAAGCGAAGGCTGGCTAAGTTTAAAGCCA
 CTGTATAGGAGGCTCTGAGCGATCGGAGCTCTCTGAGCTTCTACGGCGAGGGCTATACATCCGATGAGGAGGTTG
 CTGGAGGCTGTTTAGAAACACCCCTGAACAAAGAACAGCGAGATCTTCAGCTCCATCAAGAGGCTGGAGAGGCTGTTT
 AAGNATTTTGACGAGTACTCTAGCGCCGCGCATCTTTTGTGAAGAACCGGCCCGGCCATCAGCACAAATCTCCAAGGAT
 ATCTTGGGCGAGTGGAAAGCTGATCCGAGACAAAGTGGAAATCGCGAGTATGAGGATATCCACCTGAAGAGAGAGGCC
 GTGGTGACCGAGAGTACGAGGACGATCGAGAGAAAGTCTTCAAGAGAGATCGGCTCCTTTTCTCGAGGAGGCTG
 CAGGAGTACCCCGACCGCGATCTGTCTGTGTGTCGAGAAAGCTGAAGGAGATCATCATCCAGAGGTTGGATGAGATC
 TACAGGCTGTATGGCTCCTCTGAGAGGCTGTTTCGAGCGCGATTTTGTGCTGAGAGAGAGGCTGAAGAGAACGAC
 GCGGTGTGTCGCGCATCATGAAGGACCTGCTGGATTCTGTGAAGAGCTTCGAGGATTAATCAAGGGCTTCTTTGGC
 GAGGCAAGGAGACAAACAGGAGAGGCTCTTCTATGCGGATTTTGTGCTGGCGCTACGAGATCTGCTGAAGGTC
 GACCAATCTACGATGCCATCCGCAATTATGTGACCCAGAGAGCCCTACTCTAAGGATAAGTTCAAGCTGTATTTT
 CAGAACCCCTCAGTTTATGGCGGCTGGGACAAGGATAAGGAGACAGACTATCGGGCCACCTTCTCGAGATAGGCT
 TCCAGTACTATCTGGCCATCATGGATAAGAAAGTACCGCAAGTGGCTGCGAGAGATCGACAGGAGGATGTGAAC
 GCGAATTACGAGAGATCAACTATAAGCTGCTGCGCGGCTTAATAAGATGCTGCCAAGGTTGTTCTTTCTAAG
 AAGTGGATGGCTACTATAACCCAGCGAGGACATCCAGAGATCTACAAGNATGGCACATTCAGAGAGGGCGAT
 ATGTTTAAACCTGAATGACTGTCAAGGCTGATCGACTTCTTTAAGGATAGCATCTCCGGTATGCCAAGTGGTCC
 AATGCCATCAGATTTCAACTTTCTGAGACAGAGAAATTAAGGACATCGCGGCTTTTACAGAGAGGTTGAGGAG
 CAGGCTATAGGTTGAGCTTGGAGTGTGCCAGCAAGAGGAGGTGGATAAGCTGTTGAGGAGGGCAGGCTGTAT
 ATGTTCCAGATCTATAACAGGACCTTTCCGATAGTCTACGGCACACCCCAATCTGCACACCATGTACTTCAAG
 CTGCTGTTTGAAGAGAACATCACGACAGATCAGGTTGAGCGGAGGAGGAGGCTGTTCTATGAGGCGCGCTCC
 CTGAAGAGGAGGAGCTGGTGGTGCACCCAGGCCAATCTCCCTATCGCCAACAGAAATCCAGATAATCCCAAGAAA
 ACCACAACCCCTGTCTACGAGCTGTATAAGGATAAGAGGTTTCTGAGGAGGAGTACGAGCTGCCATCCCAATC
 GCGATCAATAAGTGCCTCAAGAACATCTTCAAGATCAATACAGAGGTGGCGGTGCTGCTGAAGCAGGAGATAAC
 CCGTATGTGATCGGCATCGATAGGGGCGAGCGCAATCTGCTGTATATCGTGGTGGTGGACGGCAAGGGCAACATC
 GTGGAGCGATTTCTTGAACGAGATCATCAACAACTTCAACGGCATCAGGATCAAGACAGATTACCACTCTCTG
 CTGACAAAGAGAGAGAGAGGTTGAGGGCCCGCCAGAACTGAGACCTCCATCGAGAGATATCAAGGAGCTGAAG
 GCGGCTATATCTCTCAGGTGGTGCACAGATCTGCGAGCTGGTGGAGAGTACGATGCCGTGATCGCCCTGGAG
 GACCTGAACCTCTGGCTTTAAGAAATAGCCCGCTGAAGGTTGAGAGAGCAGGTGTATCAGAAGTTGAGAGAGATGCTG

FIG. 76F

ATCSATAAGCTGAACTACATGGGTGACAAAGAGTCTAATTCCTTGTTGCAACAGGTCGGCTGCCCTGAAAGGGCTATCAG
 ATCACCAATAAGSTTCGAGAGGCTTTAAGTCCATGTTCTACCCAGAAAGGCTTTCATCTTTTACATGCCCTGGCTGGGCTG
 ACATCCAAAGATCGATCCATCTACCGGCTTTTGTGAACCTGCTGAAAAACCAAGTATACCAAGCATCGCCGATTCCAAG
 AAGTTTCATCAGGCTCTTTTGACAGGATCATGTTACGTTGCCCGAGGAGGATCTGTTGAGTTTGGCCCTGGACTATAAG
 AACTTCTCTCGCACAGACCGCGATTACATCAAGAACTGGAAAGCTGTACTCTTACCGCAACCGGATCAGAATCTTC
 CGCAATCTTAAGAAAGAACCAACCTGTTGCACTGGGAGGAGGTTGTGCTGACCCAGCGCTATAGAGGCTGTTCAAC
 AAGTACCGCATCAATTATCAACAGGGCGATATCAGAGCGCTGCTGTGCGAGCAGTCCGACAAAGGCTTCTACTCT
 AGCTTTATGGGCTGATGAGGCTGATGCTGCAATGCGGAACAGCATCAAGGCGCGACCGAGCGTGATTTTCTG
 ATCAGGCTCTGTGAAGAACTCCGACCGCATCTTCTACGATAGCCCGAACTATGAGGCGCGAGGAAATGCCATCCTG
 CCAAGGAACCGCGACCGCAATGGCGCTATACATCGCCAGAAAGGTGCTGTGGGCGCATCGGCTAGTTCAAGAAAG
 GCGGAGGACGAGAGAGCTGGATAAGGTGAAGATCGCCATCTCTAACAAAGGAGTGGGCTGGAGTACCGCCAGACGAGC
 GTGAAGCAC
 AAAAGGCTGGGCTCCAGAAAGAGCTGGGCGGCTAAAGAAAGAAAGGCAATCTACCGCTACATGTTTCCGAT
 TAGCTTATTCCTTACGAGCTGTGCTTGAATATGCTAACCGGCTGTGATCGGAGCTATGCTTAAAGATTC

FIG. 76G

MSKLEKFTNCYSLSKTLRPFKAI PVGKTQENIDNKRLLVEDEKPAEDYKGVKKLLDRYYLS
 FINDVLHSIKLKNLNNYISLFRKKTRTEKENKELENLEINLRKEIAKAFKGNEGYKSLFK
 KDI IETILPEFLDDKDEIALVNSFNQPTTAFTGFFDNRENMFSEEAKSTSIAFR CINENL
 TRYISNMDIFEKVDAIFDKHEVQEIKEKILNSDYDVEDFFEGEFFNFVLTQEGIDVYNAI
 IGGFVTESGEKIKGLNEYINLYNQKTKQKLPKFKPLYKQVLS DRESLSFYGEGYTSDEEV
 LEVFRNTLNKNSEIPSSIKKLEKLFKNFDEYSSAGIFVKNGPAISTISKDIFGEWNVIPD
 KWNAEYDDIHLKKKAVVTEKYEDRRRSFKKIGSFSLEQLQEYADADLSVVEKLKEII IQ
 KVDEIYKVYGSSEKLPDADFVLEKSLKKNDAVVAIMKDLLDSVKSPENYIKAFFGEGKET
 NRDES FYGDFVLAYDILLKVDHIYDAIRNYVTQKPYSKDKFKLYFQNPQFMGGWDKDKET
 DYRATILRYGSKYYLAIMDKKYAKCLQKIDKDDVNGNYEKINYLKLPGFNKM LPKVFFSK
 KWMAYYNPSEDIQKIYKNGTFKKGDMFNLDCHKLIDFFPKDSISRYPKWSNAYDFNFSET
 EKYKDIAGFYREVEEQGYKVSFESASKKEVDKLVBECKLYMFQIYNKDFSDKSHGTPNLH
 TMYFKLLFDENNHGQIRLSGGAELFMRRASLKKEELVVHPANSPIANKNPDNPKKTTTSL
 YDVYKDKRFSEDQYELHIPAIANKCPKNIFKINTEVRVLLKHDDNPFYVIGIDRGERNLLY
 IVVVDGKGNIVEQYSLNEIINNFMGIRIKTDYHSLLDKKEKERFEARQNWTSIENIKELK
 AGYISQVVHKICELVEKYDAVIALEDLNSGFKNSRVKVEKQVYQKFEKMLIDKLNVMVDK
 KSNPCATCGALKGYQITNKFESFKSMSTQNGFIFYIPAWLTSKIDPSTGFVNLLKTKYTS
 IADSKKFISSPDRIMYVPEEDLFEFALDYKNFSRTDADYIKKWKLYSYGNRIRIFRNPKK
 NNVFDWEEVCLTSAYKELFNKYGINYQQGDIRALLCEQSDKAFYSSFMALMSLMLQMRNS
 ITGRTDVDFLISPVKNSDGIFYDSRNYEAQENAILPKNADANGAYNIARKVLWAIGQFKK
 AEDEKLDKVKIAISNKEWLEYAQTSVKH

FIG. 76H

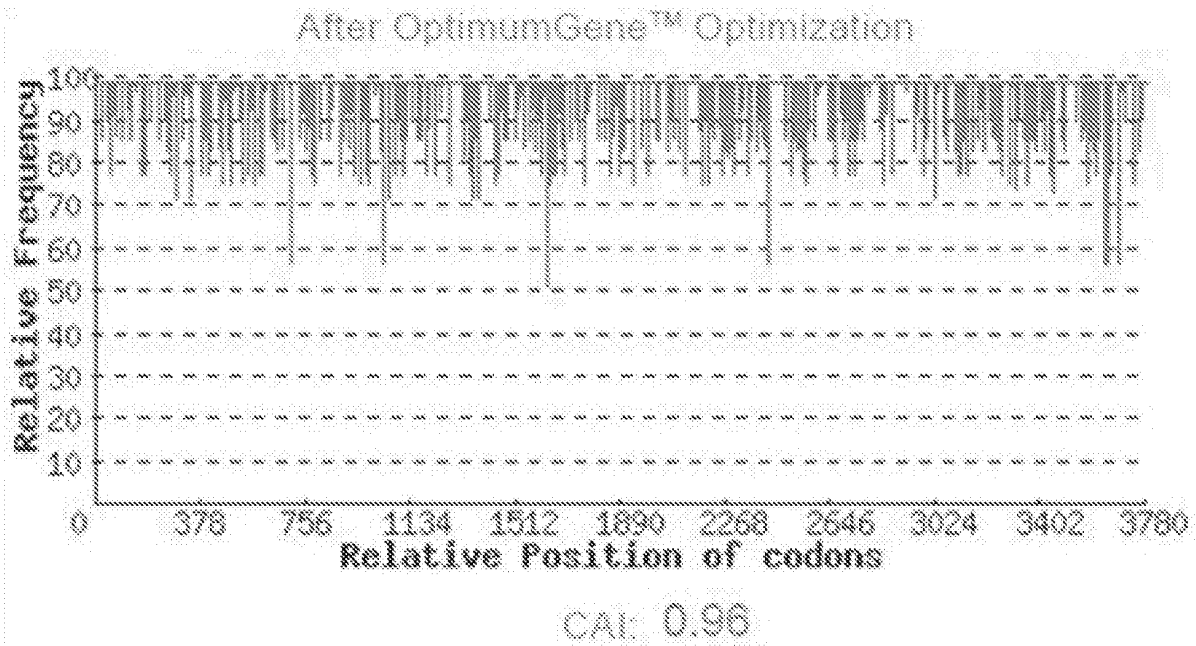


FIG. 77A

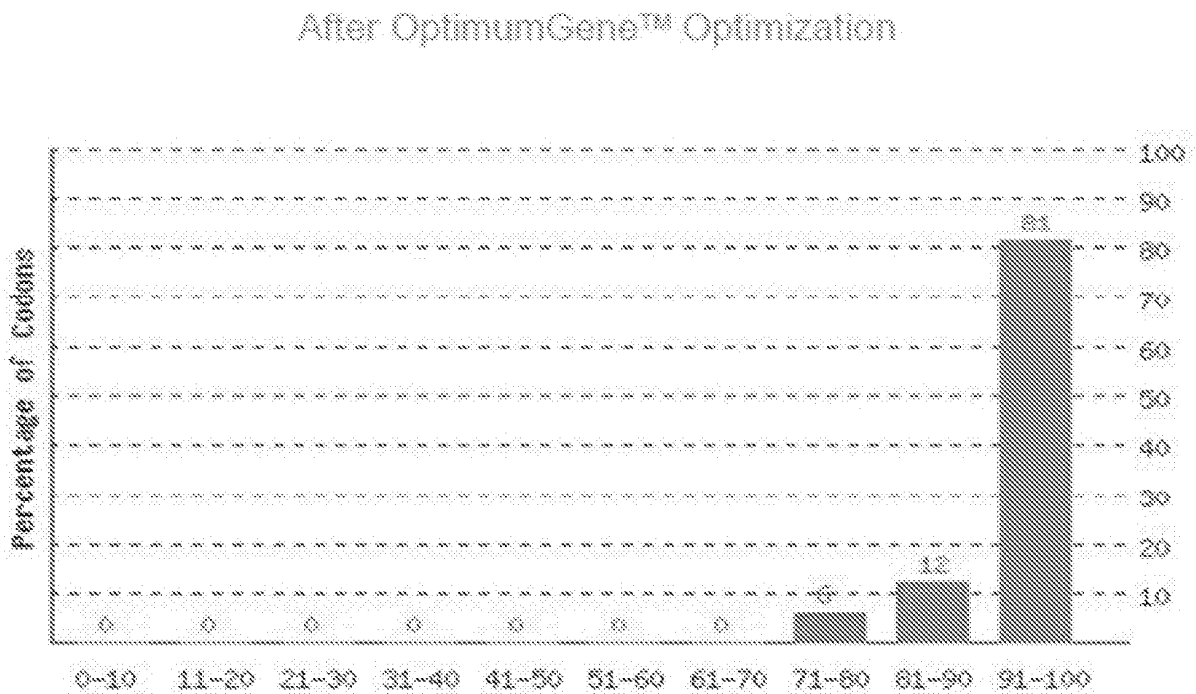


FIG. 77B

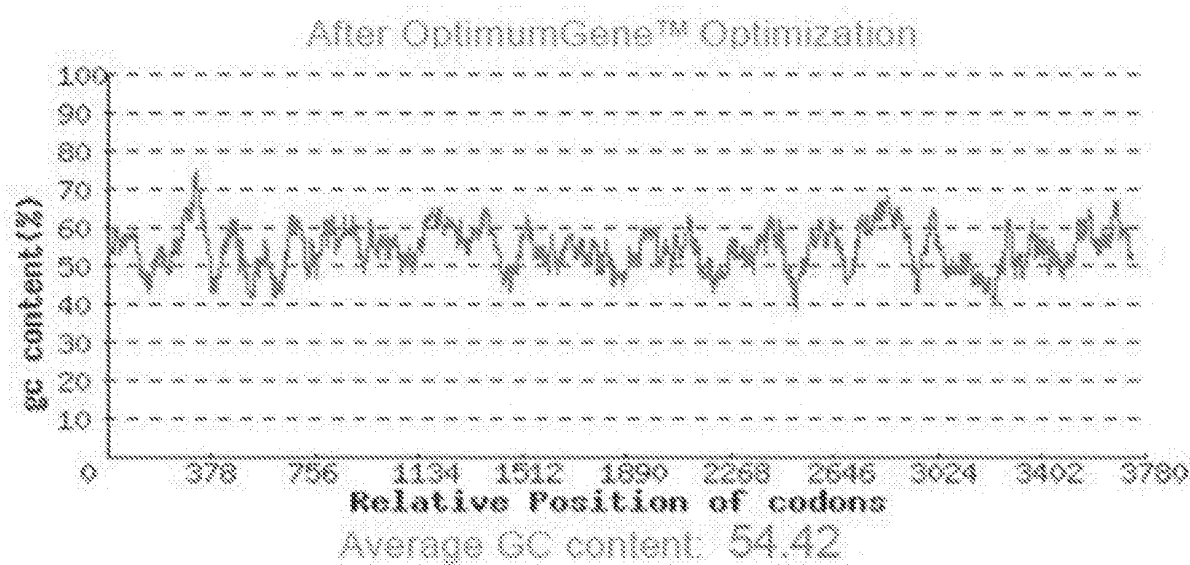


FIG. 77C

Restriction Enzymes	Optimized
<i>* Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.</i>	
BamHI(GGATCC)	1(3835)
EcoRI(GAATTC)	1(3925)
HindIII(AAGCTT)	0
BsmBI(CGTCTC)	0
Bsal(GGTCTC)	0
BbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(3899)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BsrGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0

CIS-Acting Elements	Optimized
Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0

Antiviral Motifs	Optimized
	0

FIG. 77D

After Optimization

Max Direct Repeat: Size:13 Distance:2043 Frequency:2
 Max Inverted Repeat: Size: 13 Tm: 41.3 Start Positions: 1469, 2693
 Max Dyad Repeat: None

FIG. 77E

GenBank

ATGGACAGGCTGAAGGATTTTACCAACCTGTACCCCGTGTCCAAGACACTGCGGTTTGGAGCTGAAGCCTGTGZDC
 AAGACCCCTGGAGAATATCGAGAAAGCCCGCATCCTGAAAGGAGGATGAGCCACAGAGCCGAGAGCTACCCGAGAGTG
 AAGAAGATCATCGATACATATCACAAAGTGTTCATCGACAGCTCCCTGGAGACATGGCCAGATCGGCATCGAG
 AATGAGATCAAGGCCATGCTGCACTCCTTTTGGAGCTGTATAAGAGAGGACTACAGGACCGAGGAGAGGACAAAG
 GCGCTGGATAAGATCAAGGCGGTGCTGAGGGGCGTGTGCTGAGAGCTTTCACCGGCGTGTGCGGCGCGGAG
 AACACAGTGCAGAAATGAGAACTATGAGAGCCCTGTTTAAGGAGAGCTGATCAAGGAGATCCTGCCAGATTTCTGT
 CTGTCTACAGAGGCCGAGTCCCTGCCCTTTTCTGTGAGAGGAGCCACCGAAGDCTGAAGGAGTTCGACTCCTTT
 ACATCTTACTTTCGCGGCTTTTATGAGAACCGGAGAGATATCTACTCTACCAAGCCCCGAGGACACAGCCATCGCC
 TATAGACTGATCCACGAGAACCTGCTAAGTTTCATCGAFAATATCCTGGTGTTCAGAAAGATCAAGGAGCCAAATC
 GCGAAGGAGCTGGAGCACATCAGGCGAGACTTCAGCGCCCGCGCTACATCAAGGAGGATGAGCGCTTGGAGGAC
 ATCTTTTCCCTGAAGTACTATATCCAGCTGTGTCTCAAGCCGCGCATCGAGAGGTACAATGCGCTGATCGGCAAG
 ATCCTGACCGAGGCGGATGGCGAGATGAGGCGCTGAAGGAGCACATCAACCTGTATAATCAGCAGGGGGGCTTC
 GAGGACCGGCTGCCACTGTTCAGAGCCCTGTATAAGCAGATCCTGTCTGATAGGAGCAGCTGTCTATCTGCCA
 GAGTCTTTGAGAGGAGCAGGAGCTGCTGAGGGCCCTGAAGGAGTTTACGATCACATCGCAGAGGACATCCTG
 GGAAGGACCCAGCAGCTGATGACAAGCATCTCGAGTACGATCTGTCCCGGATCTATGTGAGAAACATAGCCAG
 CTGACCGACATCTCCAAGAGATGCTGGGCGATTGGAAATGCCATCTACATGGCCCCGGAGAGAGGCTATGACCAC
 GAGCGAGCCCCCAAGGCGATCACAGCCAGTACAGAGGAGCCGATCAAGGCGCTGAAGGCGGAGGAGTCTATC
 AGCCTGGCCBACCTGAACAGCTGCATGCGCTTCTTGGACAACTGTAGGGATTTGTGCGTGGACACCTATCTGTCT
 ACCTTGGGACAGAAAGGAGGACCTCAAGGCTGAGCAACTGGTGGAGAACTGTTCGCTCTTACCAAGAGGCC
 GAGCAGCTGCTGTCTTTTCCCTATCCTGAGGAGAACATCTGATCCAGGACAAGGATAACGTGGTGTGATCAAG
 AACCTGCTGGATAATATCAGCGACTGCGAGAGGTTCTGAGAGCCACTGTGGGCGCATGGGCGATGAGCCCGACAG
 GATGAGAGGTTTTACGCGAGTACAAATTATACAGGGCGGCGCTGACCAAGTCACTCCCTCTGTATAACAGGTTG
 CGGAATTATCTGACCCGCAAGCCATACTCCACACGCAAGGTGAAGCTGAACTTCGGCAATAGCCAGCTGCTGTCC
 GGCTGGGATAGGAACAAAGGAGAAAGACAAATTCTTGGCTGATCCTGGCGAAGGGCCAGAACTTCTACCTGGCCATC
 ATGAACAATCGGCACAAAGGAGCTTCAGGAATAAGATCTGCGCGAGTATAAGGAGGGCGAGCCCTTACTTCGAG
 AAGATGGATTATAAGTTTCTGCCAGACCCCAACAAGTGTCTGCCCAAGTTTCTCTCTAAGAAAGGATCGAG
 ATCTACAAGCCTAGCCCAAGCTGCTGGAGCAGTATGCGCACCGCCACCAAGAAAGGGCGATACCTTCAGCATG
 GACGATCTGCACGAGCTGATCGACTTCTTTAAGCACTCCATCGAGGCCCAAGGAGATTGGAAGCAGTTCCGCTTT
 AAGTTCAAGGACACCGCCACATAAGAGAACTGAGCAGCTTCTACCGGAGGTTGAGGACCAAGGCTACAAGCTG
 TCTTTTAGAAAGGTTGTCGAGTCTTACGTGTATAGCTGATCGATCAGGGCAAGCTGTACCTGTTCAGATCTAT
 AACAAAGGACTTTAGCCCTTGTTCAGAGGCGACCCCAAACTCTGCACACACTGTACTGGGGGATGCTTTTCGATGAG
 AGAAACCTGGCCGACCTGATCTATAAGCTGGATGGCAAGGCCGAGATCTTCTTTCCGGAGAGTCCCTGAAGAT
 GAGCACCCCAACCCACCTGCGAGCAAGCCCATCAAGAGAGAGAGCCGGCAGAAAGAGGGCGAGAGGCGCTGTT
 GAGTACGATCTGTTGAAGGACCGGAGATATACCATGGATAAGTTTCACTTCCAGTGGCCANTCACAAATGAACCTT
 AAGTGTCTCTGCGGCAAGGAGGAGCAATGGTGAATGCCACATCAGGAGAGGCCAAGGACATGCAAGTGAATC
 GGCATCGATAGGGGCGAGCCGAATCTGCTGTATATCTGCGTATCGACAGCCCGGCGCCATCTCGATCAGATC
 TCCCTGGAACACAAATCAATGACATCGATTATCAGGATCTGCTGAGTCCAGGACAAAGGATCGCTACGAGGAGCAC
 AGGAAGTGGCAGACCATCGAGGGCATCAAGGAGCTGAAGCAAGGCTACCTGTCTCAGGCGGTGACCCGATCGCC

FIG. 77F

GAGCTGATGCTGGCCCTATAAGGCCCTTGGTGGCCCTGGAGGACCTGAACATGGGCTTCAAGCGGGGCGAGACAGAAG
 GTGGAGAGCCAGCGTGTATCCAGCACTTTGAGAAAGCAGCTGATCGACAAGCTGAATTATCTGCTGGATAAGAAGAAG
 CTGCGCTGAGGACATCGGAGGCTTGGCTGAGAGCCTTACCAGTTCAACCGGCCCTTTCAAGAGCTTTAAGGAGATGTC
 AAGCAGAAACCGCTTCTGTTCTATATCCCTGCGCTGGAACACATCCAAATATCGACCTCAACCAAGGCTTCTGTGAAC
 CTGTTTCACCTGCACTACGAGATGTGGATAAGGCCAAGAGCTTCTTTCAGAAATTCCACAGCATCTCCTTACAAC
 CCTAAGAAAGGATTTGGTTTGAGTTCCGCTTTGACTATAAGAACTTCACCAAGAAGGCCGAGGCTCTAGGAGCATG
 TGGATTCTGTGCCACCCACCGCTCCCCGATCAAGAACCTTCAGAAATTCTCAGAAAGAAATGGCCAGTGGGATAGCGAG
 GAGTTTGGCCCTGACCCGAGGCCCTTCAAGTCCCTGTTTGTGCGTTACGAGATCGATTATACCGCCGACCTGAAAACCT
 GCCATCGTGGACCTAGNAGCAGAAAGGATTTCTTTTGTGGACCTGCTGAAGCTTTTCAAGCTTGACCGTGCAGATGAGA
 AACTCCTGGAAAGGAGAAAGGACCTGGATTACCTGATCTCTCCAGTGGCTGGGCGCCGATGGCAGGTTCTTTGACACA
 CGCGAGGCGCAATAAGAGGCTTGCCTCAAGGACCTCAGATGCCAAACCGAGCCCTATAATATCCGCTTGAAGGGGCTGTGG
 GCACTGAGGCCAGATCAGACAGACCTCCGAGGCGGGCAAGCTGAAGCTGGCCATCTCTAACAAAGGAGTGGCTGCAG
 TTTGTGCAGGAGAGATCCTACGAGAAAGGAC

AAAAGGCTTCTGCTCTCTGAAAGAGGCGGCTGGGCTGAAAGAGGAAAGAGGCTCTTACCTCTTACCTGTTTCCAGCT
 TACGCTTACCTCTTACCACTTGGCTGATTAACAGACCTATAAGATGTGCGGCACTATGCTTAAGGATTC

FIG. 77G

MDSLKDFTNLYPVSKTLRFELKPVGKTLENIKAGILKEDEHRAESYRPRVKKIIDTYHKV
 FIDSSLENMAKMGIEENEIKAMLQSFCELYKKDHRTEGEDKALDKIRAVLRGLIVGAFTGV
 CORRENTVQNEKYESLFKEKLIKEILPDFVLSTEAESELPFSVEEATRSLEKFDSTSYFA
 GFYENRKNYISTKPKQSTAIAYRLIHENLPKPIDNVLVPQKIKEPIAKELEHIRADFSAGG
 YIKKDERLEDIFSLNYIYHVLSQAGIEKYNALIGKIVTEBGDGEMKGLNEHINLYNQQRGR
 EDRLPLFRPLYKQILSDREQLSYLPESFEKDEELLRALKEFYDHIAEDILGRTQQLMSTSI
 SEYDLSRIYVRNDSQLTDISKMLGDWNAIYMARERAYDHEQAPKRITAKYERDRIKALK
 GEESISLANLNSCIAFLDNVRDCRVDTYLSTLGQKEGPHGLSNLVENVFASYHEAEQLLS
 PPYPRENNLIQDKDNVVLIKNLLDNISDLQRFKPLWGMGDEPDKDERFYGEYNYIRGAL
 DQVIPLYNKVRNYLTRKPYSTRKVKLNPGNSQLLSGWDNRNKEKDNSCVILRKGQNFYLAI
 MNNRHKRSFENKMLPEYKEGEPYFEKMDYKFLPDPNKMLPKVPLSKKGIETIYKPSPKLLE
 QYGHGTHKKGDTFEMDDLHELIDFFKHSIEAHEDWKQFGPKFSDTATYENVSSFYREVED
 QGYKLSFRKVSESYVYSLIDQGKLYLFQIYNKDFSPCSKGTPLNHTLYWRMLFDERNLAD
 VIYKLDGKAEIFFREKSLKNDHPTHFAGKPIKKKSRQKKGEESLFEYDLVKDRREYTMDF
 QFHVPITMNFKCSAGSKVNDMVNAHIREAKDMHVIGIDRGERNLLYICVIDSRGTILDQI
 SLNTINDIDYHDLLESRDKDFQQEHRNWQTIEGIELKQGYLSQAVHRIAELMVAYKAVV
 ALEDLNMGFKRGRQKVESVYQQFEKQLIDKLNLYVDKKKRPEDIGQLLRAYQFTAPFKS
 FKEMGKQNGFLFYIPAWNTSNIDPTTGfVNLfHVQYENVVDKAKSPFQKFDSISYNPKKDW
 FEPAPDYKNPTKKAEGSRSMWILCTHGSRIKNFRNSQKNCQWDSEEFALTEAFKSLFVRY
 EIDYTADLKTAIIVDEKQKDFFDLLKLFLKLTVMRNSWKEKDLDYILISFVAGADGRFFDT
 REGNKSLEPKDADANGAYNIALKGLWALRQIRQTSEGGKLKLAISNKEWLQFVQERSYEK

FIG. 77H

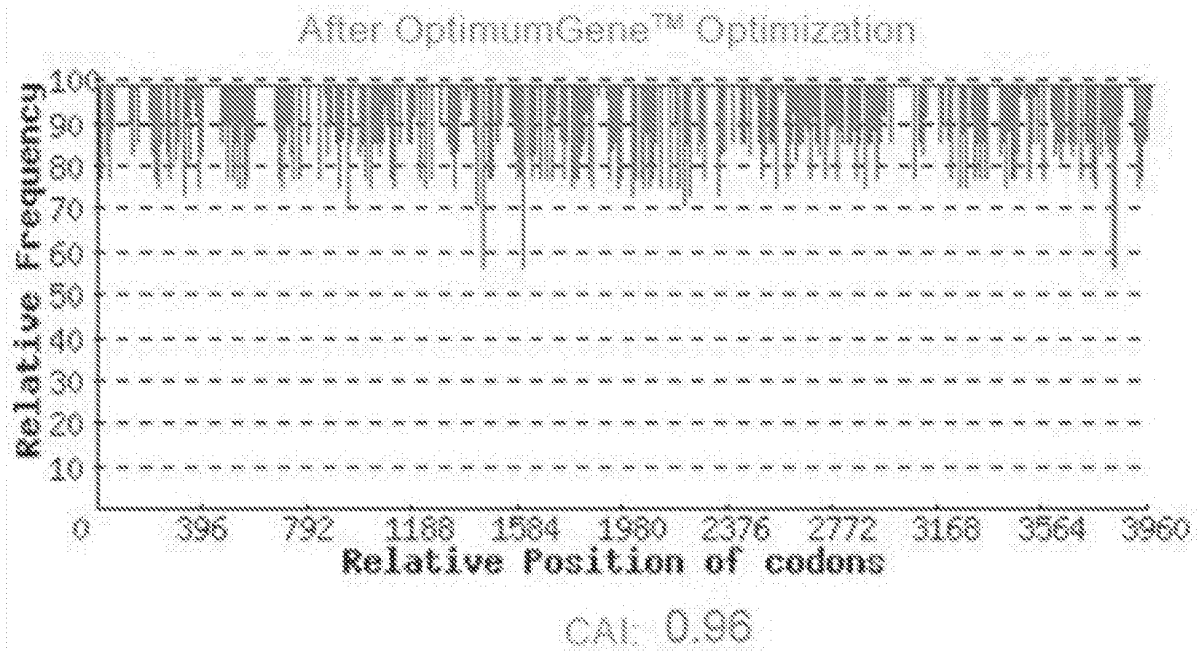


FIG. 78A

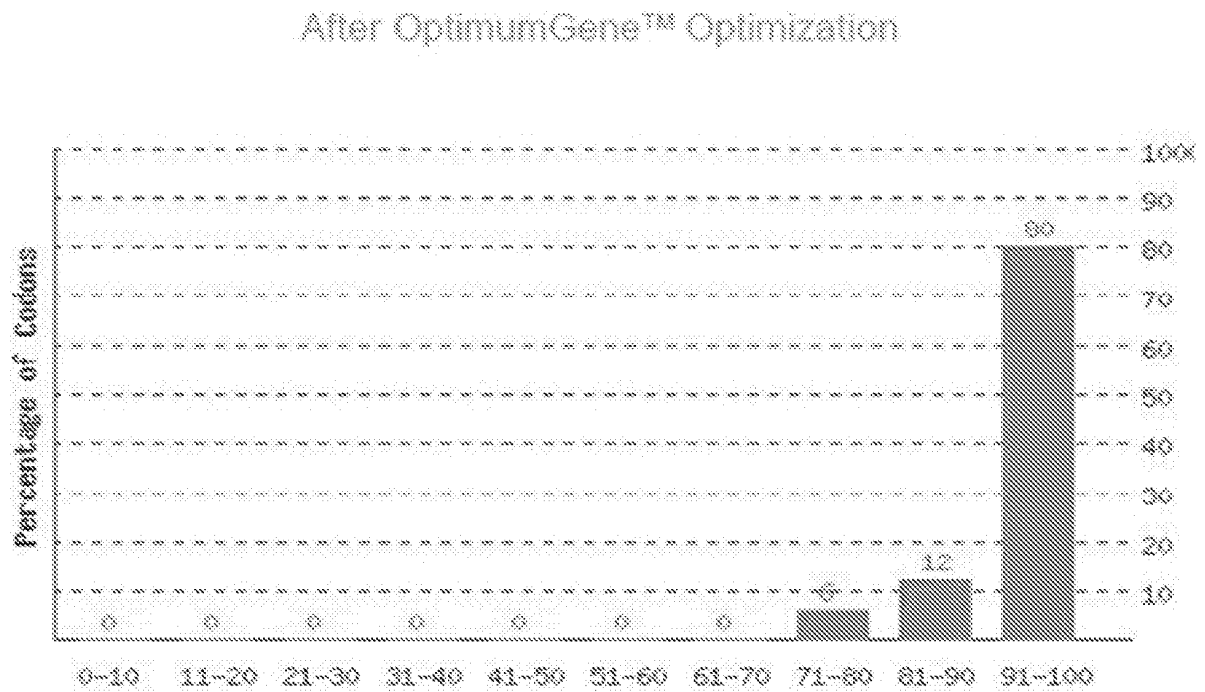


FIG. 78B

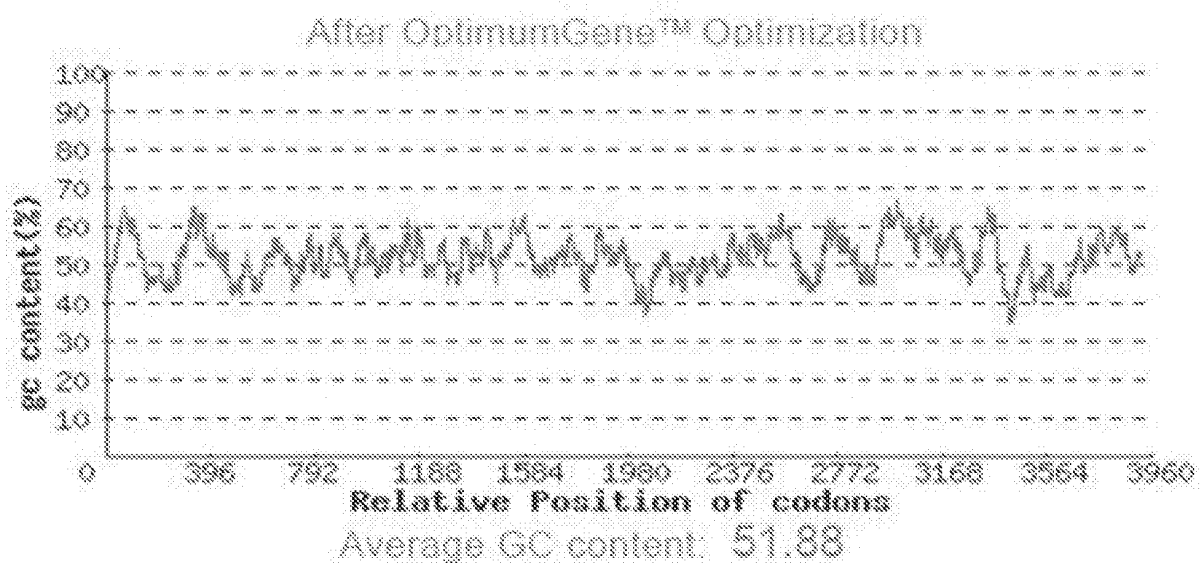


FIG. 78C

Restriction Enzymes Optimized

* Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.

BamHI(GGATCC)	1(4024)
EcoRI(GAATTC)	1(4114)
HindIII(AAGCTT)	0
BsmBI(CGTCTC)	0
BsaI(GGTCTC)	0
BbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(4088)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BsrGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0

CIS-Acting Elements Optimized

Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0

Antiviral Motifs Optimized

0

FIG. 78D

After Optimization

Max Direct Repeat: Size:15 Distance:1842 Frequency:2

Max Inverted Repeat: None

Max Dyad Repeat: None

FIG. 78E

GGTACG

ATGGAGAACTATCAGGAGTTCAACAACTGTTTCAGCTGAATAAGACACTGAGATTGAGCTGAAGCCCATGXC
 AAGACCTGCGAGCTGCTGGAGGAGGGCAAGATCTTCGCCAGCGGCTCCTTTCTGGAGAAGGACAAGGTGAGGGCC
 GATAACCTGAGCTACTGGAAGAAGGAGATCGACAGAAGACACAAGATCTTTATCGAGGAGSACACTGAGCTCCTTC
 TCTATCAGCAACGATCTGCTGAAGCAGTACTTTGACTGCTATAATGAGCTGAAGGCCCTTCAAGAAGGACTGTAAAG
 AGCGATGAGGAGGAGGTGAAGAAAACCGCCCTGCGGCAACAAGTGTACCTCCATCCAGAGGCCCATGCGCGAGGCC
 ATCTCTCAGGCCCTTTCTGAAGAGCCCCCAGAAGAAGCTGCTGGCCATCAAGAACCTGATCGAGAACGTGTTCAGG
 GCCGACCGAGAATGTGCAACACTTTCTCCGAGTTTAACASCTATTTCTCCGCTTTGAGACAAACAGAGAGAATTTTC
 TACTCTGACCGAGGAGAGTCCACATCTATCGCCCTATAGGCTGGTGCCACGATAACCTGCCCTATCTTCATCAAGAAC
 ATCTACATCTTGGAGAAGCTGAAGGAGCAGTTGAGGCCCAAGACCTGAGCGAGATCTTGGAGAACTACAGCTG
 TATGTGGCCGCTCTAGCCTGGATGAGGTGTTCTCCCTGGAGTACTTTAACAATAACCTGACACAGAGGGCATC
 GACAACTATAATGCGCTGATCGGCCAAGATCGTGAAGGAGGATAAGCAGGAGATCCAGGGCTGGAACGACCATC
 AACCTGTATAATCAGAAGCCACAAGGACCGGAGACTGCCCTTCTTTATCTCCCTGAAGAAGCAGATCCTGTCCGAT
 CCGGAGGCCCTGTCTTGGCTGCCCTGACATGTTCAAGAAATGATTCTGAAGTGATCAAGGCCCTGAAGGGCTTCTAC
 ATCGAGGACCGCTTTGAGAACAATGTGCTGACACCTCTGGCCACCCCTGCTCTCTCTGATAAGTACAACCTG
 AATGGCATCTTTATCCGCAACAATGAGGCCCTGAGCTCCCTGTCCGAGAACGTGTATCGGAATTTTTCTATCGAC
 GAGGCCATCGATGCCAAACCGCGAGCTGCGAGACCTTCAACRATTACGAGCTGATCGCCAAATGCCCTGCGGCCAAG
 ATCAAGAAAGGAGACAAAGCAGGGCCGGAAGTCTTTGAGAAGTACGAGGAGTATATCGATAAGAAAGGTGAAGGCC
 ATCGACAGCCCTGTGCTCCAGGAGATCGACAGGCTGGTGGAGAAATTACCTGAGCGAGTTTAACTCTAATAGCGGC
 AACATGCCAAGAAAGGTGGAGGACTACTTCAGGCTGATGAGGAAAGGGCTACTTCGGCTCCAAAGATCTGATCGAA
 AATATCAGACCAAGCTGAGCGCCGCGAGGAAGCTGCTGGGCAAAAGTACCAAGGAGACAGCCCAAGGACATCTTC
 AAGAAAGATGAGAACTCCAGCTGATCAAGGAGCTGCTGGAGCGCAACCAAGCAATTCGAGCACTTTATCGAGCCA
 CTGCTGGGCAACAGGCGAGGAGGAGATCGGGACCTGGTGTCTACGGCGATTTTCTGCCCTGTATGAGAACTT
 GAGGAGCTGACCCCTGCTGTATATAACAAGGTGCGGAATGAGTGGACACGGAAGCCCTATTCCAGGACAGATCCGC
 CTGTGCTTCAACAAGCCTAAGCTGATGACAGGCTGGGTGGATTCCAAGACCGAGAAGTCTGACAAACCGCACACAG
 TACGGCGGCTATCTGTTTCGGAAGAAGAAATGAGATCGGCGAGTACGATTATTTCTGGGCATCTCTAGCAAGGCC
 CAGCTGTTTCAGAAAGAAAGAGCCGTGATCGGCCACTACGAGAGGCTGATTAATCTATCAGCCAAAGGCCAATACC
 ATCTACGGCTCTGCCCTATGAGGGCGAGAACAGCTACAGGAGGACAGGAAGCGGCTGACAAAGTGATCATCGCC
 TATATCGAGCAGATCAAGCAGACAAACATCAAGAAGTCTATCATCGAGTCCATCTCTAAGTATCCTAATATCGAC
 GACGATGACAGGTGACCCCATCTCTCTGCTGGAGAGATCAAGAAGGTGTCTATCGACAGCTACAACGGCATC
 CTGTCTTCAAGTCTTTTCAGAGGCTGACAAAGGAAGTGAATCGATAACCTGCTGAAAACCTTCAGGCCCTGAGG
 AACAGGCCGAGTTTCTGAGCTGATCAATAAGGATTATCAGATCTTCACCGAGGTGCGAGCCGTGATCGAGGAG
 ATCTGCAAGCAGAAACCTTCATCTACTTTCCAATCTCCAAAGTGGAGCTGGAGAAGGAGATGGGCGATAAGGAC
 AAGCCCTGTGCTGTTCAGATCAGCAATAAGGATCTGTCTTCCGCAAGACCTTTAGCGCCAACTGCGGGAAG
 AAGAGAGGCGCGCGAGATCTGCACACAATGCTGTTTAAGGCCCTGATGGAGGCGCAACCAAGATAATCTGGACCTG
 GGCTCTGGCGCCATCTTCTACAGAGCCAAAGGCTGGACGGCAACAAGCCACACACCCTGCCAATGAGGCCATC
 AAGTGTAGGAACGTGGCCCAATAAGGATAAGGTGTCCCTGTTCACCTACGACATCTATAAGAACAGGCCCTACATG
 GAGAAEAAATTCCTGTTTCACCTGAGCATCGTGAGAACTATAAGGCCGCCAATGACTCGGCCAGCTGGAACAGC
 TCCGCCACCGAGTATATCAGAAAGGCCGATGACCTGCACATCATGGGCATCGATAGGGGCGAGCGCAATCTGCTG

FIG. 78F

FIG. 78G

MENYQEPTNLPQLNKTLRFELKPIGKTCELLLEEGKIFASGSFLEKDKVRADNVSYVKKBI
 DKXHKIFIEETLSSFSISNDLLKQYFDCYNELKAFKKDCKSDREEVKKTALRNKCTSIQR
 AMREAISQAFLKSPQKKLLAIKNLIENVFKADENVQHFSEFTSYFSQFETNRENFYSDDEE
 KSTSIAYRLVHDNLPFIPIKNIYIFEKLKEQFDAKTLSEIFENYKLYVAGSSSLDEVFSLEY
 FNNLTLTQKGIDNNAVIGKIVKEDKQEIQGLNEHINLYNQKHKDRRLPFFISLKKQILSD
 REALSWLPDMFKNDSEVIKALKGFIYIEDGFENNVLTPLATLLSSSLDKYNLNGIFIRNNEA
 LSSLSQNVYRNFSIDEAIDANAELQTFNNYELIANALRAKIKKETKQGRKSFEKYREYID
 KKVKAIDSLSIQEINELVENYVSEFNSSNGNMPRKVEDYPSLMRKGDGFSNDLIENIKTK
 LSAAEKLLGTKYQETAKDIFKKDENSCLIKELLDATKQFQHFIPKPLLGTGEEADRDLVFY
 GDFLPLYEKFSBELTLLYNKVRNRLTQKPYSKDKIRLCFNKPKLMTGWVDSKTEKSDNGTQ
 YGGYLFRKKNSEIGEYDYFLGISSKAQLFRKNEAVIGDYERLDYYQPKANTIYGSAYEGEN
 SYKEDKKRLNKVIIAYIEQIKQTNIKKSLIESISKYPNISDDDKVTPSSLLEKIKKVSID
 SYNGILSFKSPQSVNKEVIDNLLKTISPLKNKAEFLDLINKDYQIFTEVQAVIDEICKQK
 TFIYFPFISNVELEKEMGDGDKPLCLFQISNKDLSPAKTFSANLRKKRGAENLHTMLPKAL
 MEGNQDNLDLGSGAIFYRAKSLDGNKPTHFANEAIKCRNVANKDKVSLFTYDIYKNRRYM
 ENKFLFHLISIVQNYKAANDSAQLNSSATEYIRKADDLHIIGIDRGERNLLYYSVIDMKGN
 IVEQDSLNIIRNNDLETDYHDLLDKREKERKANRQNWAVEGIDKDKGYLSQAVHQIAQ
 LMLKYNAILALEDLGQMFVTRGQKIEKAVYQOFEKSLVDKLSYLVDKKRFYNELGGILKA
 YQLASSITKNNSDKQNGFLFYVPAWNTSKIDPVTGPTDLLRPKAMTIKEAQDFFGAFDNI
 SYNDKGYFEPETNYDKFKIRMKSAQTEWTICTFGNRIKRKKDKNYWNYEEVELTEEFKKL
 FKDSNIDYENCNLKEEIQNKDNRKFFDDLIKLLQLTLQMRNSDDKGNDYIISFVANAEGQ
 FFDSRNGDKKLPLDADANGAYNIARKGLWNIRQIKQTKNDKKLNLSISSTEWLDFVREKP
 YLK

FIG. 78H

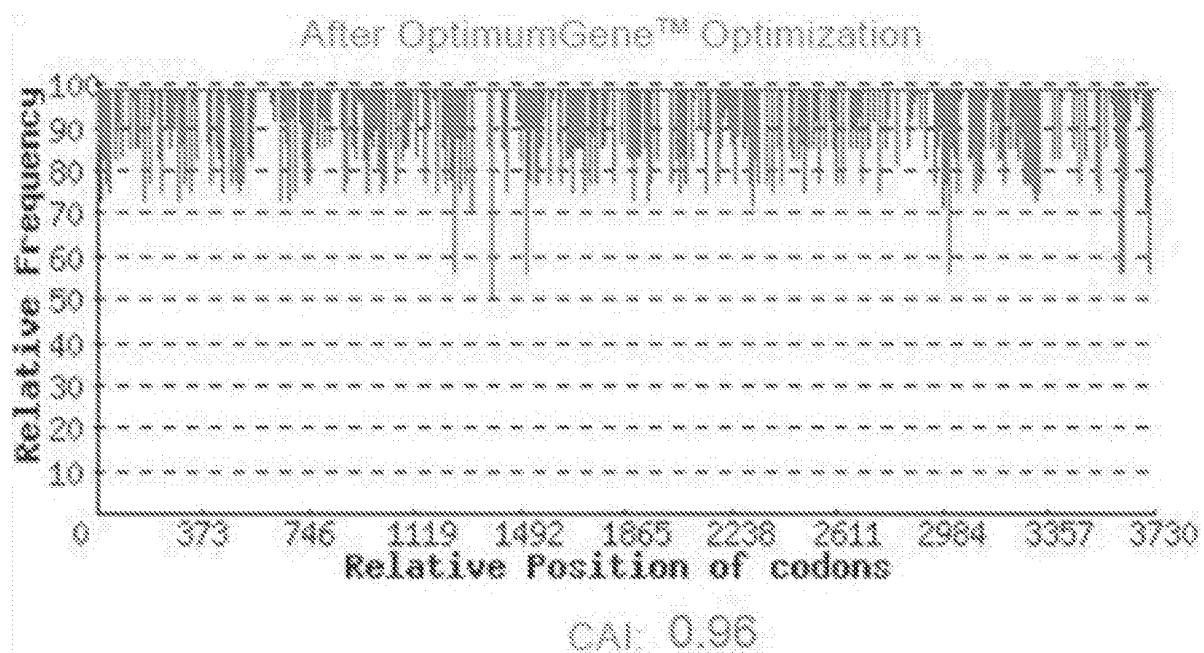


FIG. 79A

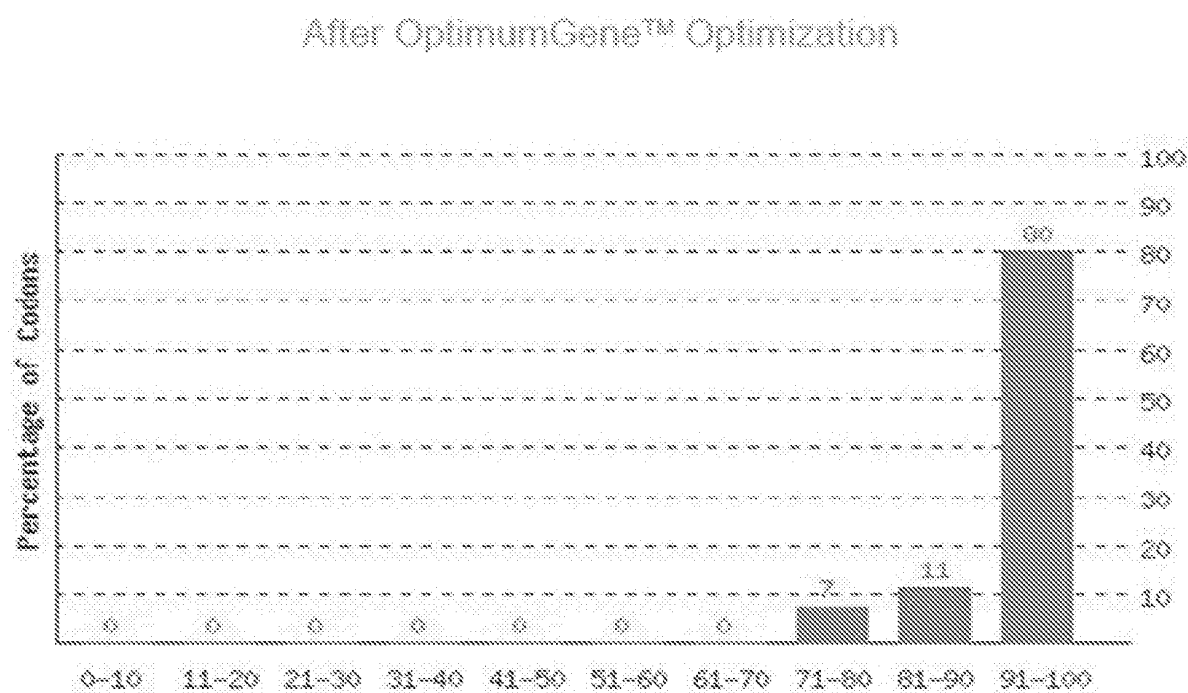


FIG. 79B

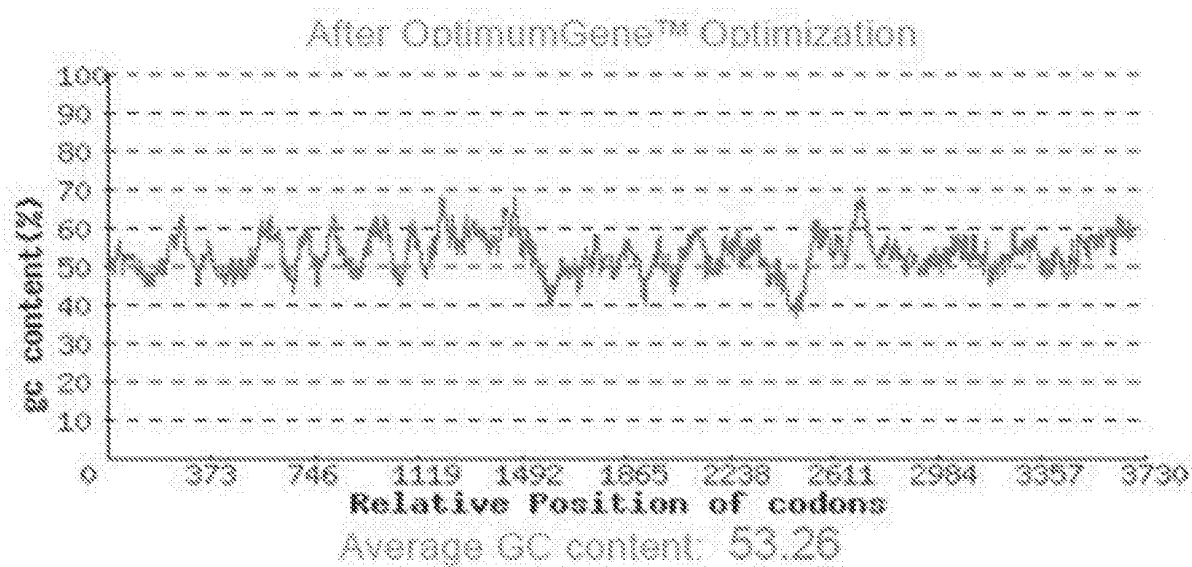


FIG. 79C

Restriction Enzymes	Optimized
* Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.	
BamHI(GGATCC)	1(3793)
EcoRI(GAATTC)	1(3883)
HindIII(AAGCTT)	0
BsmBI(CGTCTC)	0
BsaI(GGTCTC)	0
BbsI(GAAGAC)	0
AgeI(ACCGGT)	0
XhoI(CTCGAG)	0
NdeI(CATATG)	1(3857)
NotI(GCGGCCGC)	0
KpnI(GGTACC)	1(1)
BsrGI(TGTACA)	0
SpeI(ACTAGT)	0
XbaI(TCTAGA)	0
NheI(GCTAGC)	0
ARE	0

CIS-Acting Elements	Optimized
Splice(GGTAAG)	0
Splice(GGTGAT)	0
PolyA(AATAAA)	0
PolyA(ATTAAA)	0
Destabilizing(ATTTA)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0

Antiviral Motifs	Optimized
	0

FIG. 79D

After Optimization

Max Direct Repeat:	Size:13 Distance:99 Frequency:2
Max Inverted Repeat:	None
Max Dyad Repeat:	None

FIG. 79E

FIG. 79F

GAGAACTTGAACTCTGAGCTTCATGAAGGGCCAGGAAGGAGGTTGGAGAGTCCCTGTGTACGAGAACTTTGAGCGCATG
 CTGCTGGACAAAGCTGAACCTATCTGGTGGGATAAGGAAGAACTGTTCCAATGAGGCCAGGAGGCTGTGTATGCAGCA
 TACCAGCTGACCAATCCACTTCTCTCTTTGAGGAGCTGCACAGATAACCCCTCAGAGCGGCATCTCTGTTTTCTG
 GACCCATGGAAACACCTCTCTGACAGATCCCAGCACAGGCTTCCTGGAATCTGCTGGGACAGATCNACTACACCAAT
 GTGGGCGAGCGCCCGCAGTTTTTCTGATCTTTTAAAGGACATCAGATATGACGGCAAGGGCAATATCTCTTTCTGAC
 CTGGAATCTGTCCAGATTTGATTTGAGGCTGGAGACACAGAGGAAAGCTGTGGAACACTGACCACTTCGCTCTGAC
 ATCGCCAAATCCAAGAGTCTGGCAAGTGTGATGCTGGAGCGGATCGAGAACTTGAGGCTGTGCTTCTGAGGCTG
 TTGAGGCACTTTTAATATCCGCTACAGAGTGGAGAAAGGACCTGAGAGAAAGGCCATCTGAGGCCAGGATAGGAAAGGAG
 TTCTATGTTGCGGCTGATCTACCTGTGTTAACTGATGATGACAGATCCCGAAGCAGCGAGCGGCTAGGAGGATTAATC
 CTGTTCTCCCGCCCTGAACGAGAAAGAACTCTGCACTTCCAGACAGGCTGATCGAGGCCAAGGATCTGCTGTGTGGAC
 GCAGATGCAAAACCGAOCATACAAATGTGCCCCCAAGGCGCTGATGCTGCTGCAGAGAAATCAAGAGCGGCGGACCCAC
 GAGTCCATCCACAGGATCCGAAAGGGCACAGTGGCTGAGATATGTCCAGGAGGSCATCTGTGGAG
 AAGAGGCTGCTCTACTGAAGAGCTGTGCTAGCTAAGAAAGAAAGGCTGATCTGATACATCTGCTGAG
 TACGCTTATTTCTAGGAGCTGCTGATTAATGATAGGCTGATGATCTGCTGAGCTATGCTGAGCTATGCTGAGCTATG

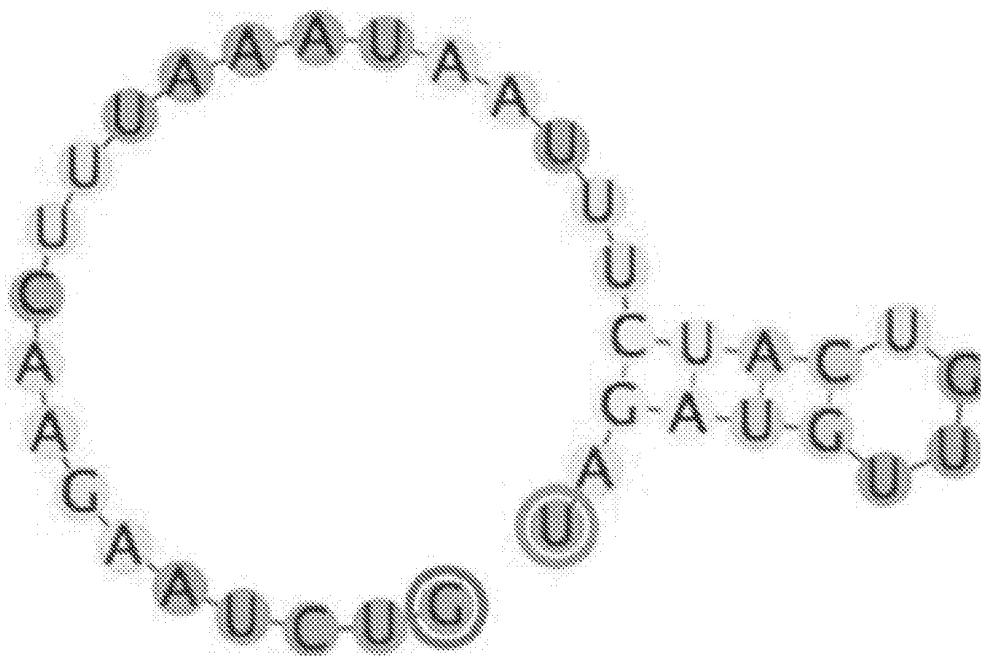
FIG. 79G

MKTQHFFEDFTSLYSLSKTIRFELKPIGKTLENIKKNGLIIRDEQRLDDYEKLKKVIDEY
 HEDFIANILSSFSFSEEILQSYIQNLSESEARAKIEKTMRD TLAKAPSEDERYKSIFKKE
 LVKKDIPVWC PAYKSLCKXFDNFTTSLVPPHENRKNLYTSNEITASIPYRIVHVNLPKFI
 QNIEALCELQKKMGADLYLEM MENLRNVWPSFVKTPDDL CNLKTYNHLMVQSSISEYNRF
 VGGYSTEDG TKHQGINEWINIYRQENKEMRLPGLVFLHKQILAKVDSSSFISDTLENDQ
 VFCVLRQPRKLPWNTVSSKEDDAASLKD LFCGLSGYDPEAIYVSDAHLATISKNI FDRWN
 YISDAIRPKTEVLMPRKKESVERYAEKISKQIKKRQSYSLAELDDLLAHYSEESLPAGFS
 LLSYFTSLGGQKYLVS DGEVILYEEGSNIWDEVLIAFRDLQVILDKDPTEKKLGKDEEAV
 SVIKKALDSALRLRKF FDLLSGTGAEIRDSSFYALYTD RMDKLEGLLKM YDKVENYLT
 KPYSIEKFKLHFDNPSLLSGWDKNKELNNLSVIFRQNGYYYLGIMTPKGKNLFKTL PKLG
 AEEMFYEKMEYKQIAEPMLMLPKVFFFPKTKPAFAPDQSVVDIYNKKTFTKTGQKGPNKKD
 LYRLIDFYKEALT VHEWKLFNF SFSPT EQYRNIGEFFDEVREQAYKVS MVNVPASYIDEA
 VENCKLYL FQIYNKDFSPYSKGI PNHTLYWKALPSEQNQSRVYKLCGGGELFYRKASLH
 MQDTTVHPKGISIHKKNLNKKGETSLFN YDLVKDKRFTEDKFFFHVPISINYKNKKITNV
 NQMVRDYIAQNDDLQIIGIDRGERNLLYISRIDTRGNLLEQFSLNVIESDKGDLRTDYQK
 ILGDREQERLRRRQEWKSIESIKDLKDG YMSQVVHKICNMVVEHKAIVVLENLNL SPMKG
 RKKVEKSVYEKFERMLVDKLN YLVVDKKNLSNEFGGLYAA YQLTNPLPSFEELHRY PQSG
 ILPFVDPWNTSLTDPSTGFVNLLGRIN YTNVGDARKFFDRFN AIRYDGKGNILFDL DLSE
 FDVRVETQ RKLWTLTTFGSRIAKSKKSGKWMVERIENLSLCFLELFRQFNIGYRVEKDLK
 KAILSQDRKEFYVRLIYLFNLMMQIRNSDGEEDYILSPALNEKNLQFDSRLIEAKDLPVD
 ADANGAYNVARKGLMVVQRIKRGDHESIHRIGRAQWLRYVQEGIVE

FIG. 79H

>1 DR

GUCUAAGAACUUUAAAUAUUUUCUACUGUUGUAGAU



>2 DR

GCUAUAAGCCUAUAUAUUUUCUACUAUUGUAGAU

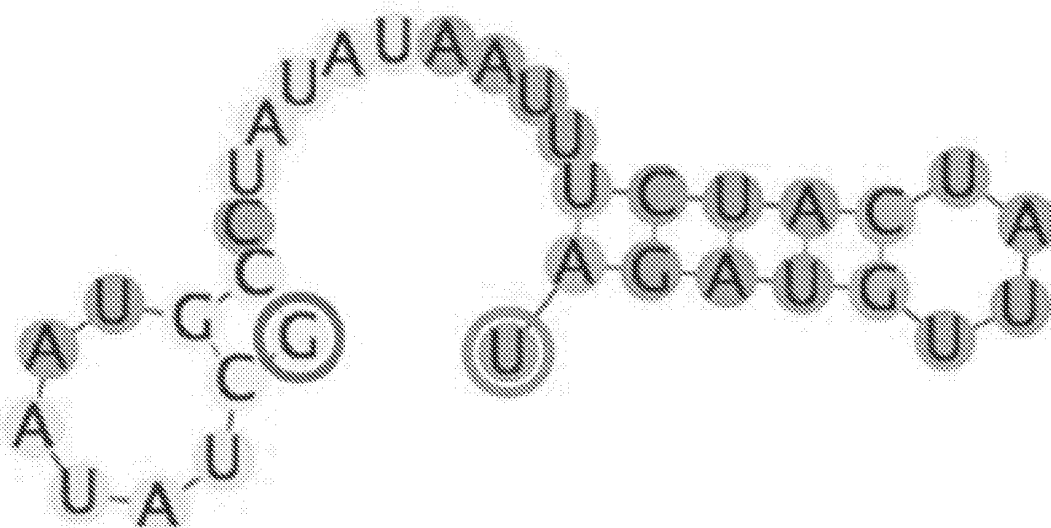
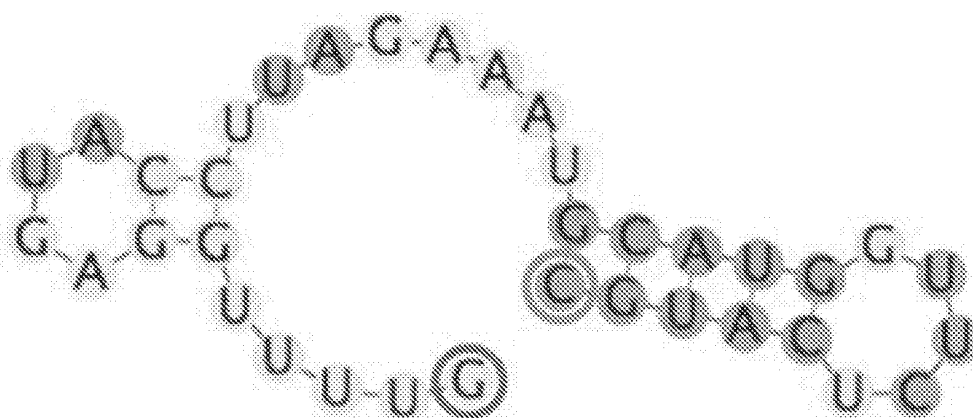


FIG. 80A

>3 DR

GCCUAUAAGGCUUUAGUAAUUUCUACUUAUUGUAGAU



>4 DR

GUCAAUAAGACUCAUUUAAUUUCUACUUCGGUAGAU

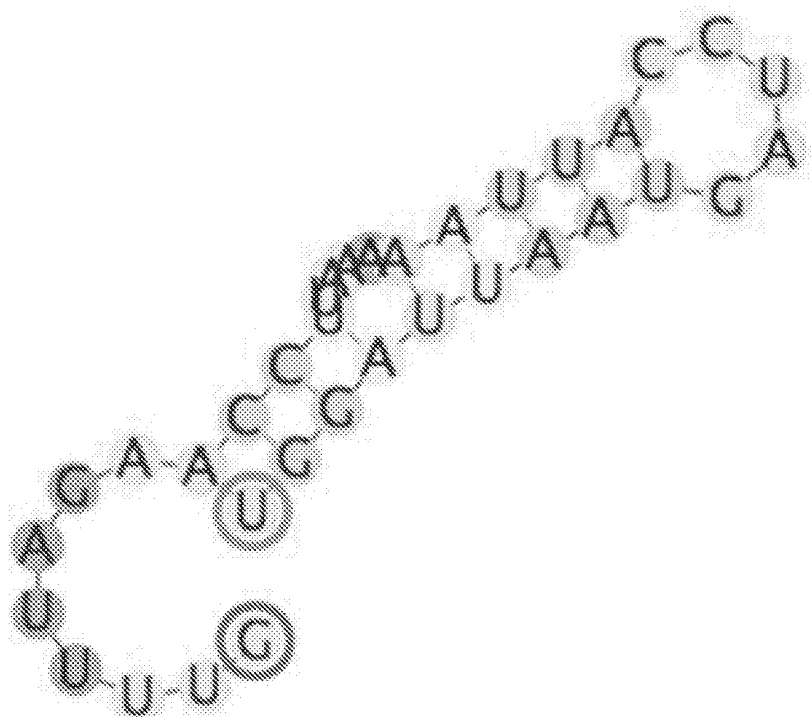
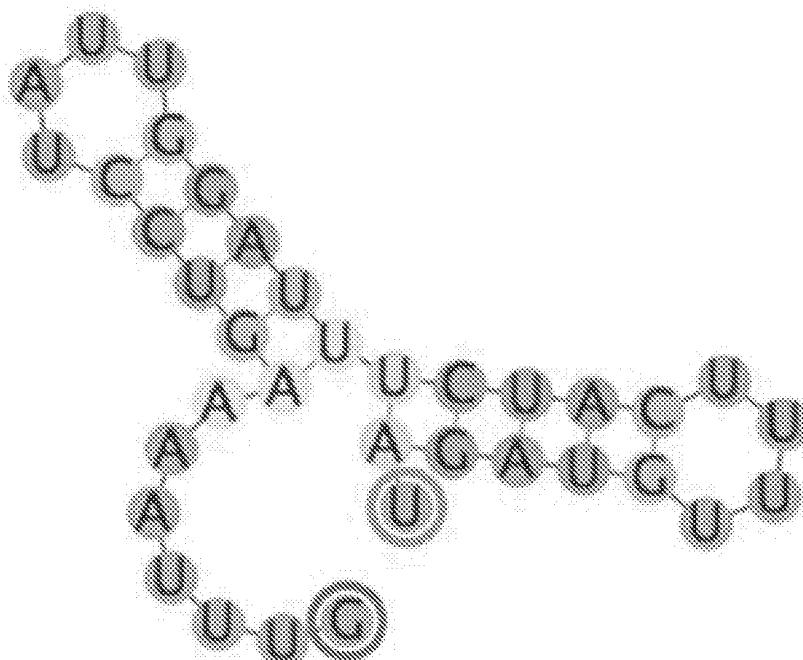


FIG. 80B

>5 DR

GUCUAGGUACUCUCUUUAAUUUCUACUAAUUGUAGAU



>6 DR

GUUUCAAAGAUUAAAUAUUUCUACUAAGUGUAGAU

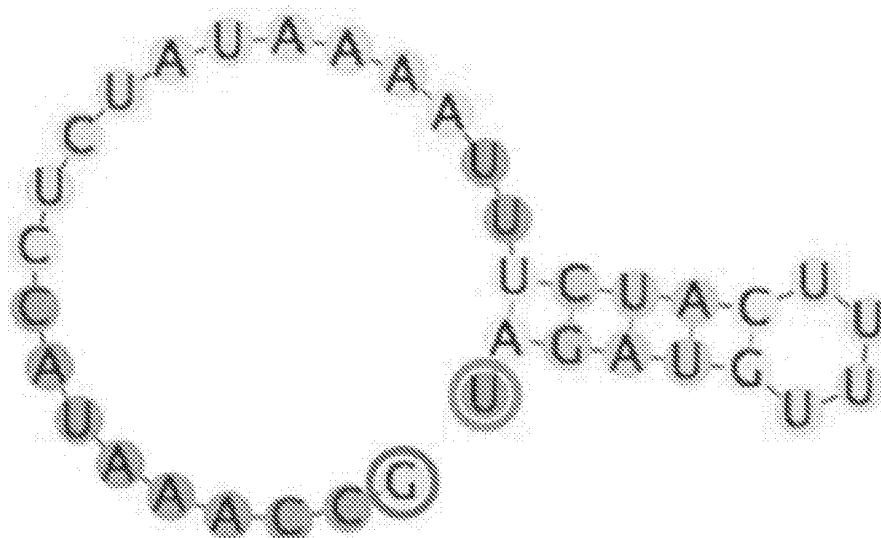
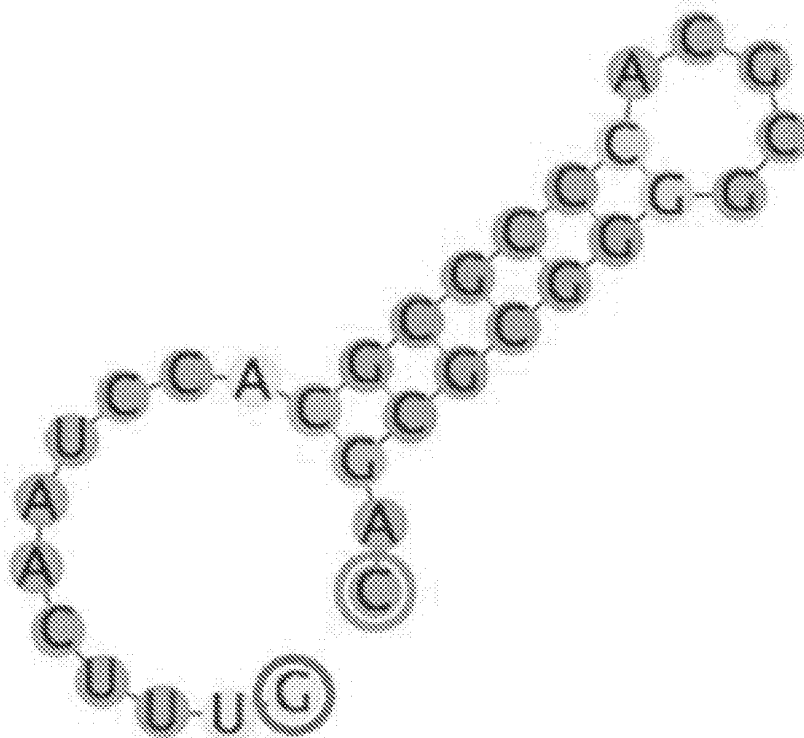


FIG. 80C

>7 DR

CUCUAAAGAGAGGAAAGAAUUUCUACUUUUUGUAGAU



>8 DR

GUCUAACGACCUUUUAAAUUUCUACUGUUUGUAGAU

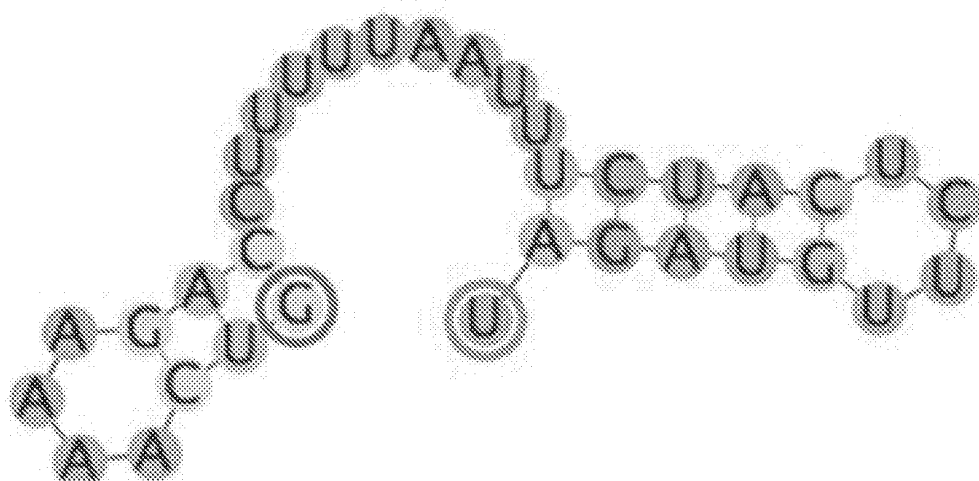
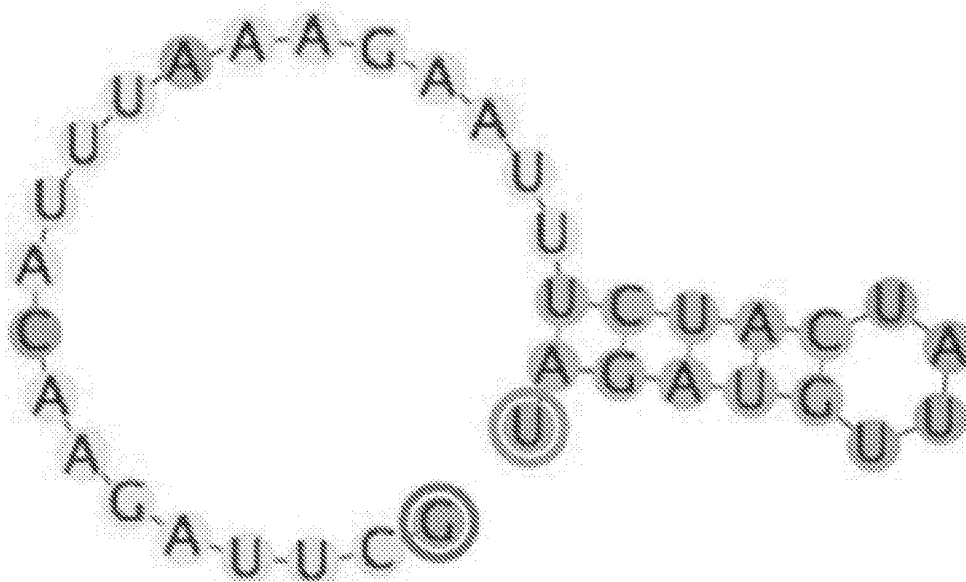


FIG. 80D

>9 DR

GUUUGAAUAACCUUAAAUAUUUCUACUUUGUAGAU



>10 DR

CUCAAAACUCAUUCGAAUCUCUACUCUUUGUAGAU

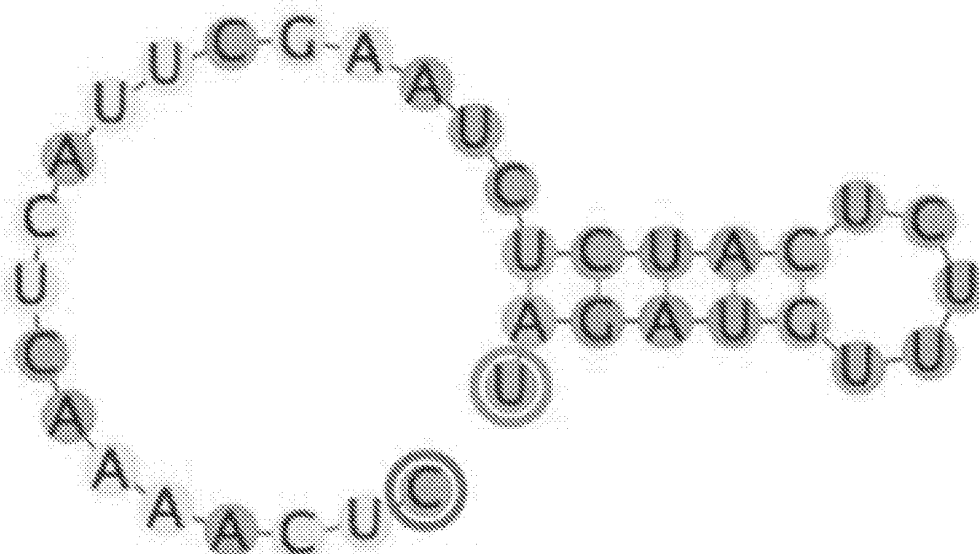
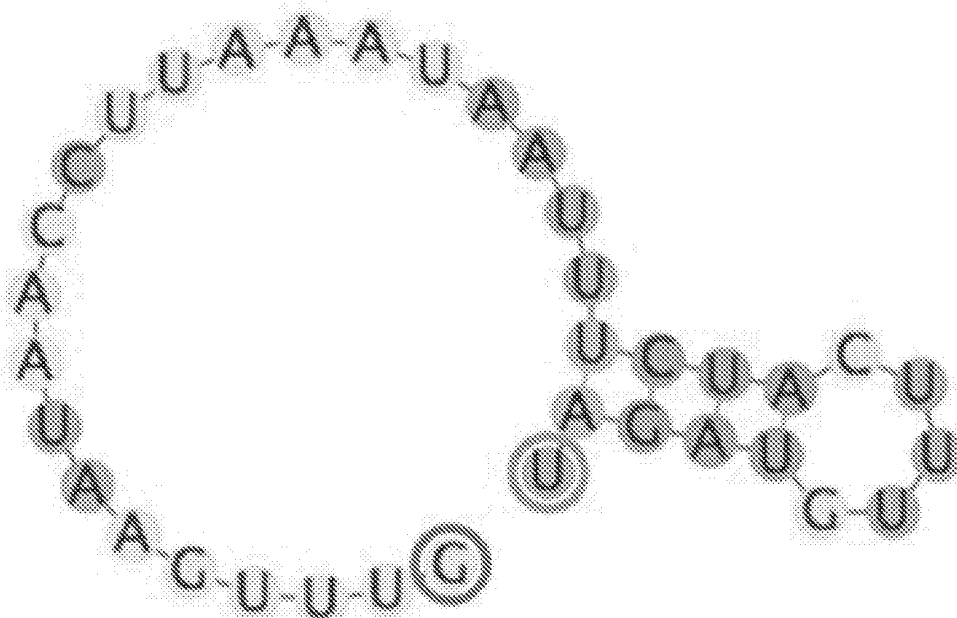


FIG. 80E

>11 DR

GCUUAGAACAUUUAAAAGAAUUUCUACUAAUUGUAGAU



>12 DR

GUCAAAAGACCUUUUUAAUUUCUACUCUUGUAGAU

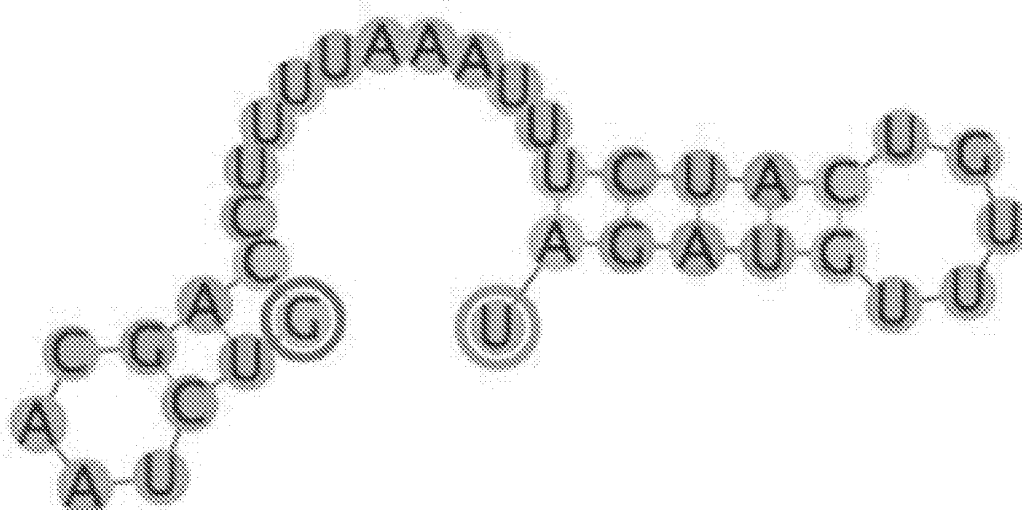
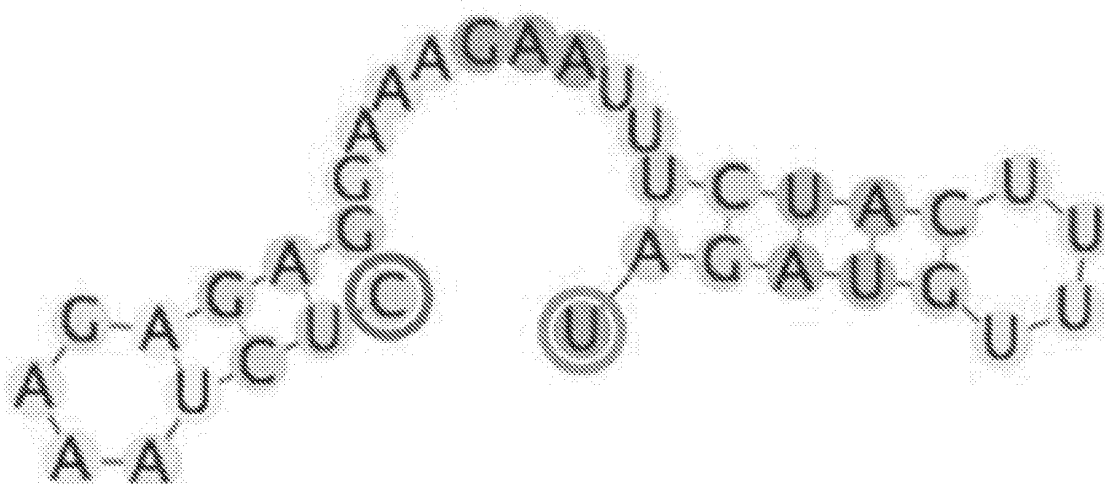


FIG. 80F

>13 DR

GUUUCAAUCCACGCGCCACGCGGGGCGCGAC



>14 DR

GCCAAAUACCUCUAUAAAAUUUCUACUUUUGUAGAU

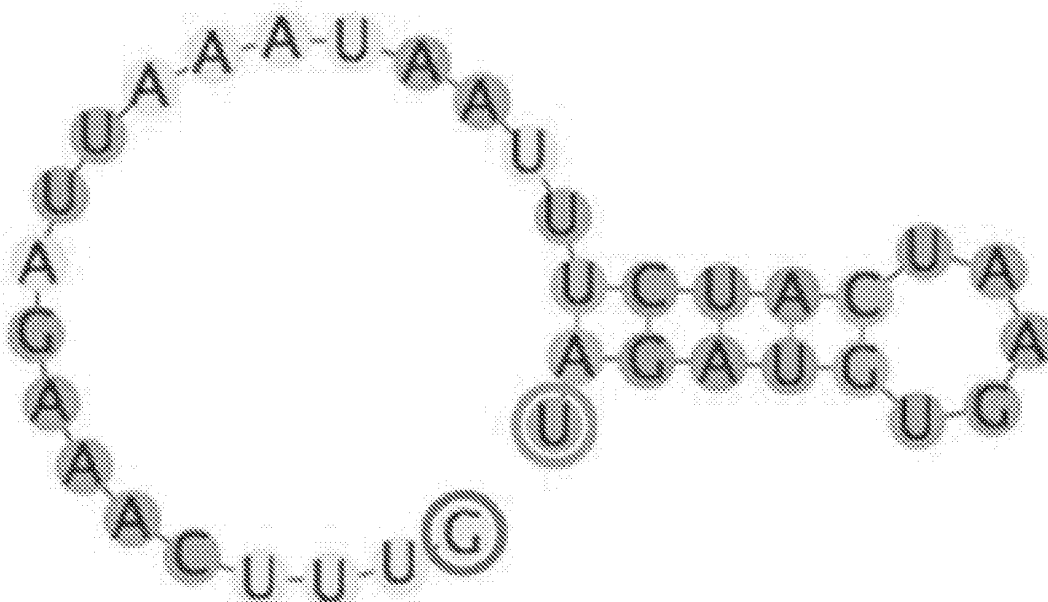
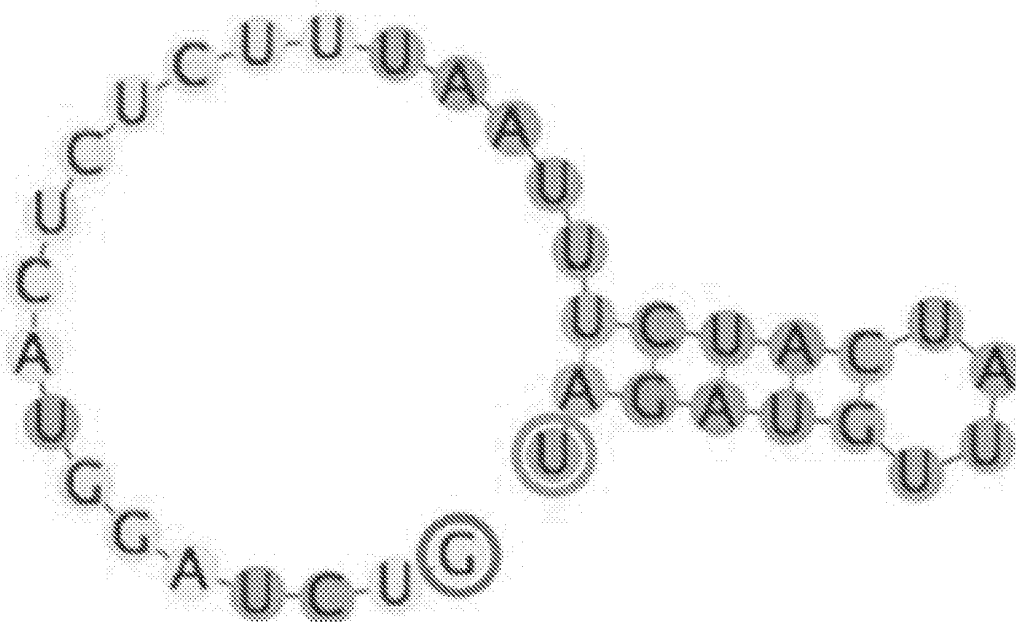


FIG. 80G

>15 DR

GUUUAAAAGUCCUAUUGGAUUUCUACUUUUGUAGAU



>16 DR

GUUUUAGAACCUUAAAAAUUACCUAGUAAUUAGGU

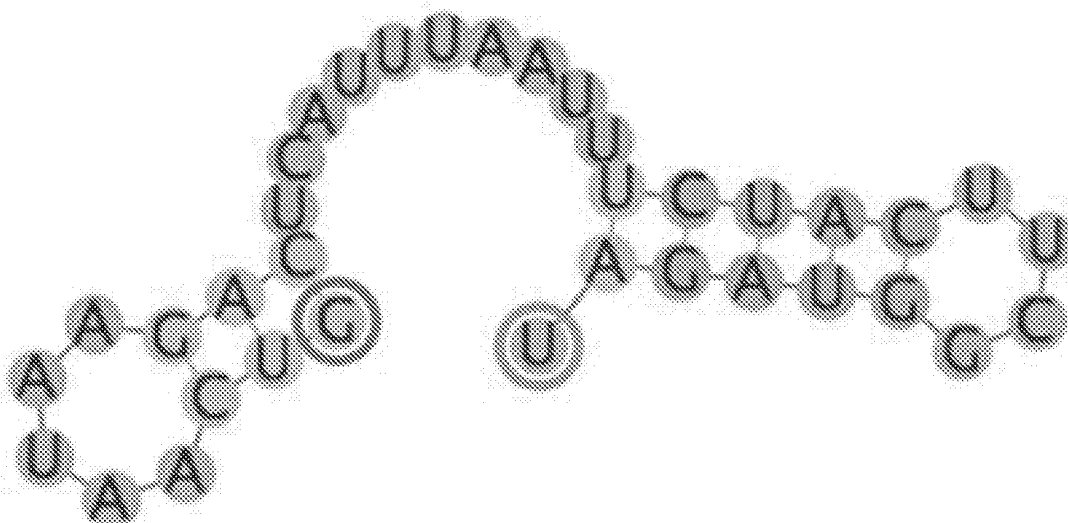


FIG. 80H

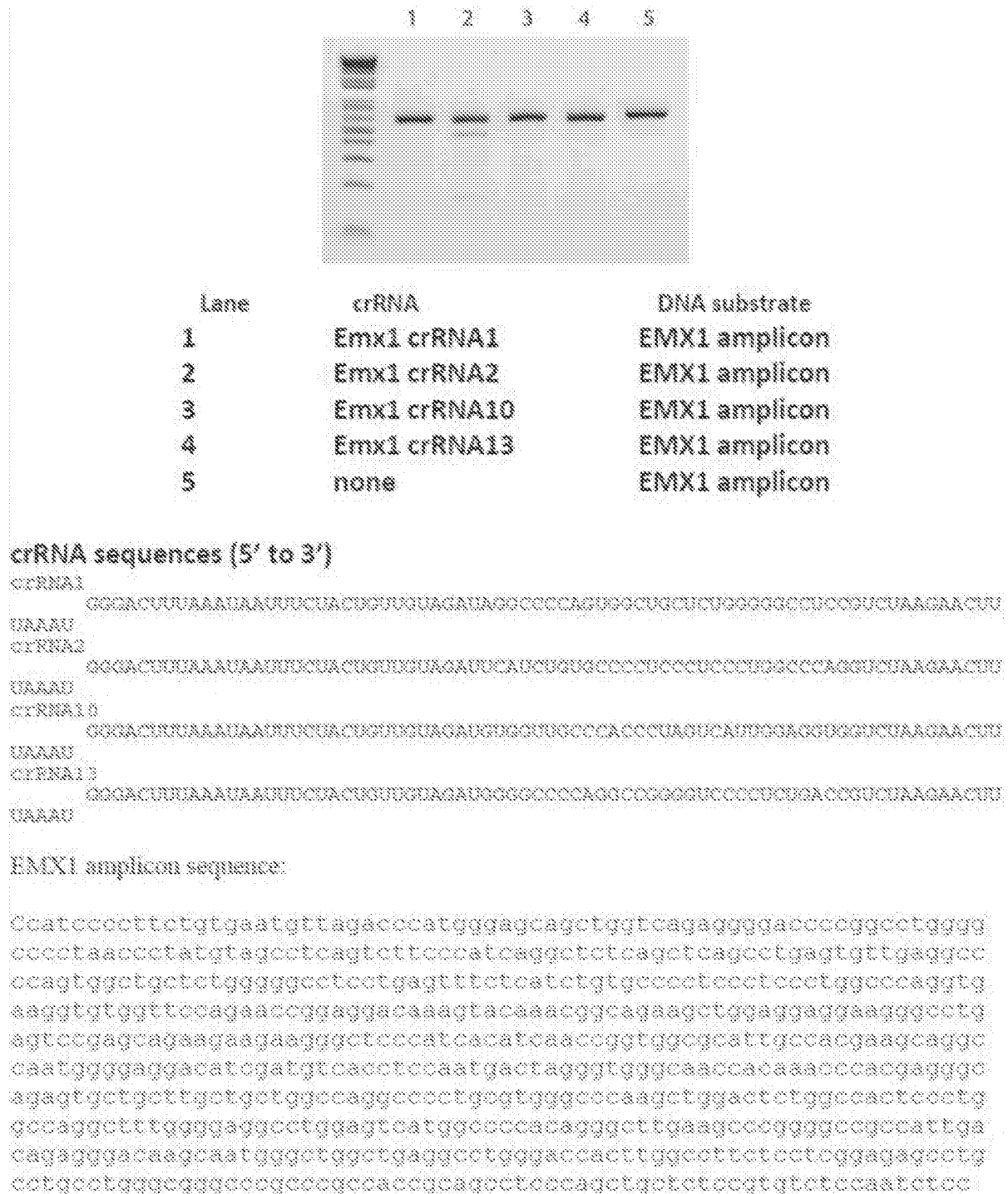
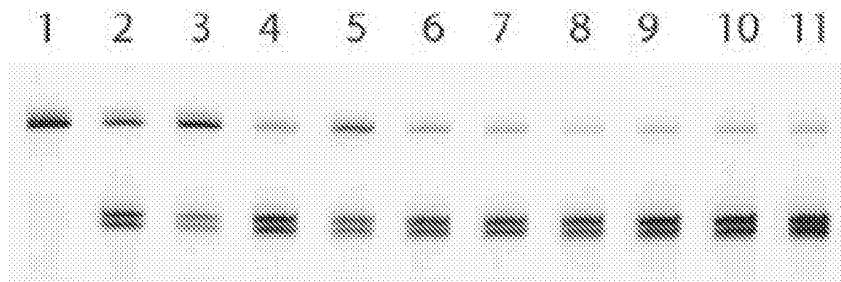


FIG. 81



Each lane consists of Cpf1-containing cell lysate, pUC19 with TTe protospacer, and the corresponding crRNA, indicated as 1-11.

crRNA sequences (5' to 3', mismatch base underlined)

1. GGGACTTTAAATAATTCTACTGTTGTAGATGAGAGTCATTTAATAAGGCCACTGTTAAAAGTCTAAGAACTTTAAAT
2. GGGACTTTAAATAATTCTACTGTTGTAGATGAGAGTCATTTAATAAGGCCACTGTTAAAAGTCTAAGAACTTTAAAT
3. GGGACTTTAAATAATTCTACTGTTGTAGATGAGAGTCATTTAATAAGGCCACTGTTAAAAGTCTAAGAACTTTAAAT
4. GGGACTTTAAATAATTCTACTGTTGTAGATGAGAGTCATTTAATAAGGCCACTGTTAAAAGTCTAAGAACTTTAAAT
5. GGGACTTTAAATAATTCTACTGTTGTAGATGAGAGTCATTTAATAAGGCCACTGTTAAAAGTCTAAGAACTTTAAAT
6. GGGACTTTAAATAATTCTACTGTTGTAGATGAGAGTCATTTAATAAGGCCACTGTTAAAAGTCTAAGAACTTTAAAT
7. GGGACTTTAAATAATTCTACTGTTGTAGATGAGAGTCATTTAATAAGGCCACTGTTAAAAGTCTAAGAACTTTAAAT
8. GGGACTTTAAATAATTCTACTGTTGTAGATGAGAGTCATTTAATAAGGCCACTGTTAAAAGTCTAAGAACTTTAAAT
9. GGGACTTTAAATAATTCTACTGTTGTAGATGAGAGTCATTTAATAAGGCCACTGTTAAAAGTCTAAGAACTTTAAAT
10. GGGACTTTAAATAATTCTACTGTTGTAGATGAGAGTCATTTAATAAGGCCACTGTTAAAAGTCTAAGAACTTTAAAT
11. GGGACTTTAAATAATTCTACTGTTGTAGATGAGAGTCATTTAATAAGGCCACTGTTAAAAGTCTAAGAACTTTAAAT

FIG. 83

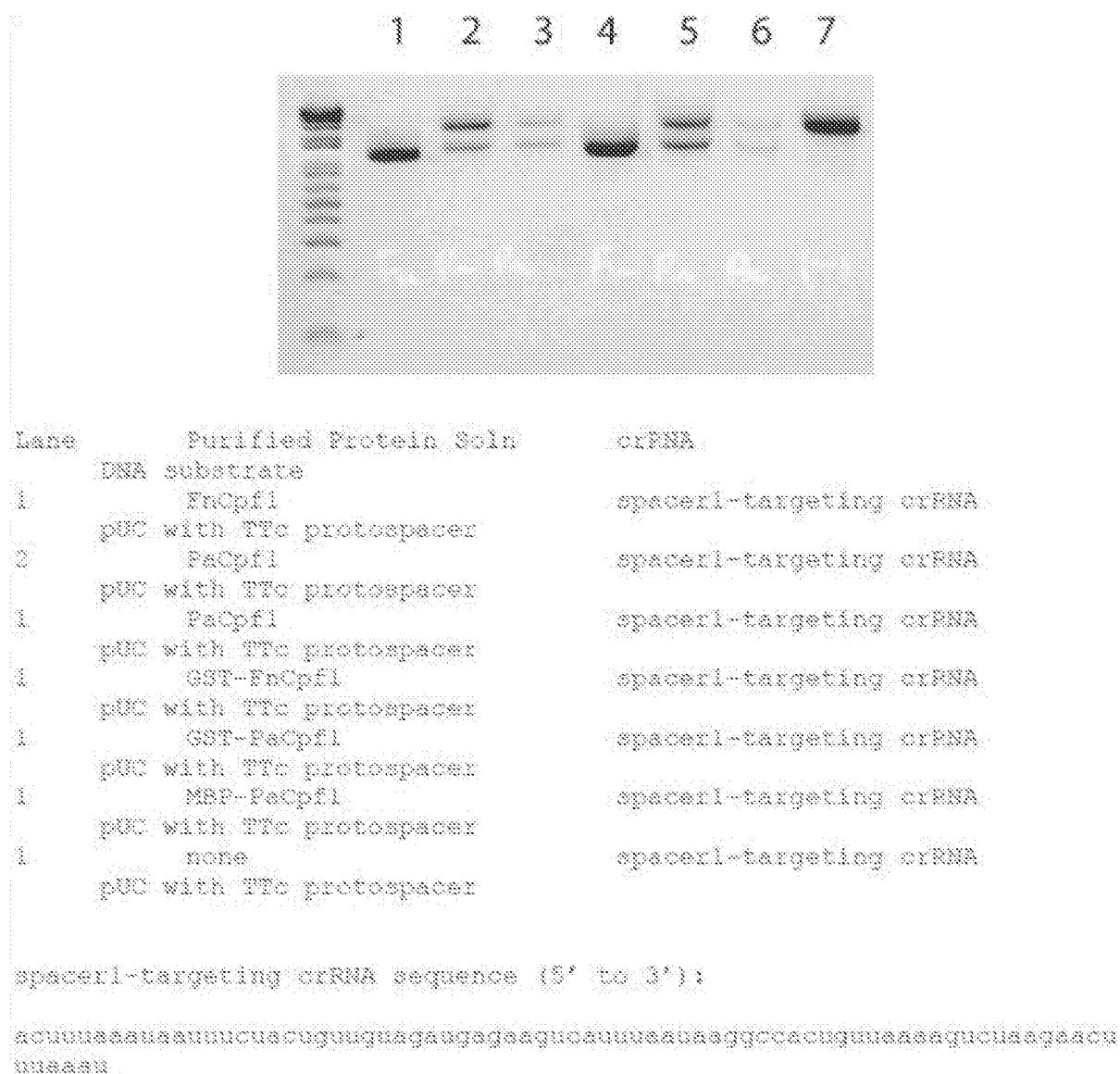


FIG. 84

FnCpf1 DR secondary structure (stem loop highlighted):

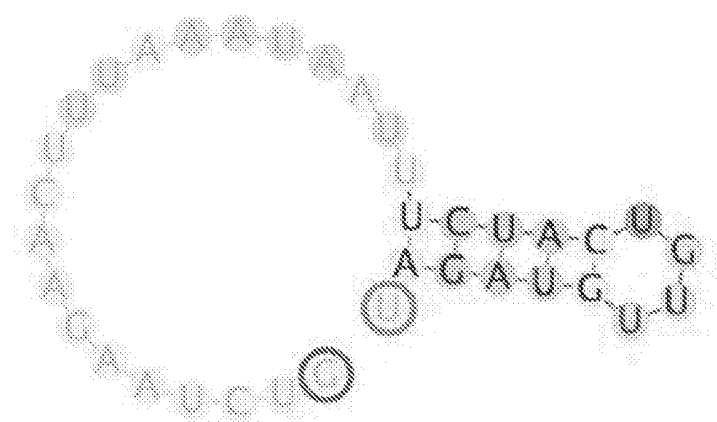


FIG. 85A

PaCpf1 DR secondary structure (stem loop highlighted, identical except for a single base difference in the loop region):

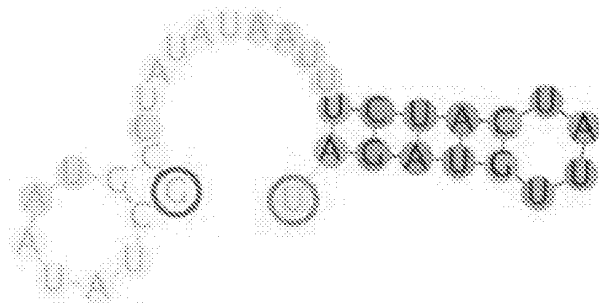


FIG. 85B

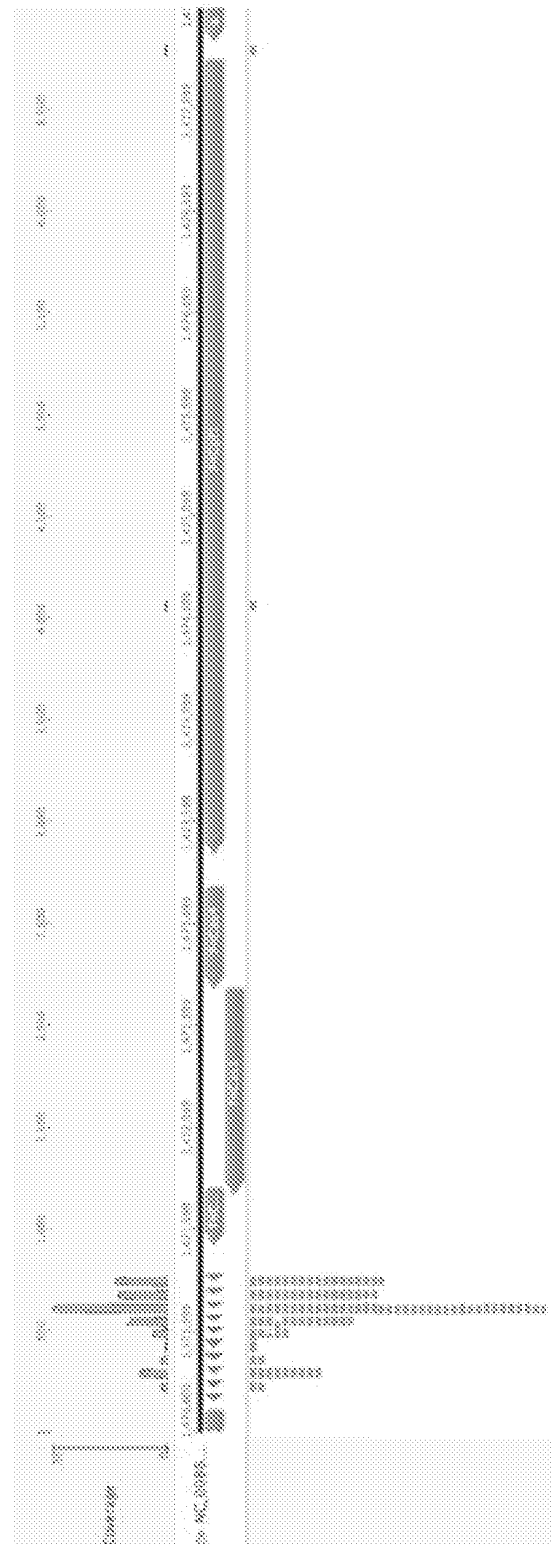


FIG. 86

Mature crRNA sequences for FxCpf1 have the form of:

5'-AAUUUCUACUGUUGUAGAUN₂₀₋₂₄-3'

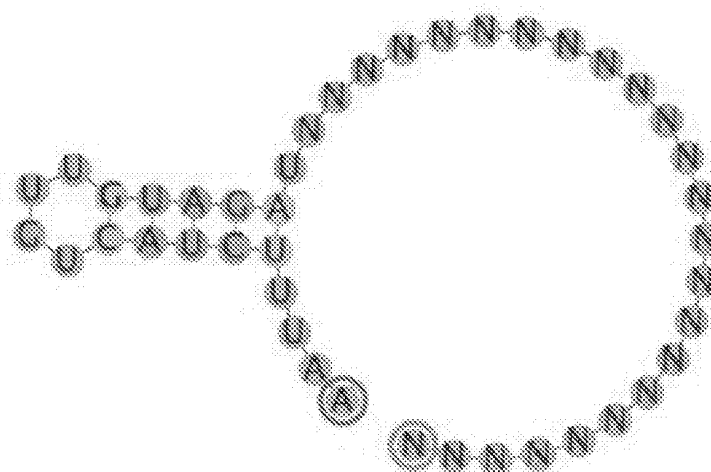


FIG. 87A

Mature crRNA sequences for PaCpf1 have the form of:

5'-AAUUUCUACUUUGUAGAUN₂₀₋₂₄-3'

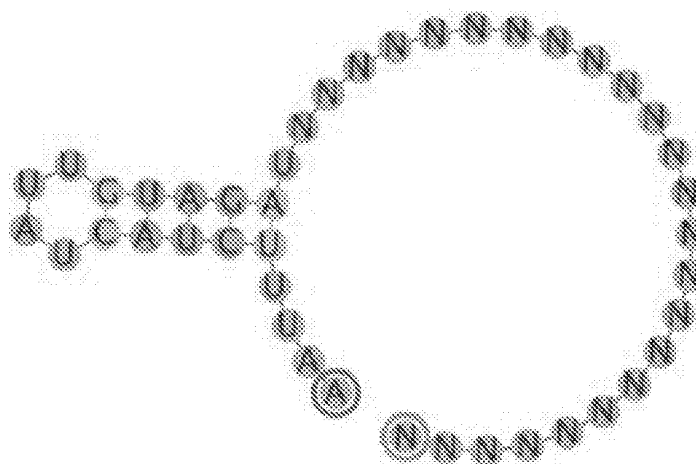


FIG. 87B

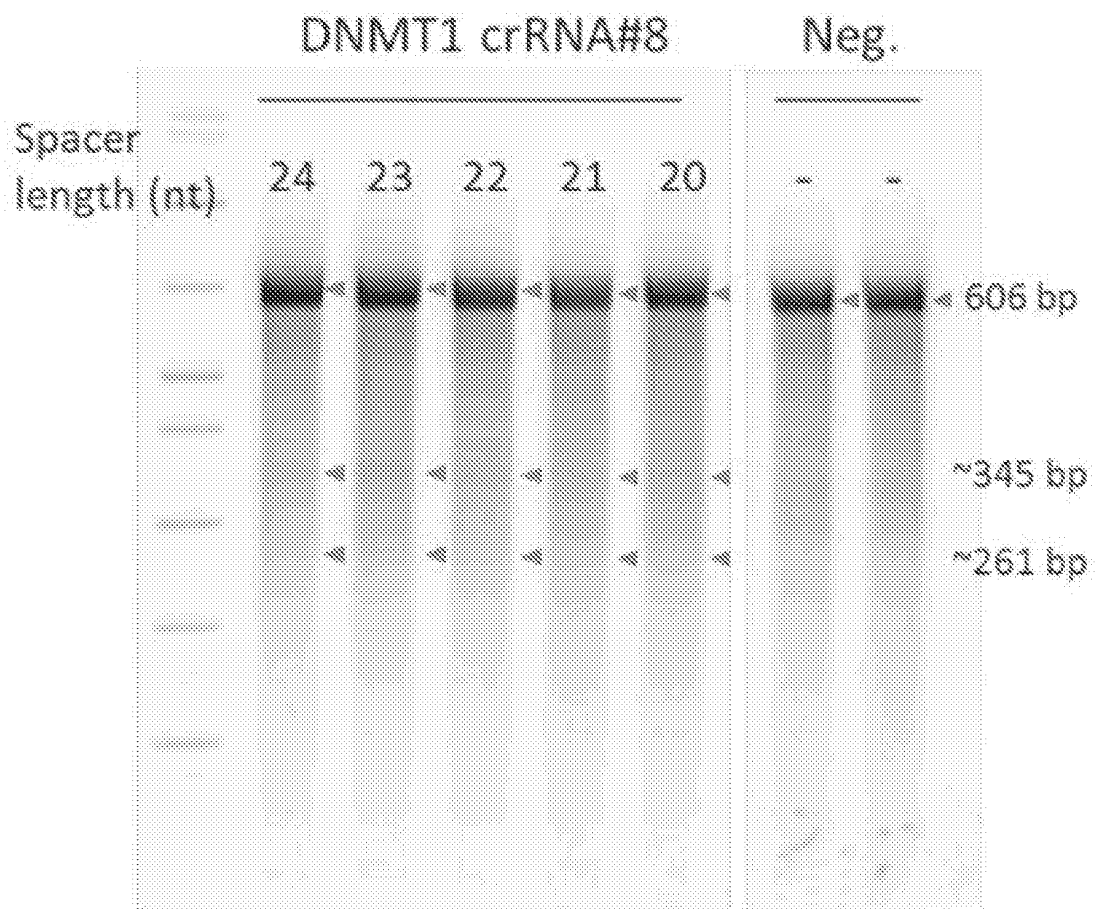


FIG. 88

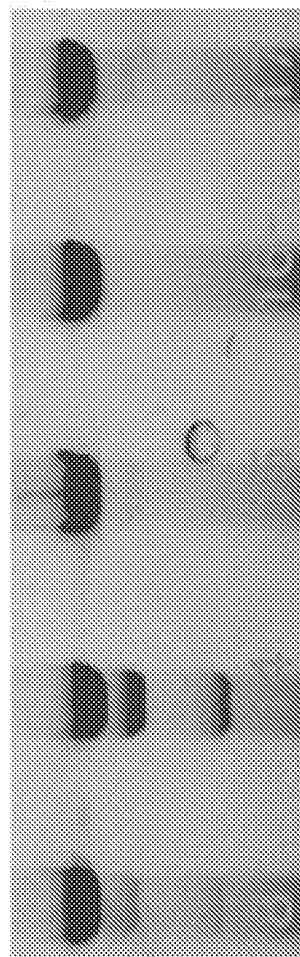
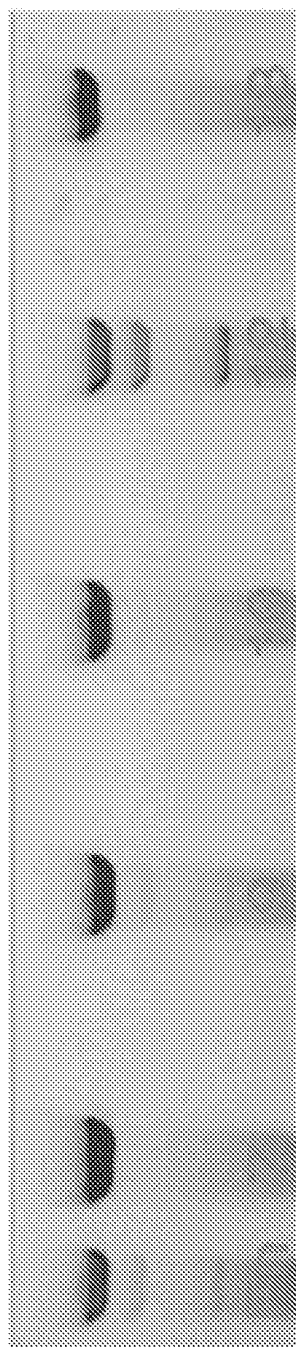
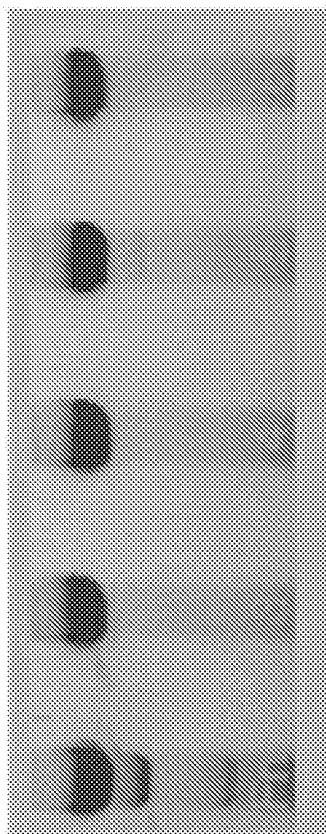


FIG. 89

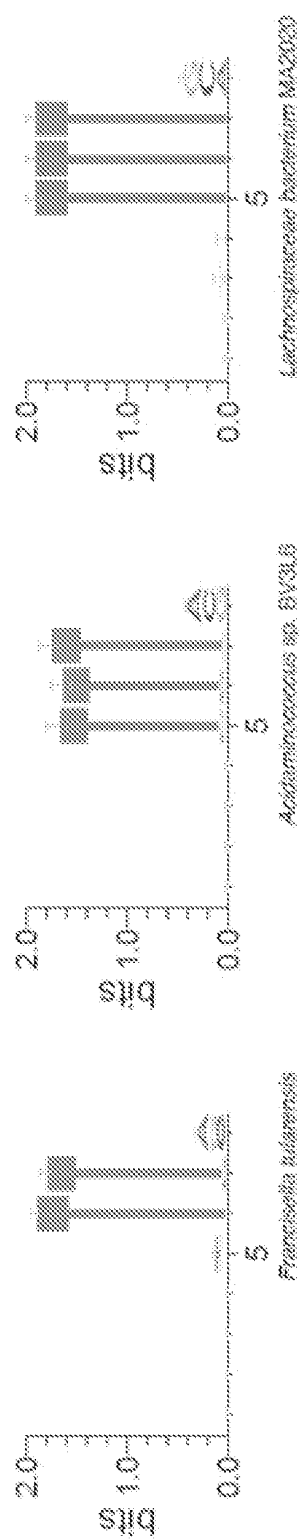


FIG. 90A

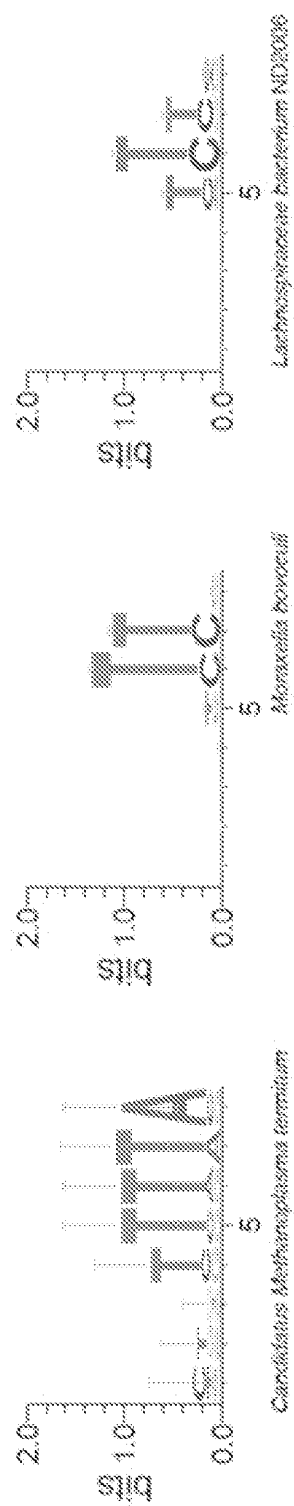


FIG. 90B

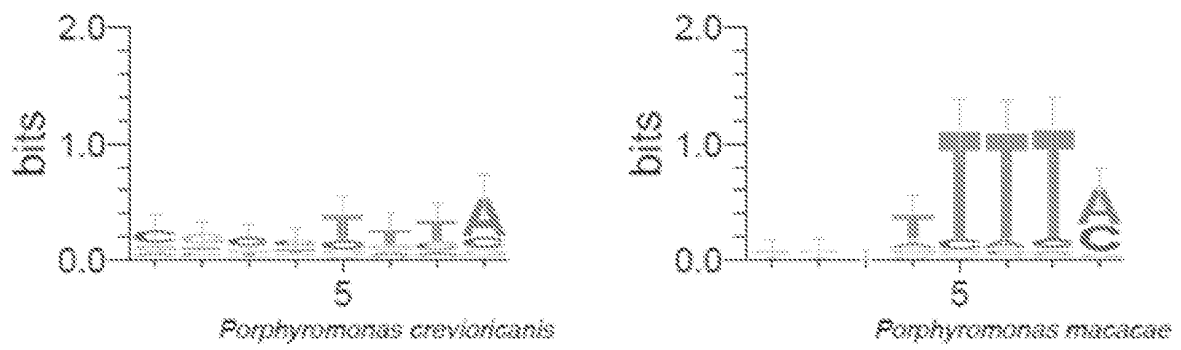
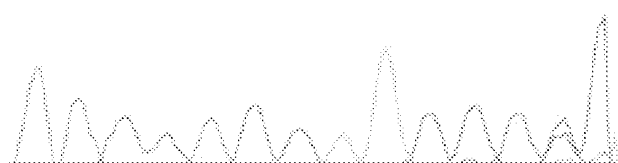


FIG. 90C

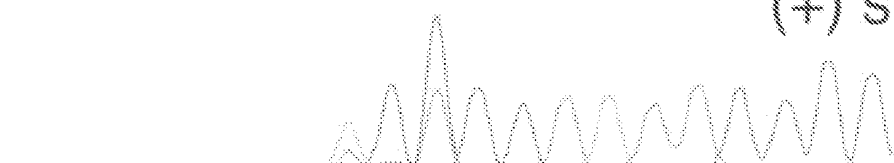
protospacer (1)

. . TTAATAAGGCCACTGTTAAAA . . – 3'



(–) strand

. . TTAATAAGGCCACN



(+) strand

GCCACTGTTAAAA . .

FIG. 91

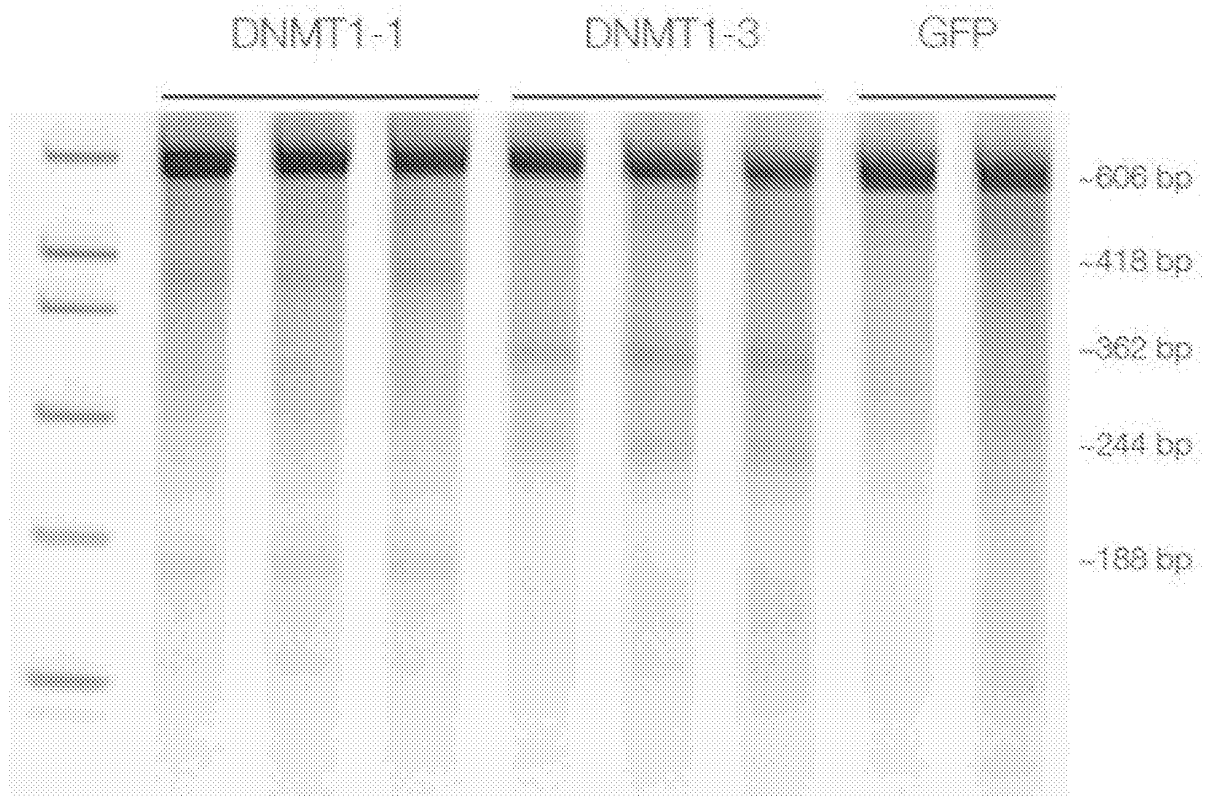


FIG. 93

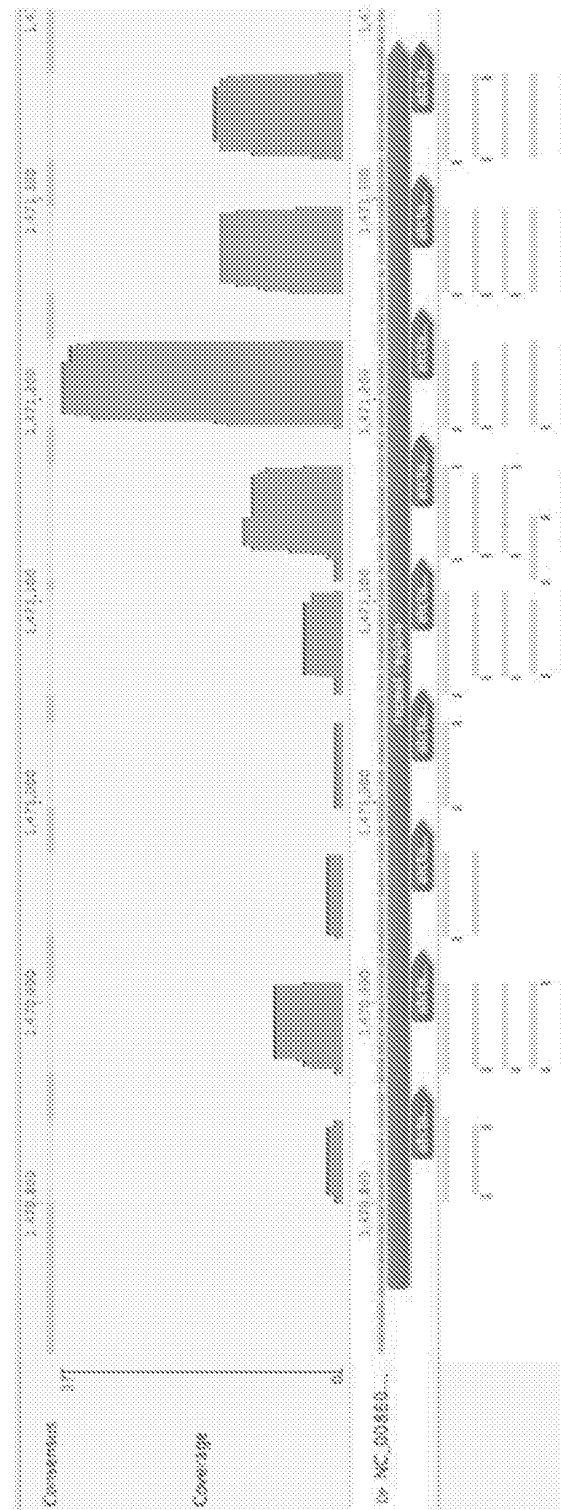


FIG. 94A

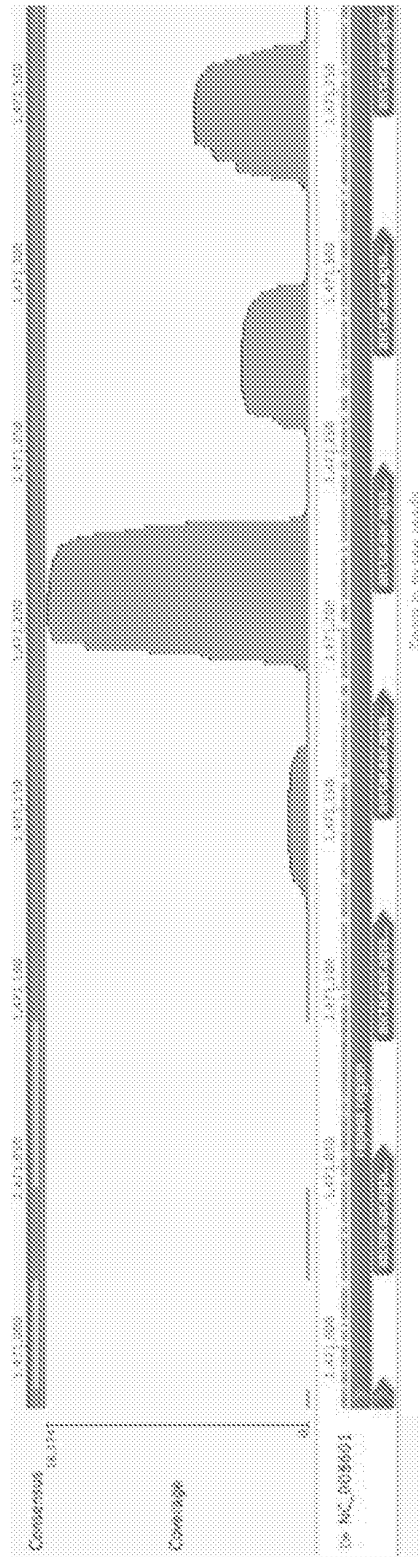


FIG. 94C

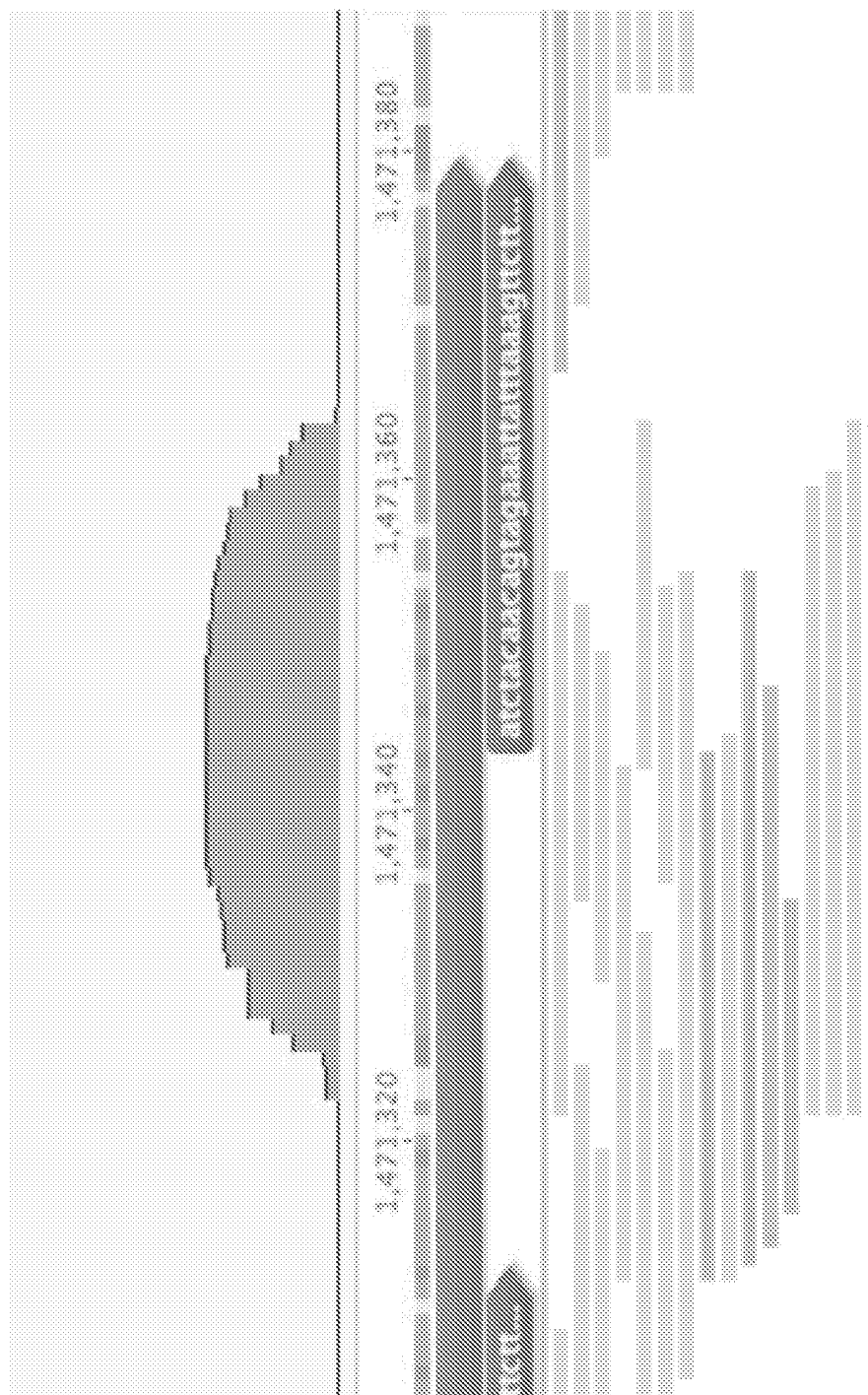


FIG. 94F

Francisella tularensis subsp. *novicida* U112

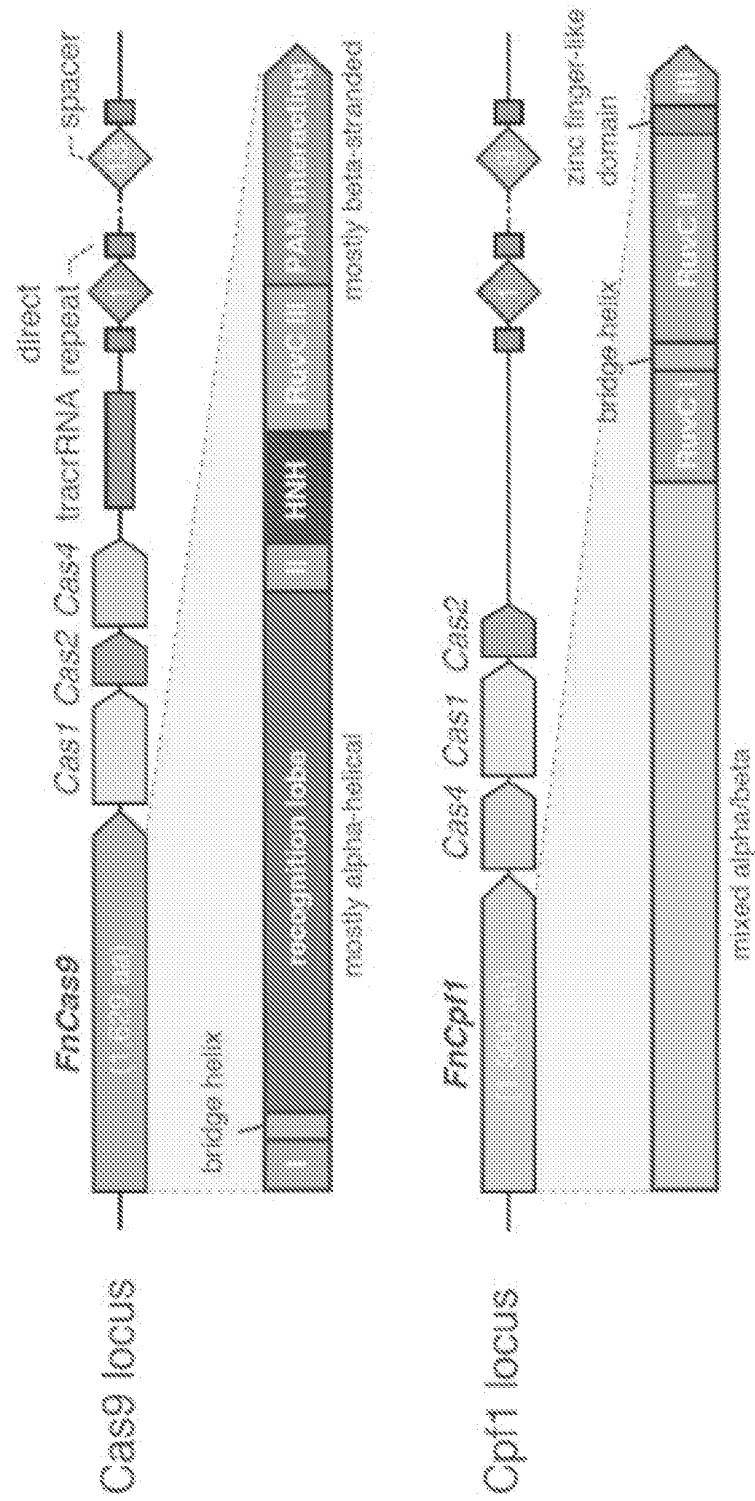


FIG. 95A

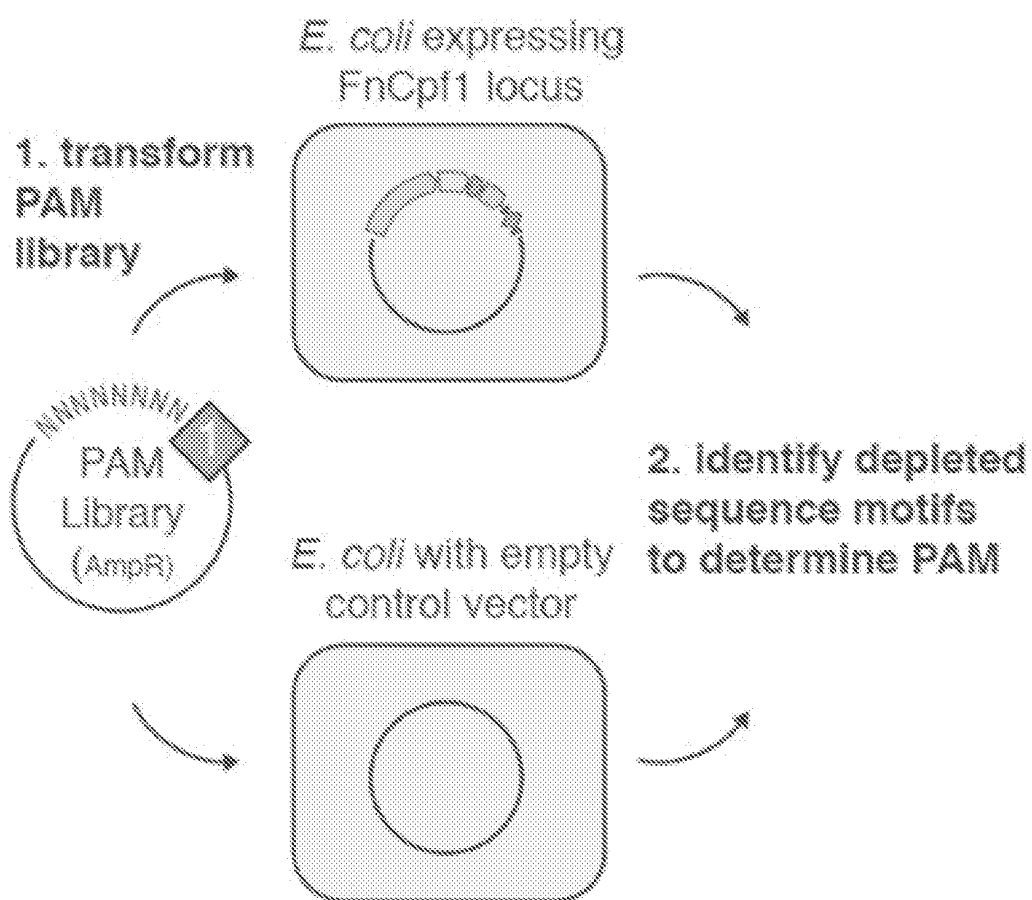


FIG. 95B

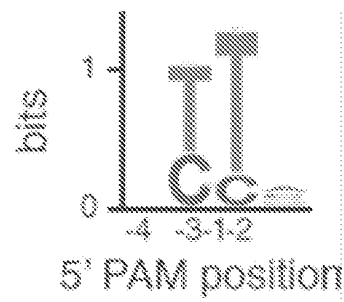


FIG. 95C

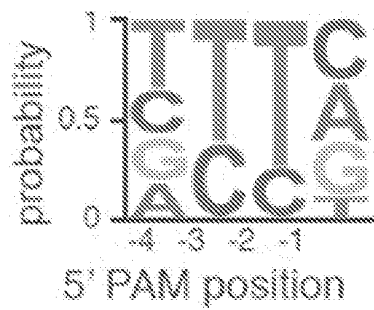


FIG. 95D

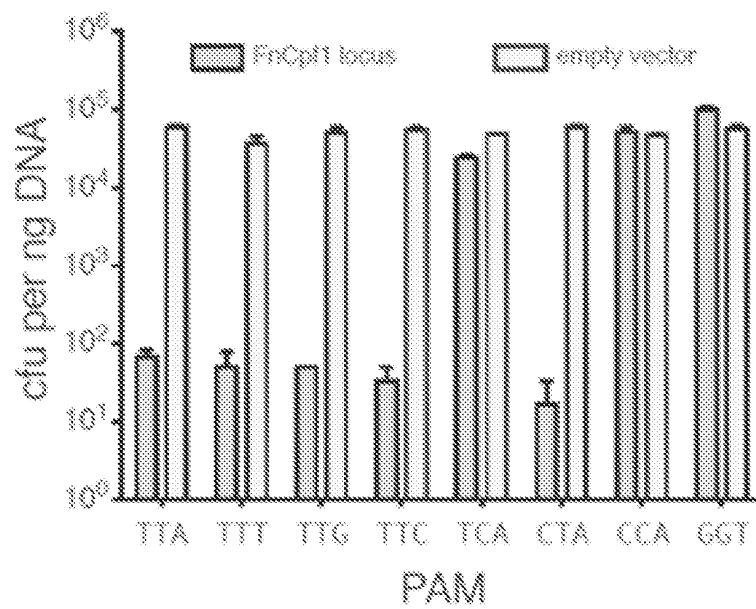


FIG. 95E

Francisella tularensis subsp. *novicida* U112

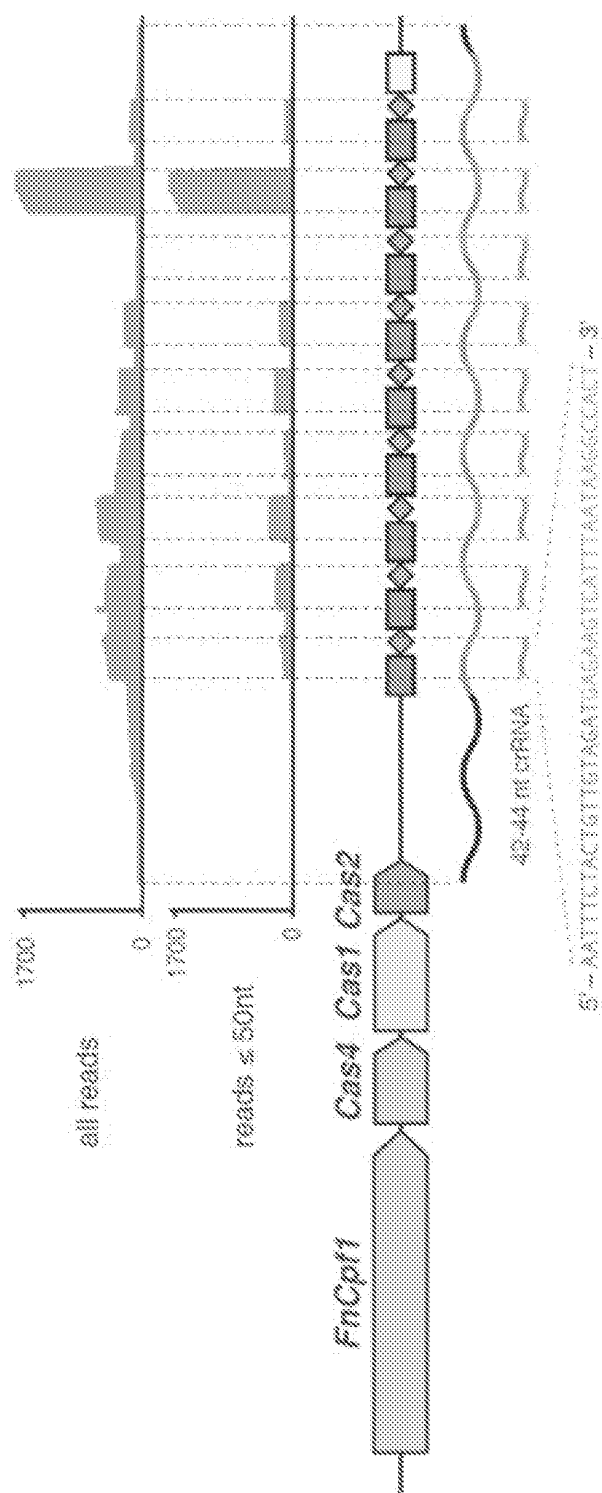


FIG. 96A

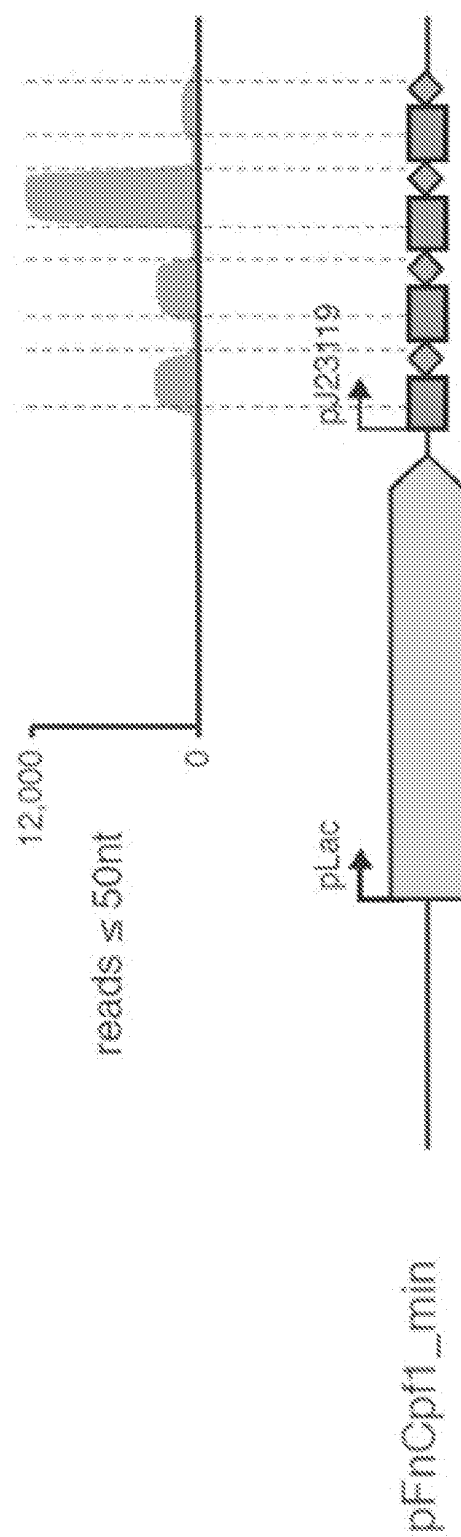


FIG. 96B

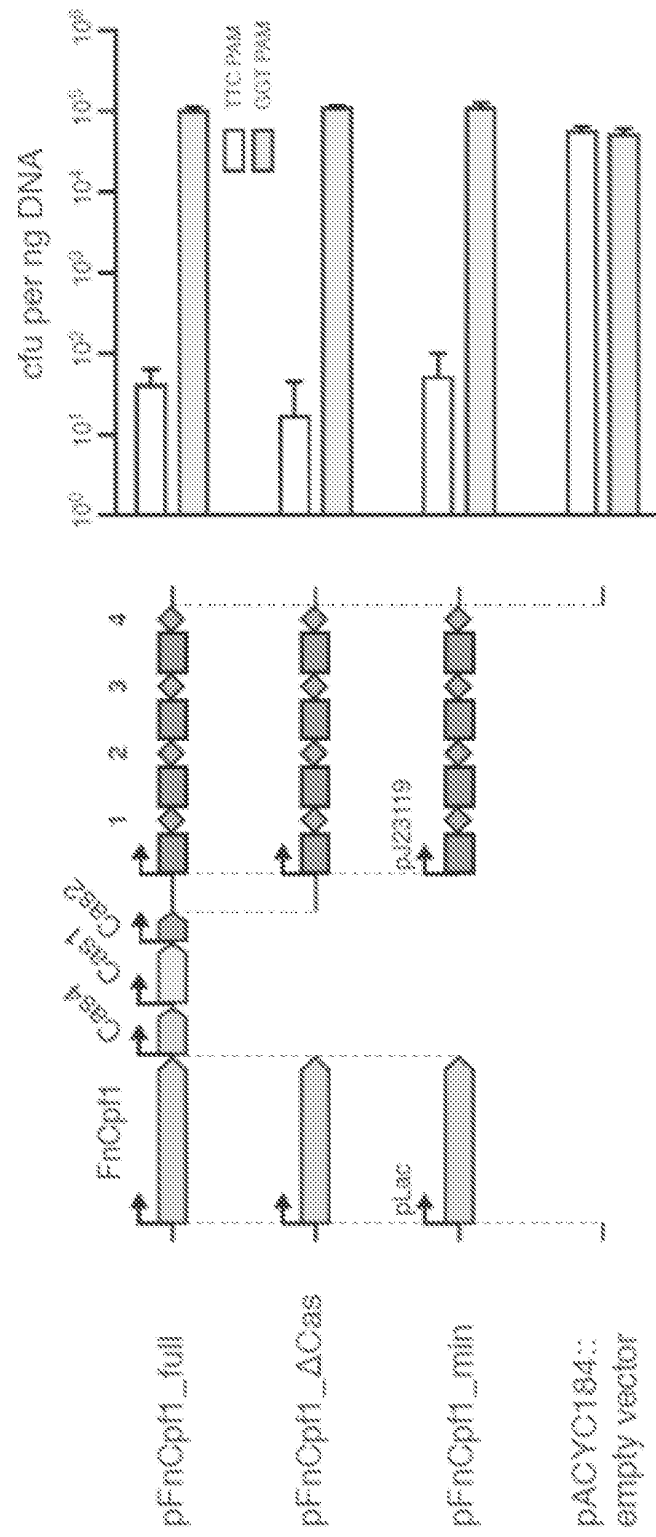


FIG. 96C

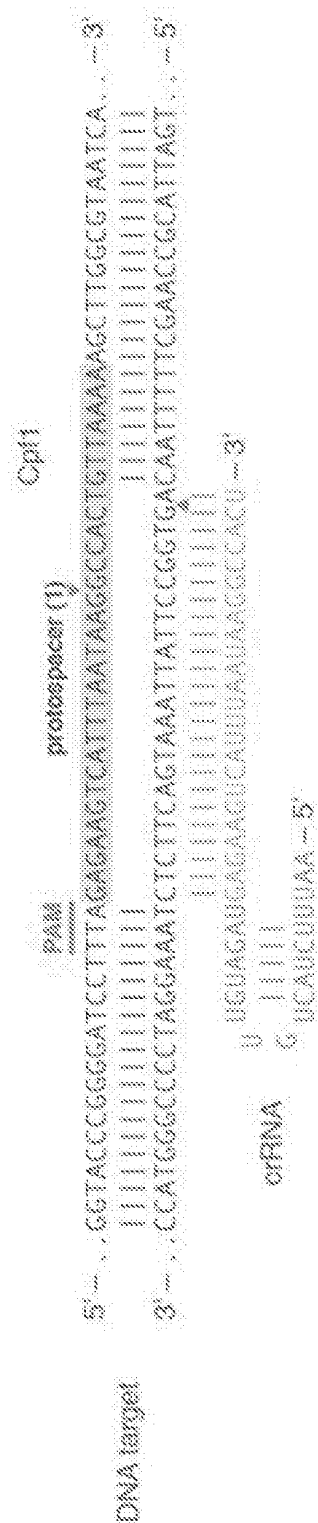


FIG. 97A

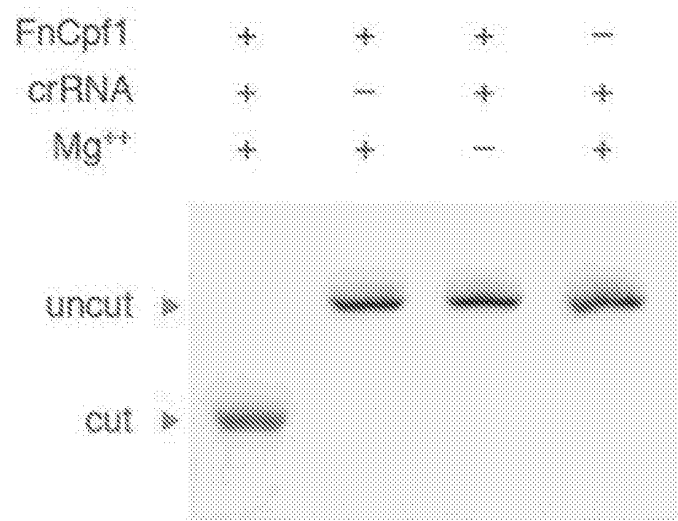


FIG. 97B

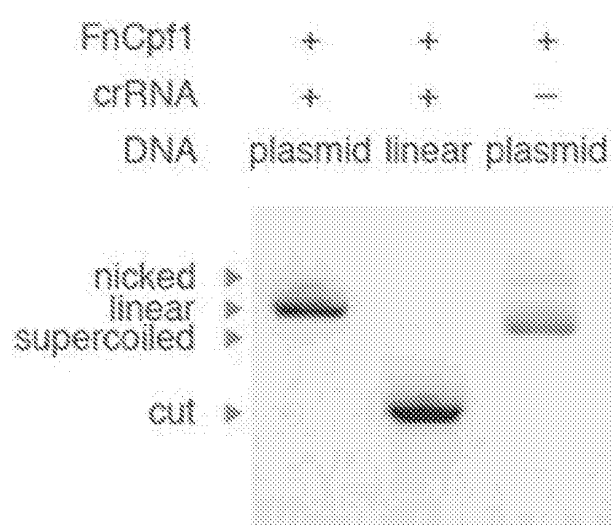


FIG. 97C

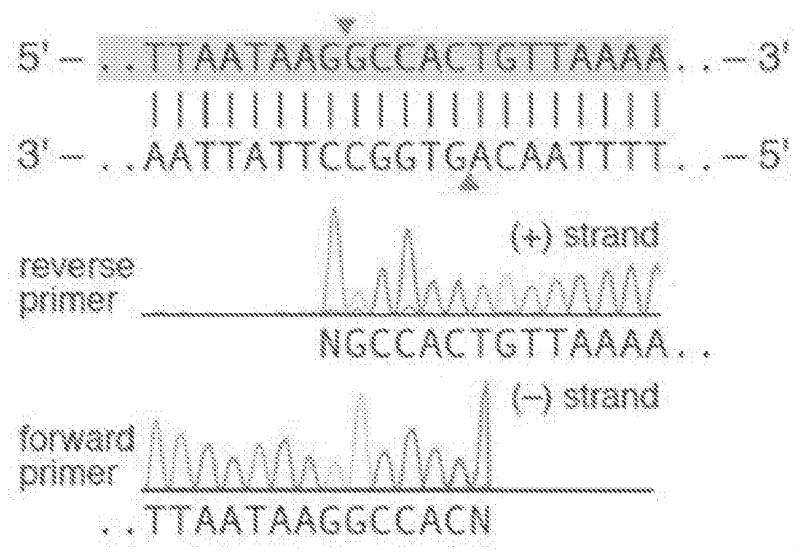


FIG. 97D

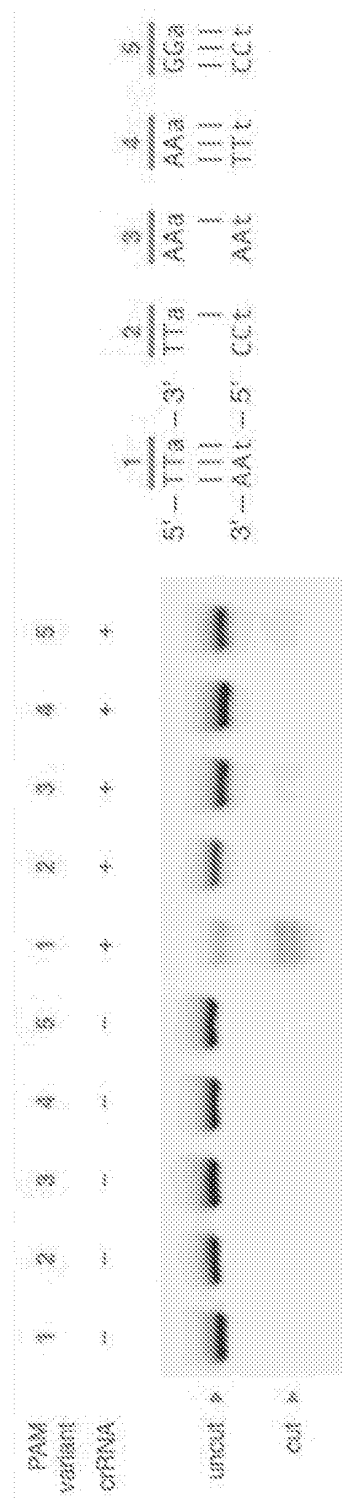


FIG. 97E

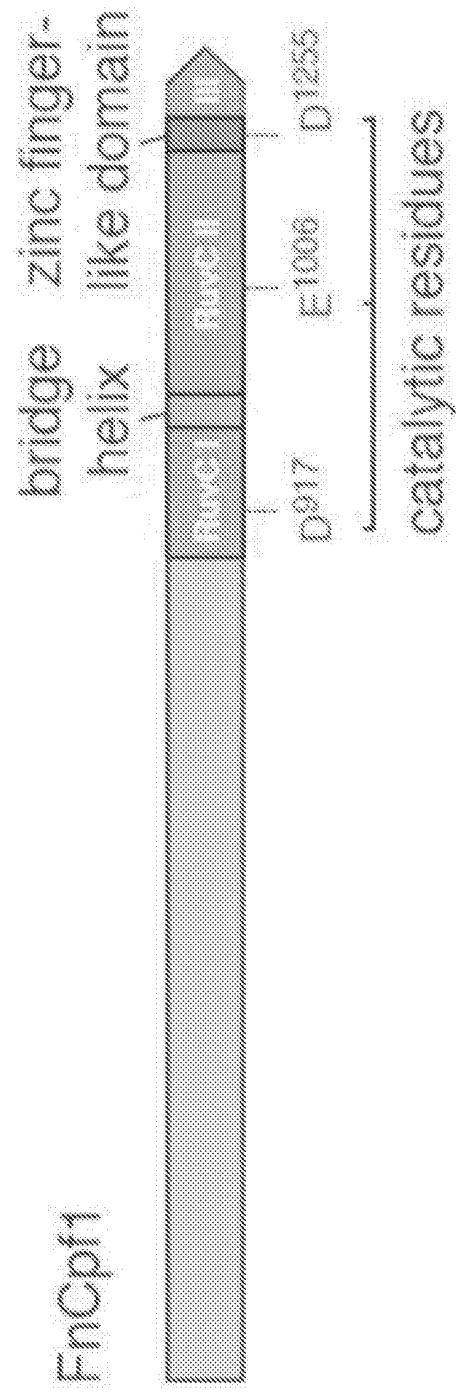


FIG. 98A

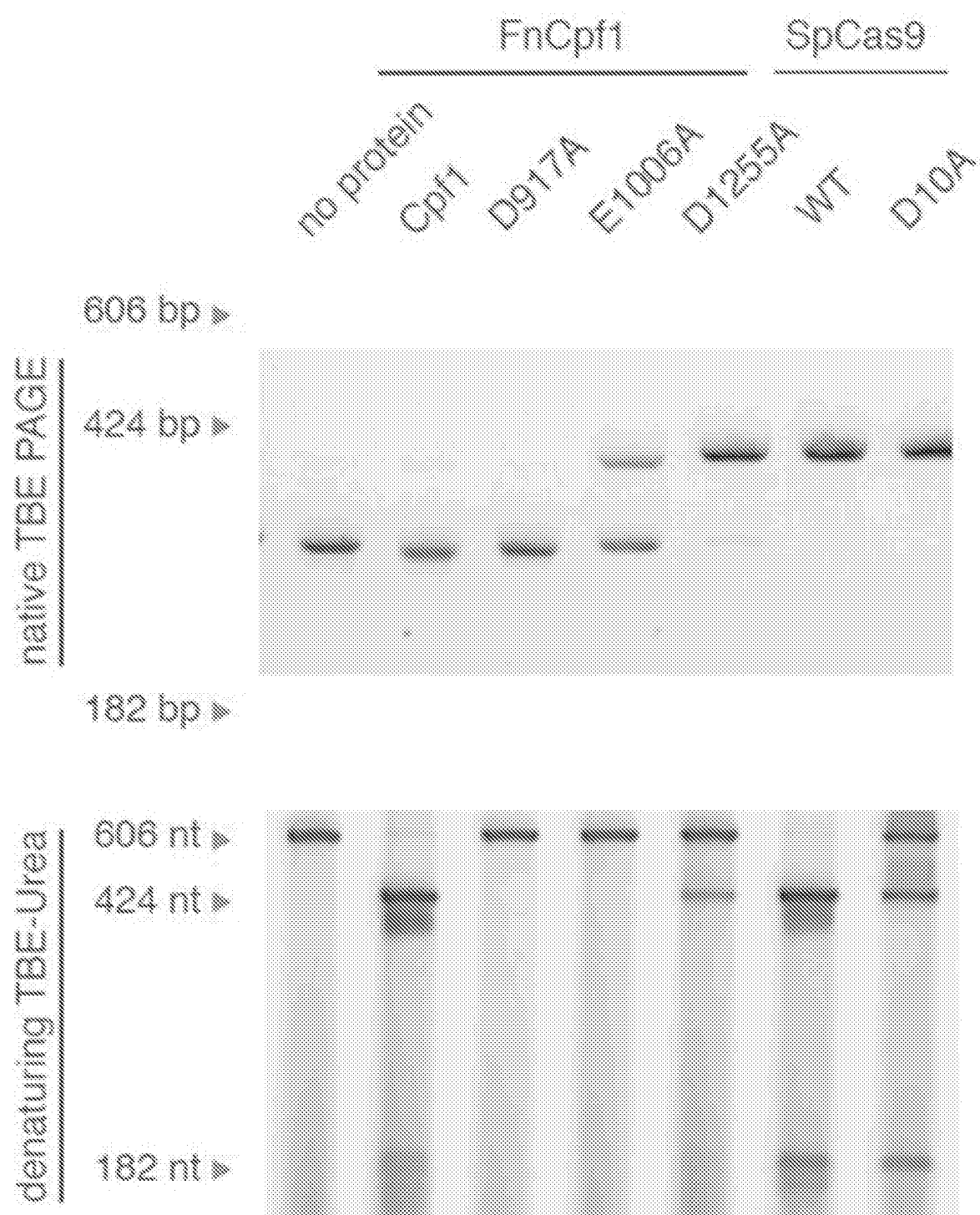


FIG. 98B

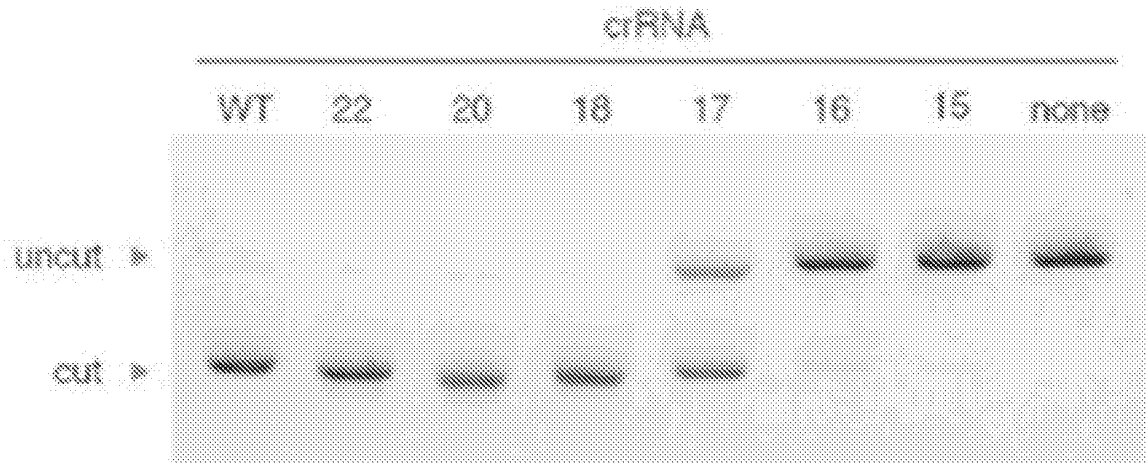
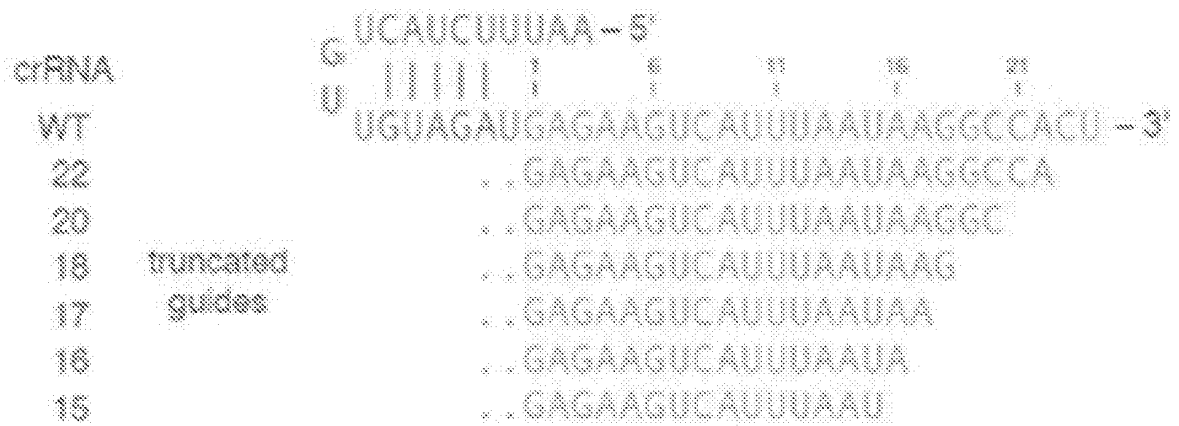


FIG. 99A

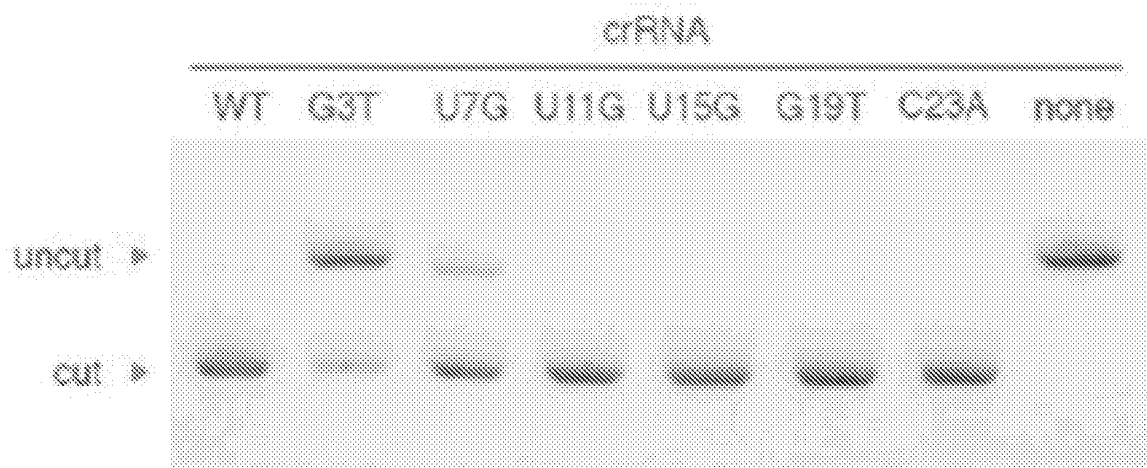


FIG. 99B

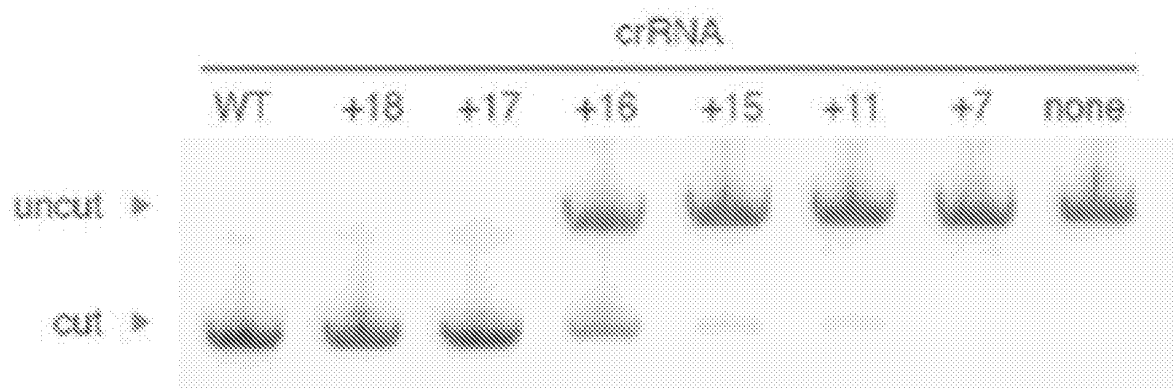
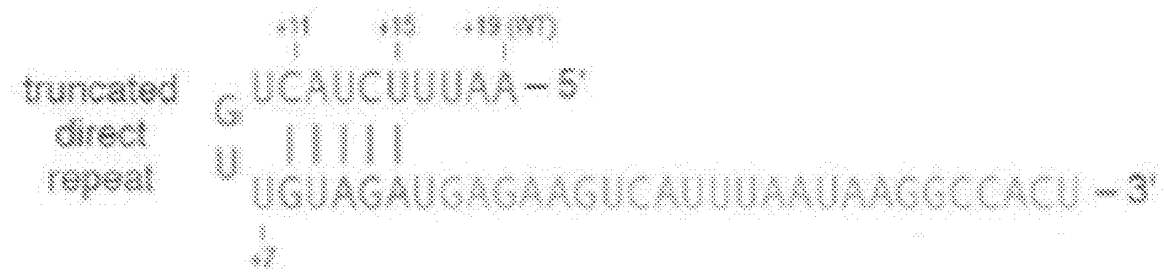


FIG. 99C

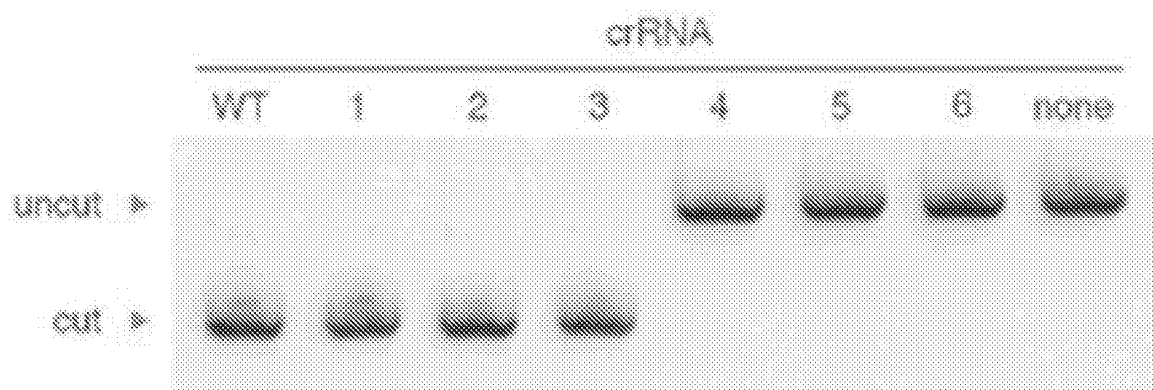
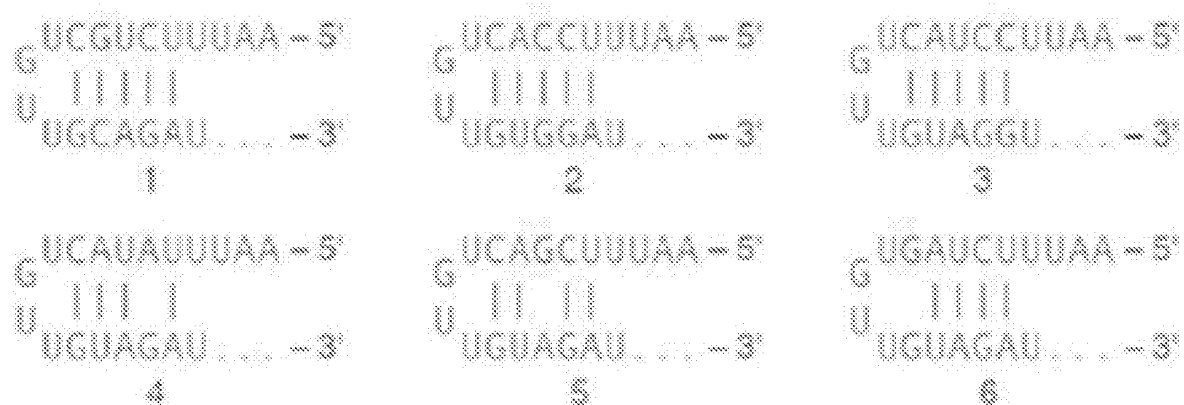


FIG. 99D

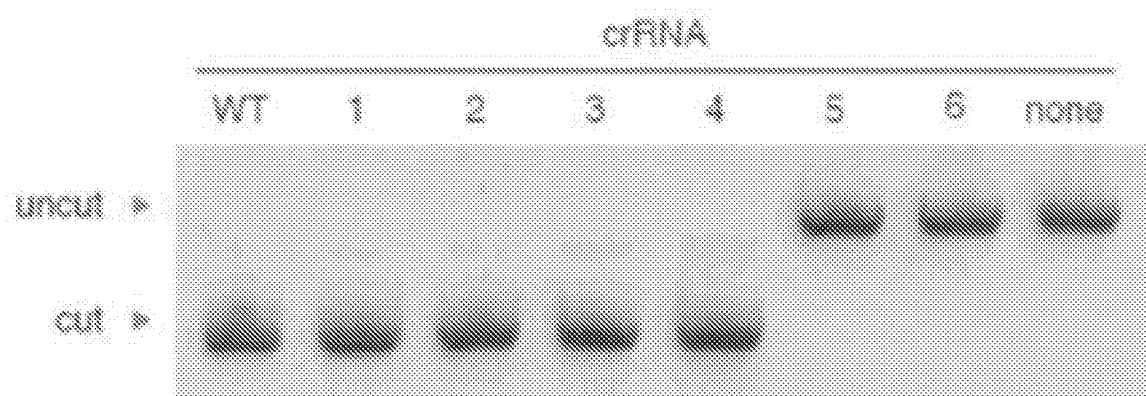
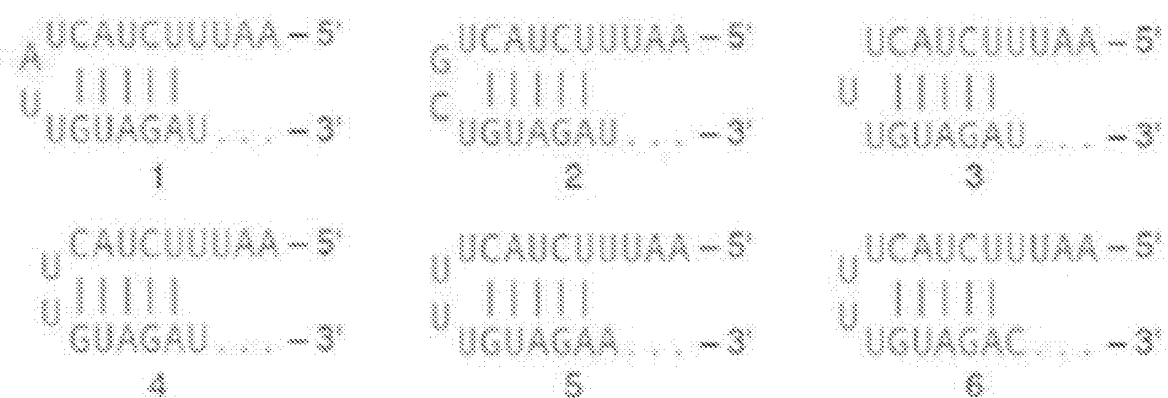
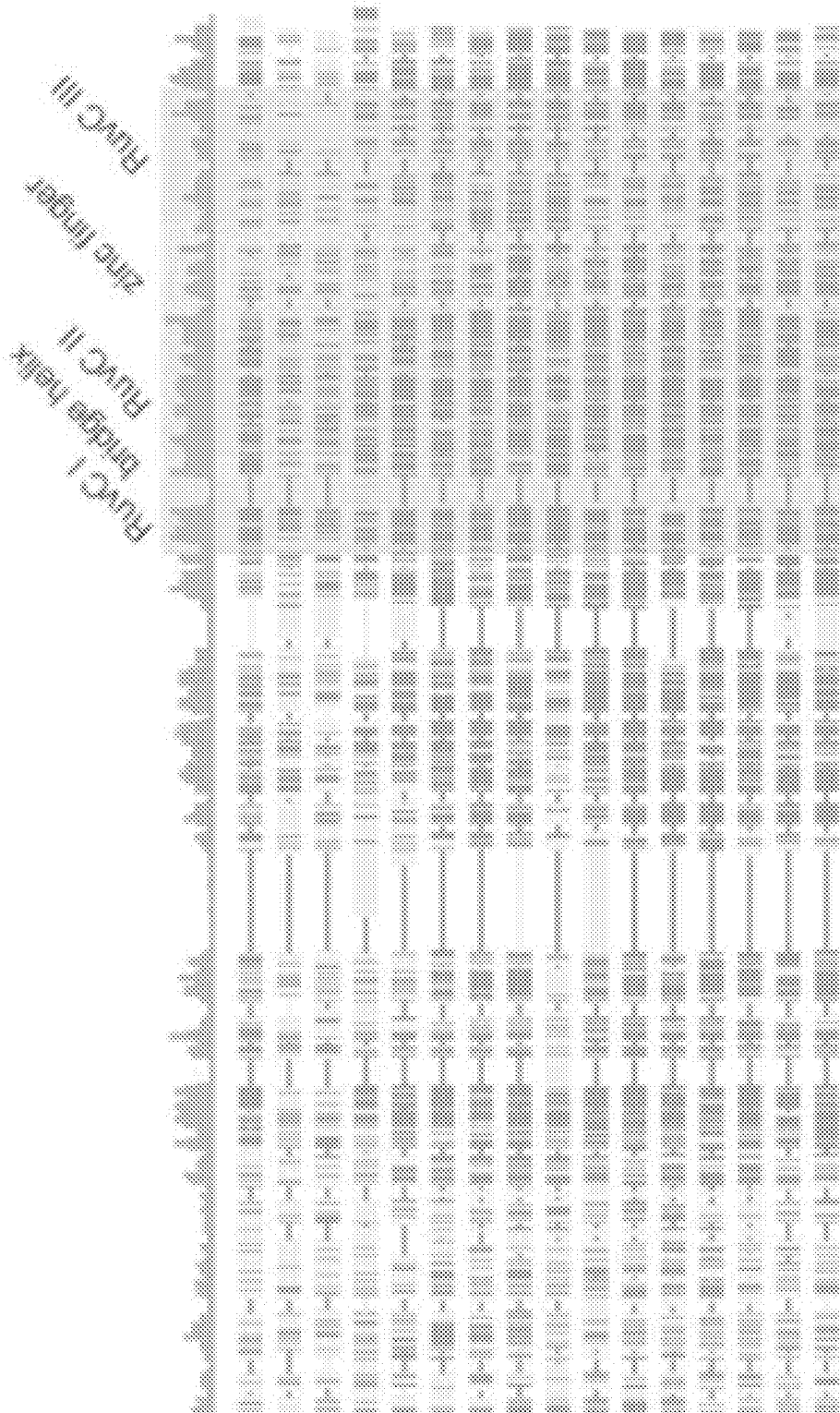


FIG. 99E



LEGO®

		direct repeat sequences of Cpf1 orthologs		
ID				
1	FrnCpf1	S' —	GUUUAAGAAAUUUUAAUUAUUUUUCUACU - GUUGUAGAU	— 3'
2	Lb3Cpf1		GUUUUUGGAGUACCUUAGAAAUGCAUGGUUCUCAUGC	
3	BpCpf1		GUUUUAGAACCUIUAAAAUUACCUAGUAAUUAGGU	
4	PeCpf1		GUUUAAAAGUCCUAUUUGGAUUUCUACU - UUUGUAGAU	
5	PbCpf1		GCCAAAUACCUCUAUAAAAUUUCUACU - UUUGUAGAU	
6	SsCpf1		GUUUCAAUCCACGCGCCACGCGGGGCGCGAC	
7	AsCpf1		GUCAAAAGACCUUUUUAUUUUUCUACU - CUUGUAGAU	
8	Lb2Cpf1		GCUUAGAACAUUUAAAAGAAUUUCUACU - AUUGUAGAU	
9	CMCpf1		GUCAAAACUCAUUCGAAUCUCUACUCUUUGUAGAU	
10	EeCpf1		GUUUGAAUAACCUUAAAUAUUUUUCUACU - - UUUGUAGAU	
11	MbCpf1		GUCUAACGACCUUUUAAAAUUUCUACUGUUUGUAGAU	
12	LiCpf1		CUCUAAAGAGAGGAAAGAAUUUCUACU - UUUGUAGAU	
13	LbCpf1		GUUUCAAAGAUUAAAUAUUUUUCUACUAAGUGUAGAU	
14	PoCpf1		GUCUAGGUACUCUCUUUAUUUUUCUACU - AUUGUAGAU	
15	PdCpf1		GUCAAUAAAGACUCAUUUAUUUUUCUACU - UCGGUAGAU	
16	PsiCpf1		GCCUAUAAGGCUUUACUUAUUUUUCUACU - AUUGUAGAU	
		removed post crRNA maturation	stem left	stem right

FIG. 100C

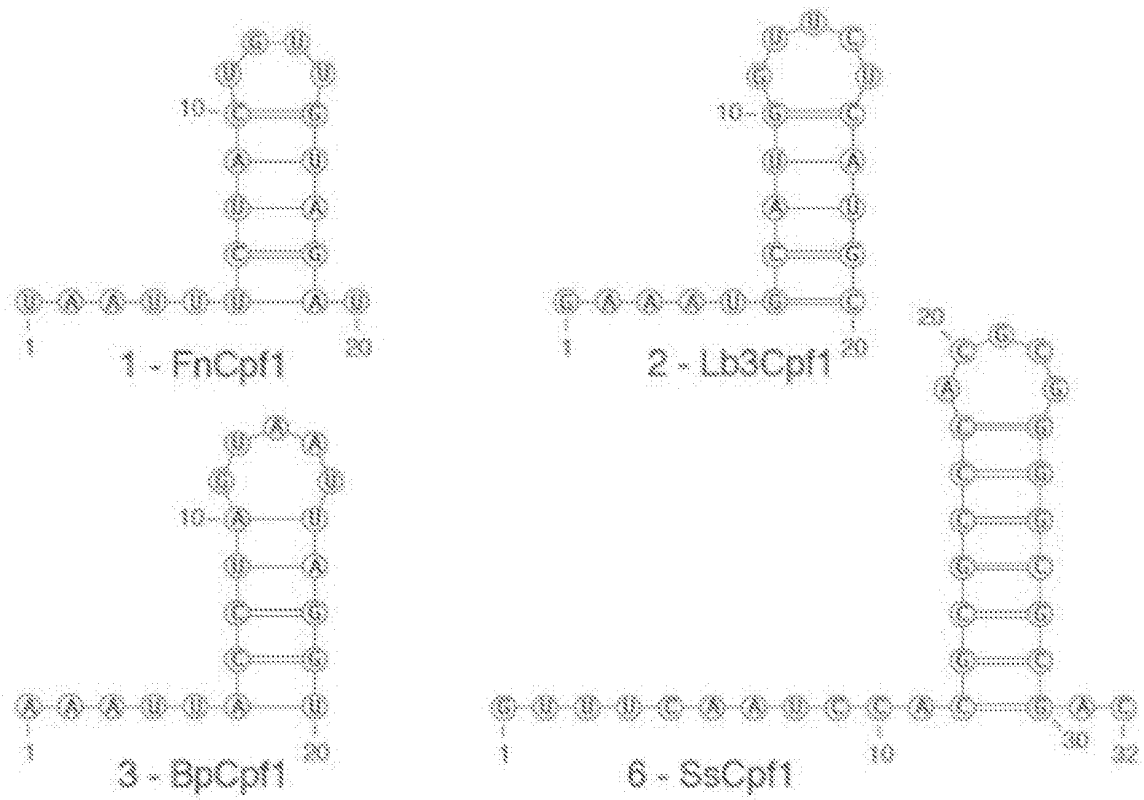


FIG. 100D

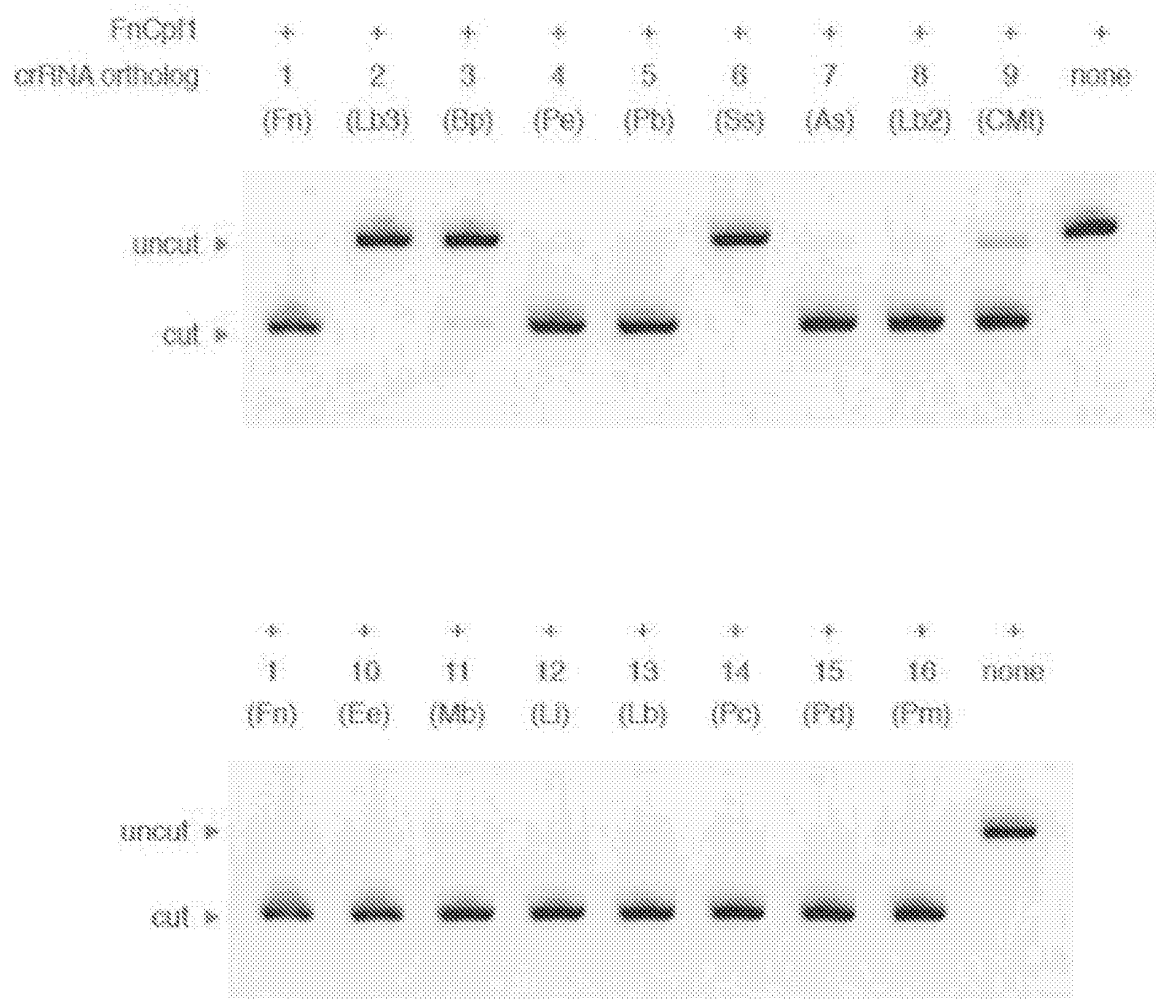


FIG. 100E

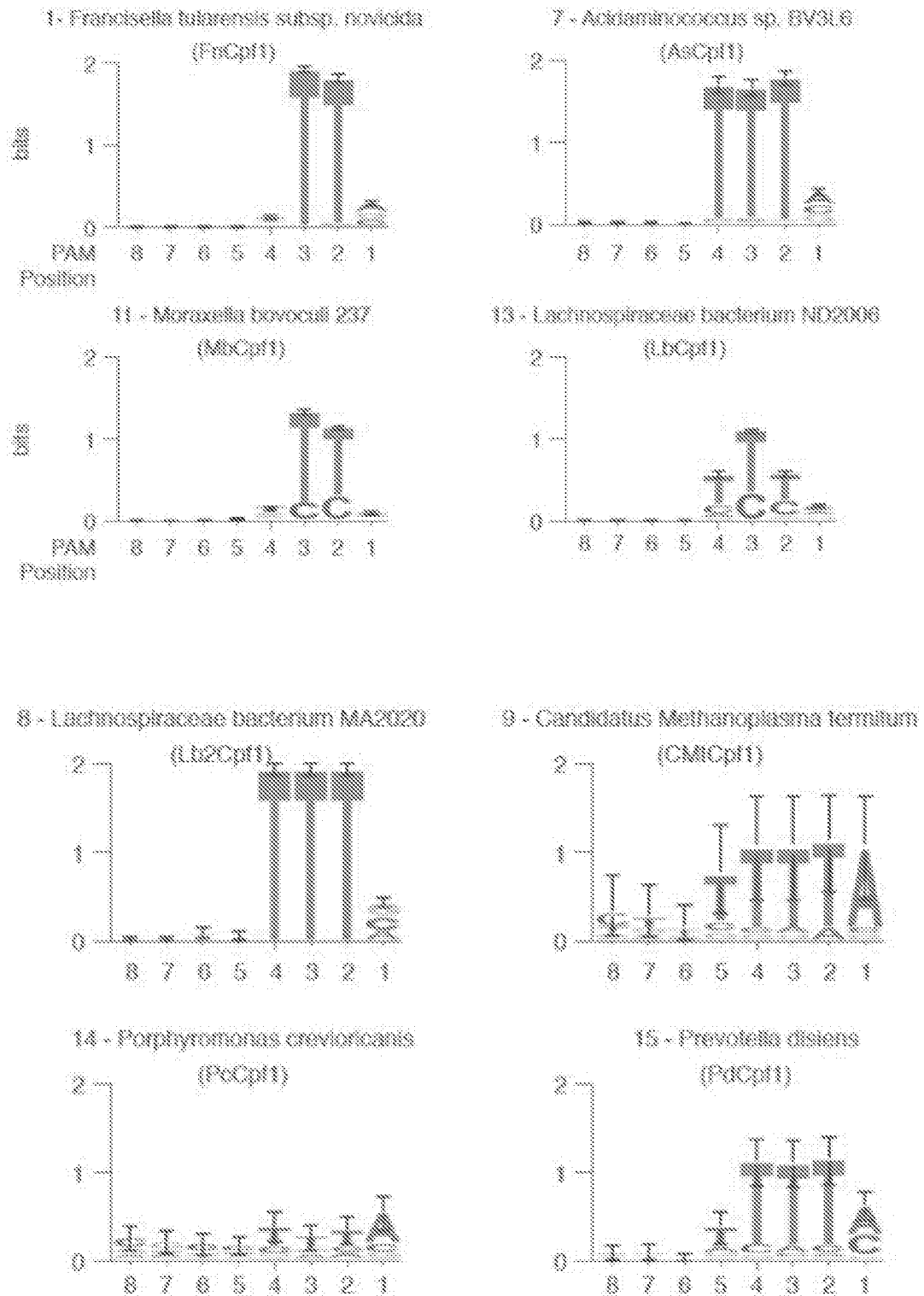


FIG. 100F

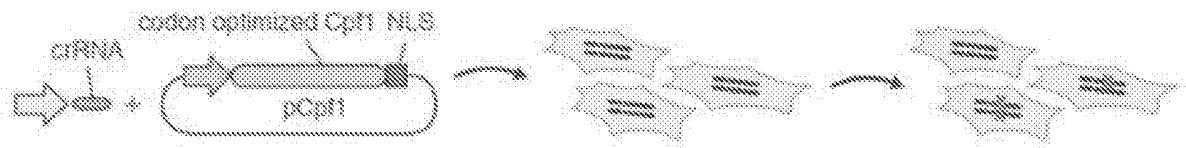


FIG. 101A



sequencing reads

```

AATGTTTCCTGA-----CTCGCCTG
AATGTTTCCTGA-----
AATGTTTCCTGATGGTCCATGTCTGT----CGCCTG
AATGTTTCCTGATGGTCCATGTCTGT----CGCCTG
AATGTTTCCTGATGGTCCATGT-TGT-----GCCTG
AATGTTTCCTGATGGTCCATGT-----ACTCGCCTG
AATGTTTCCTGATGGTCCATGTCTGT----CGCCTG
AATGTTTCCTGATGGTCCATGT----TACTCGCCTG
AATGTTTCCTGATGGTCCATGTCTGTT---CGCCTG
AATGTTTCCTGATGGTCCATGTCT--TACTCGCCTG
  
```

FIG. 101B

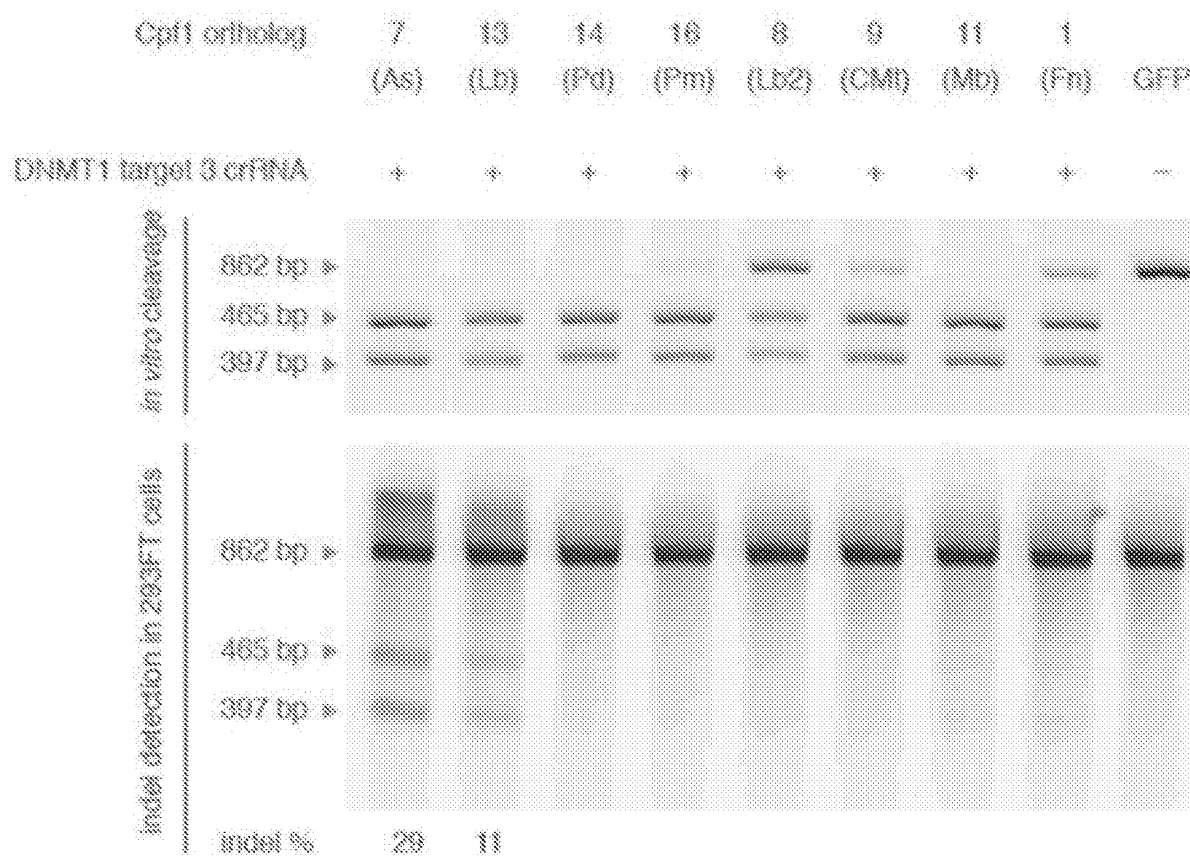


FIG. 101C

		PAM	guide sequence
Cpf1 DNMT1 target sites	1	TTTCCCTCACTCCTGCTCGGTGAATTT	
	2	TTTGAGGAGTGTTTCAGTCTCCGTGAAC	
	3	TTTCCTGATGGTCCATGTCTGTACTC	
	4	TTTATTTCCCTTCAGCTAAAATAAAGG	

		guide sequence	PAM
SpCas9 DNMT1 target sites	1	TCACTCCTGCTCGGTGAATTTGG	
	2	AACCCTCTGGGGACCGTTTGAGG	
	3	AGTACGTTAATGTTTCCTGATGG	
	4	TTTCCCTTCAGCTAAAATAAAGG	

FIG. 101D

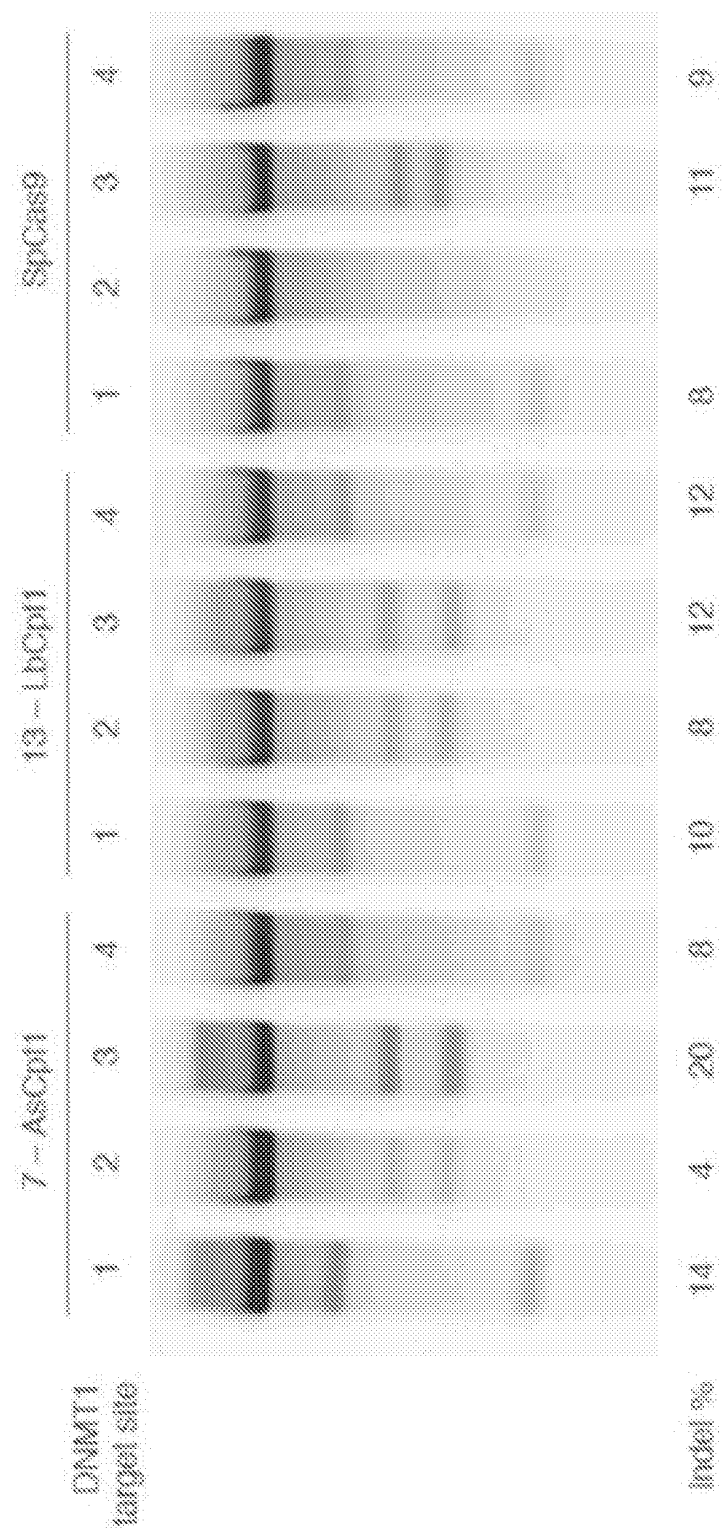


FIG. 101E

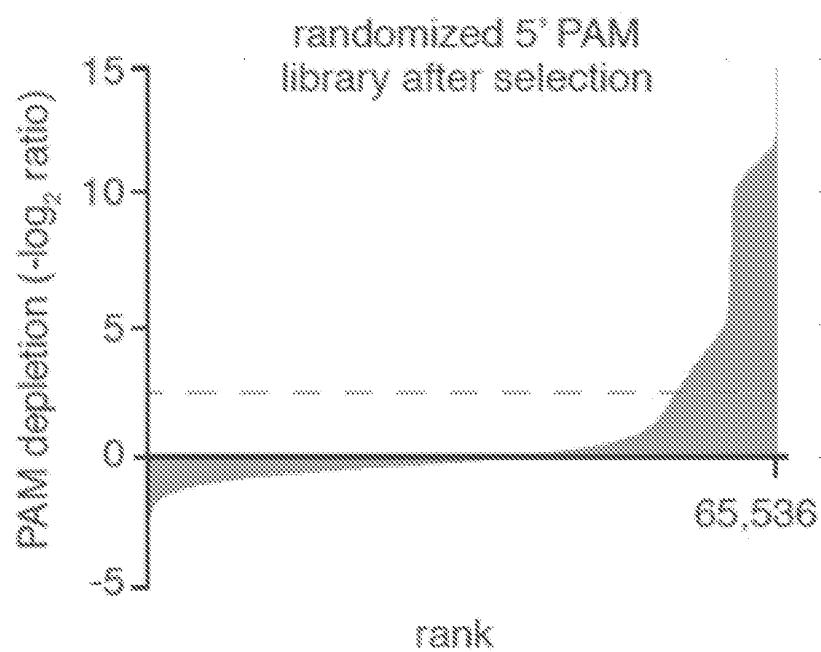


FIG. 102A

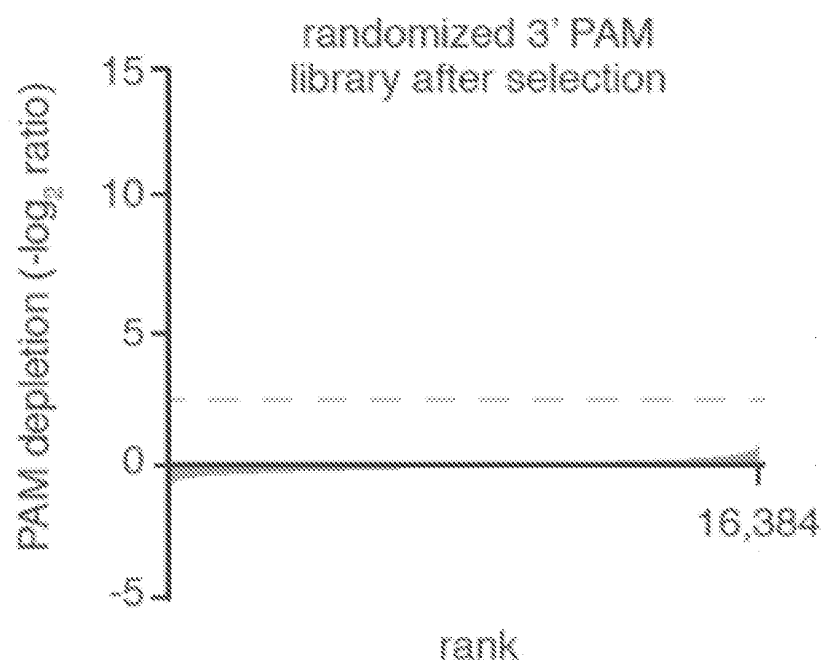


FIG. 102B

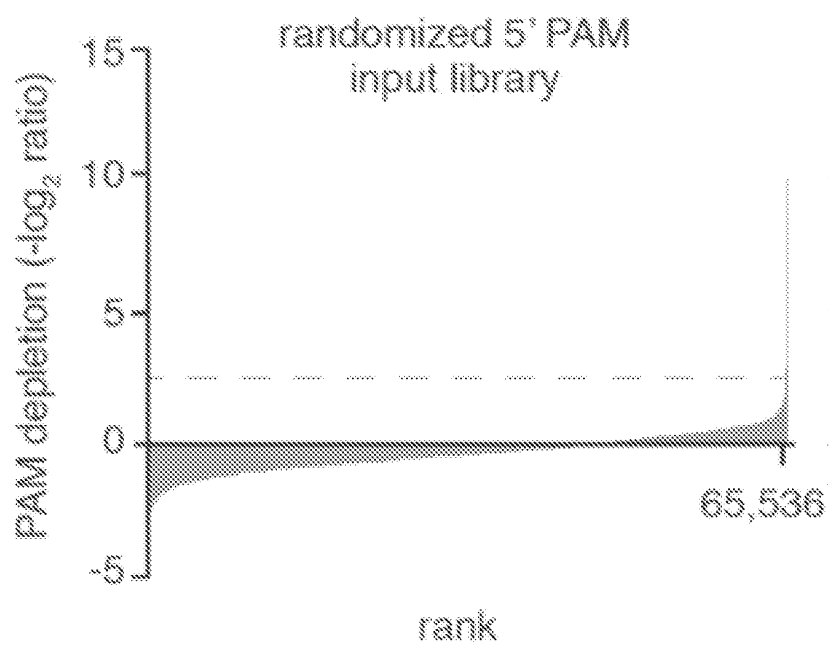


FIG. 102C

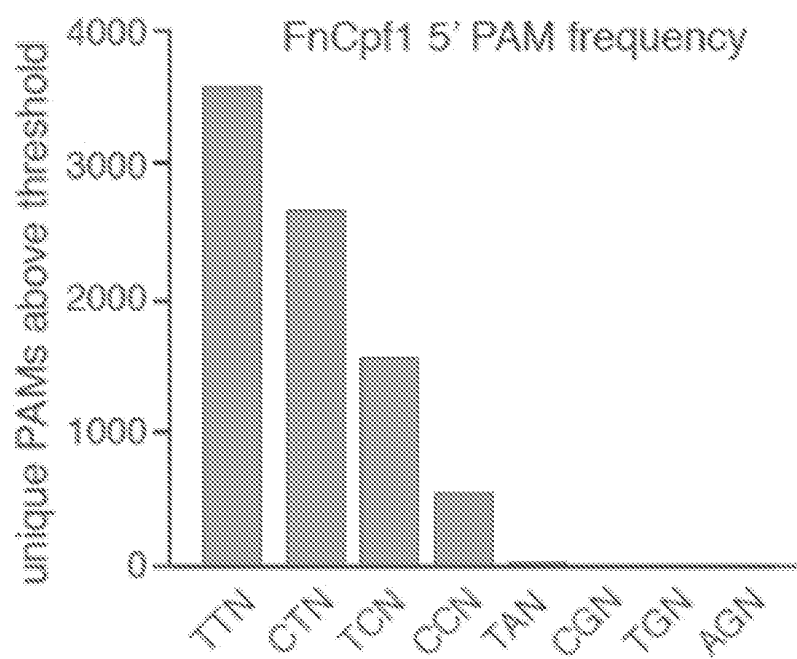


FIG. 102D

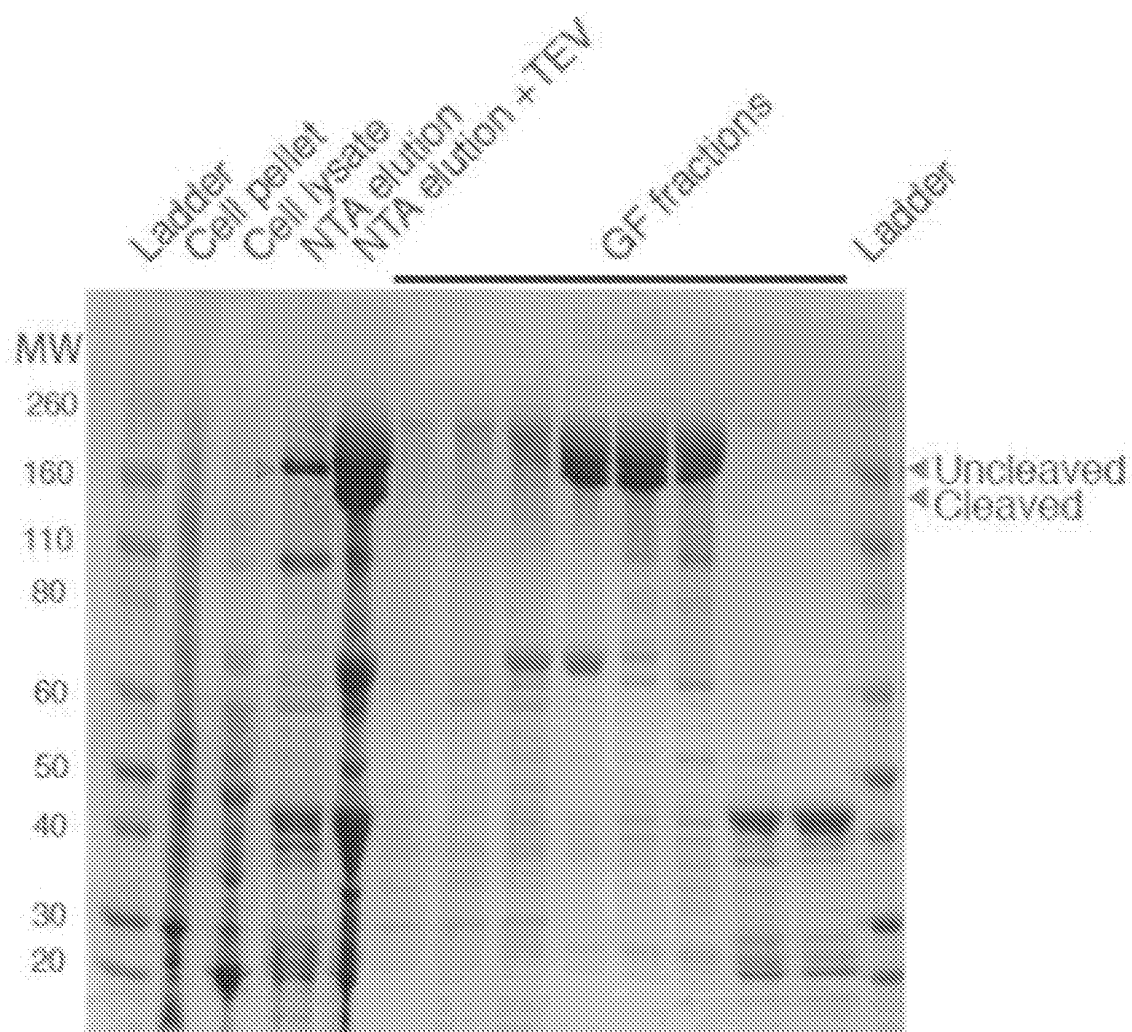


FIG. 103A

SEC elution of FnCpf1

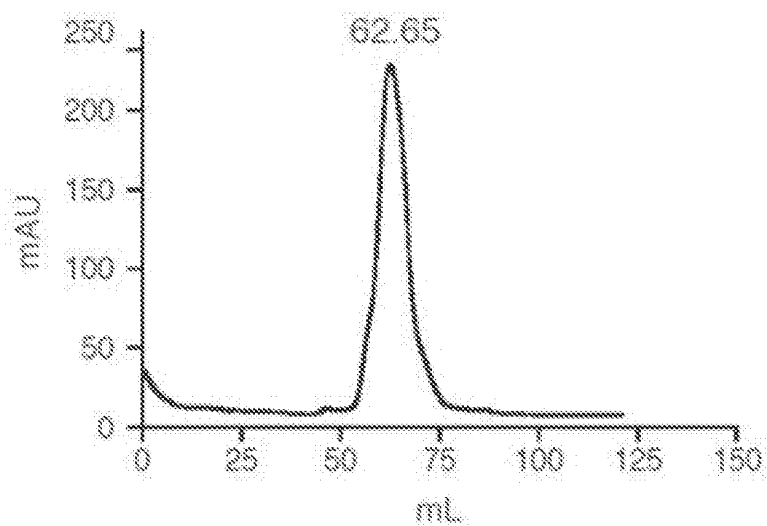


FIG. 103B

SEC calibration standards

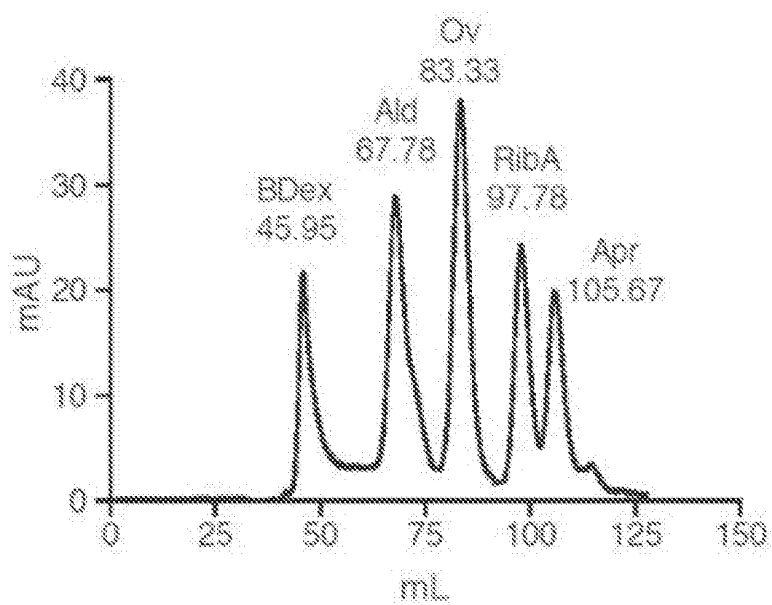


FIG. 103C

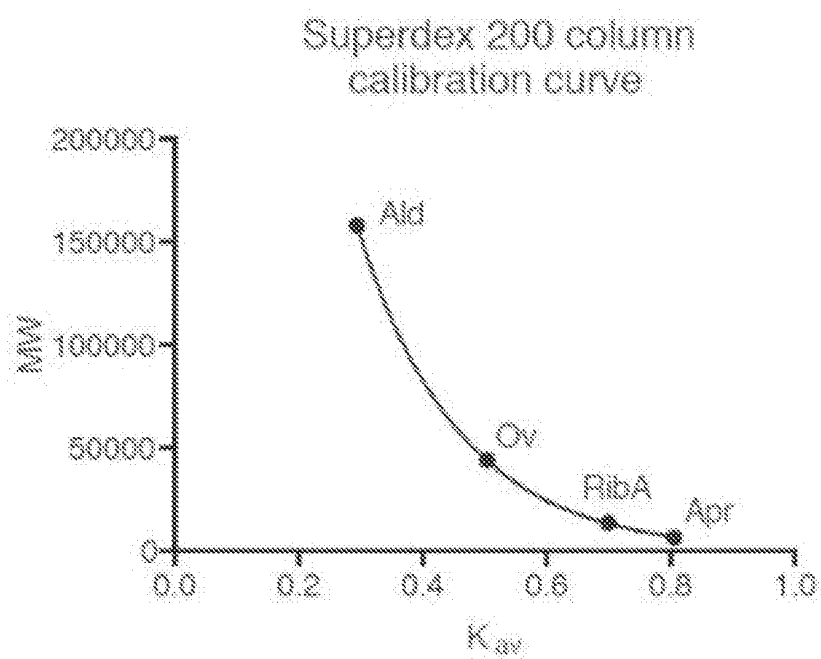


FIG. 103D

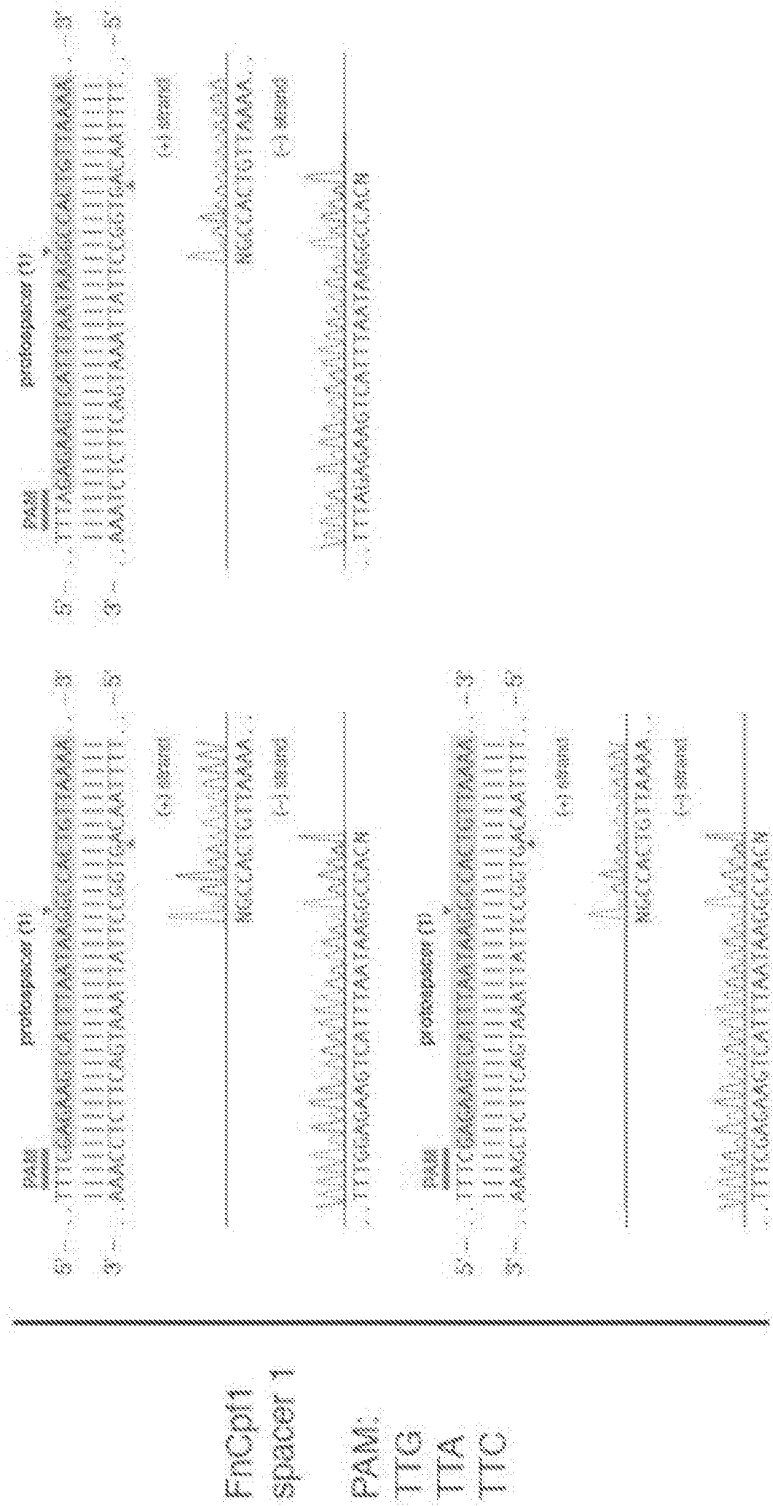


FIG. 104A

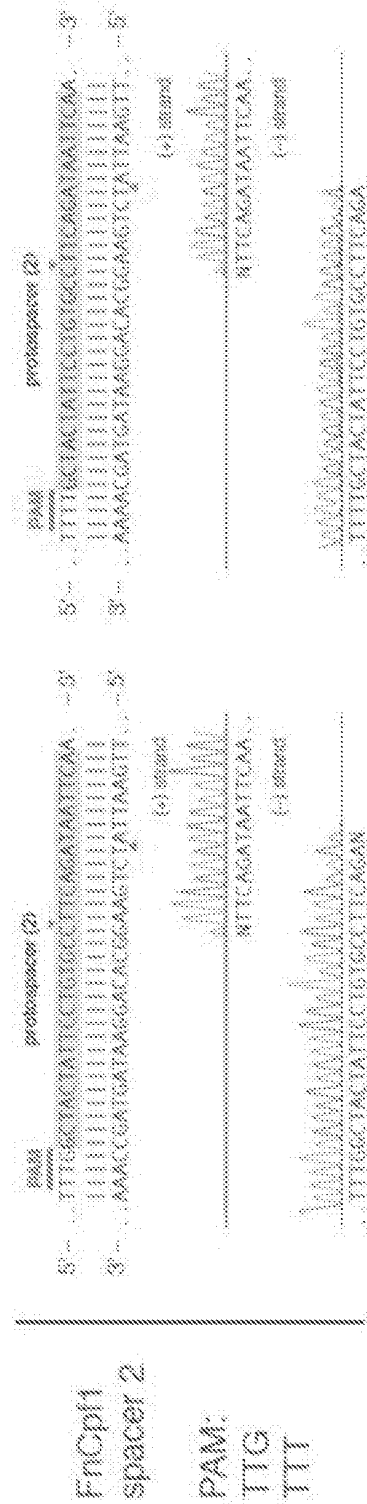


FIG. 104B



FIG. 104D

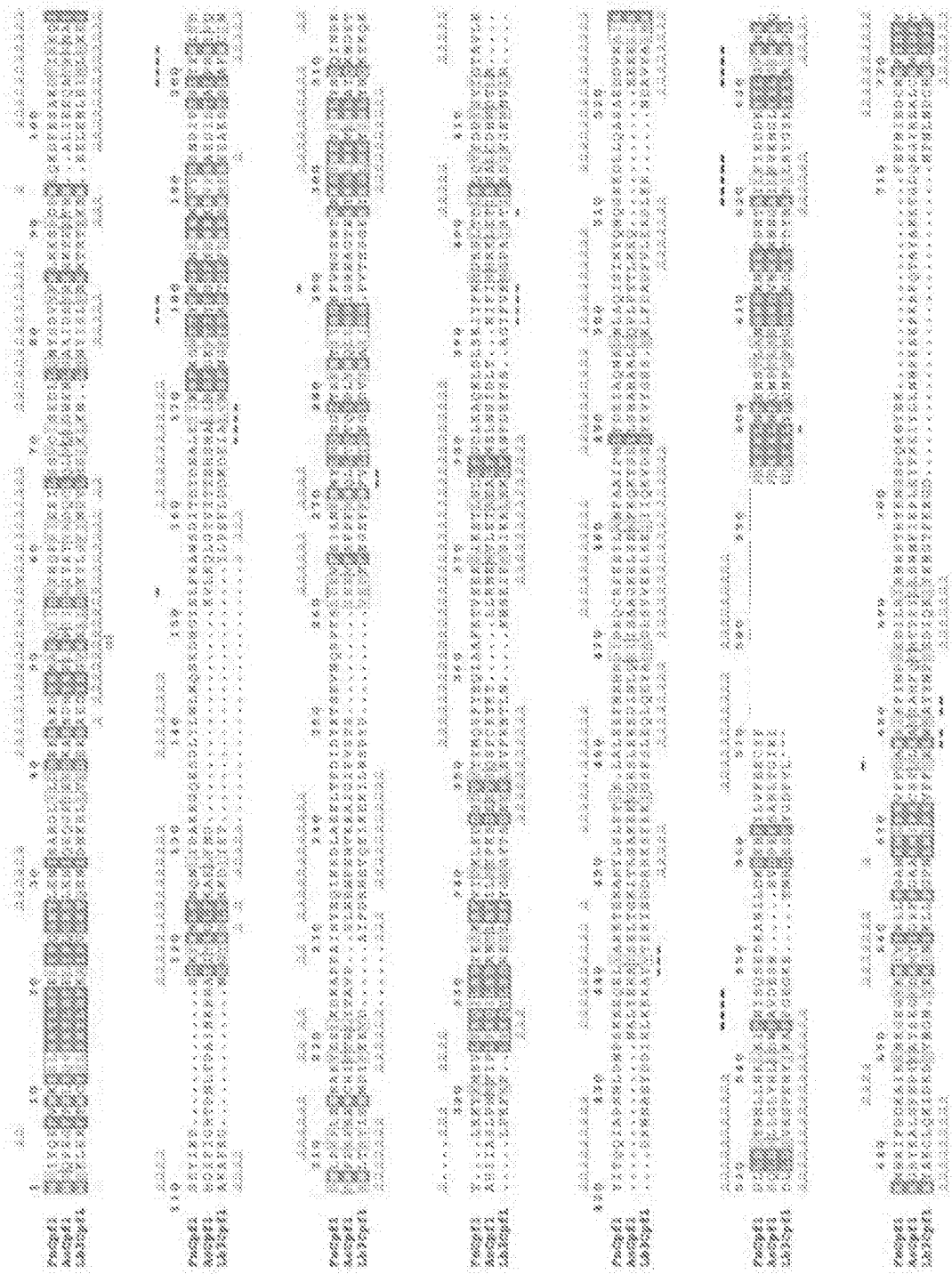


FIG. 105A

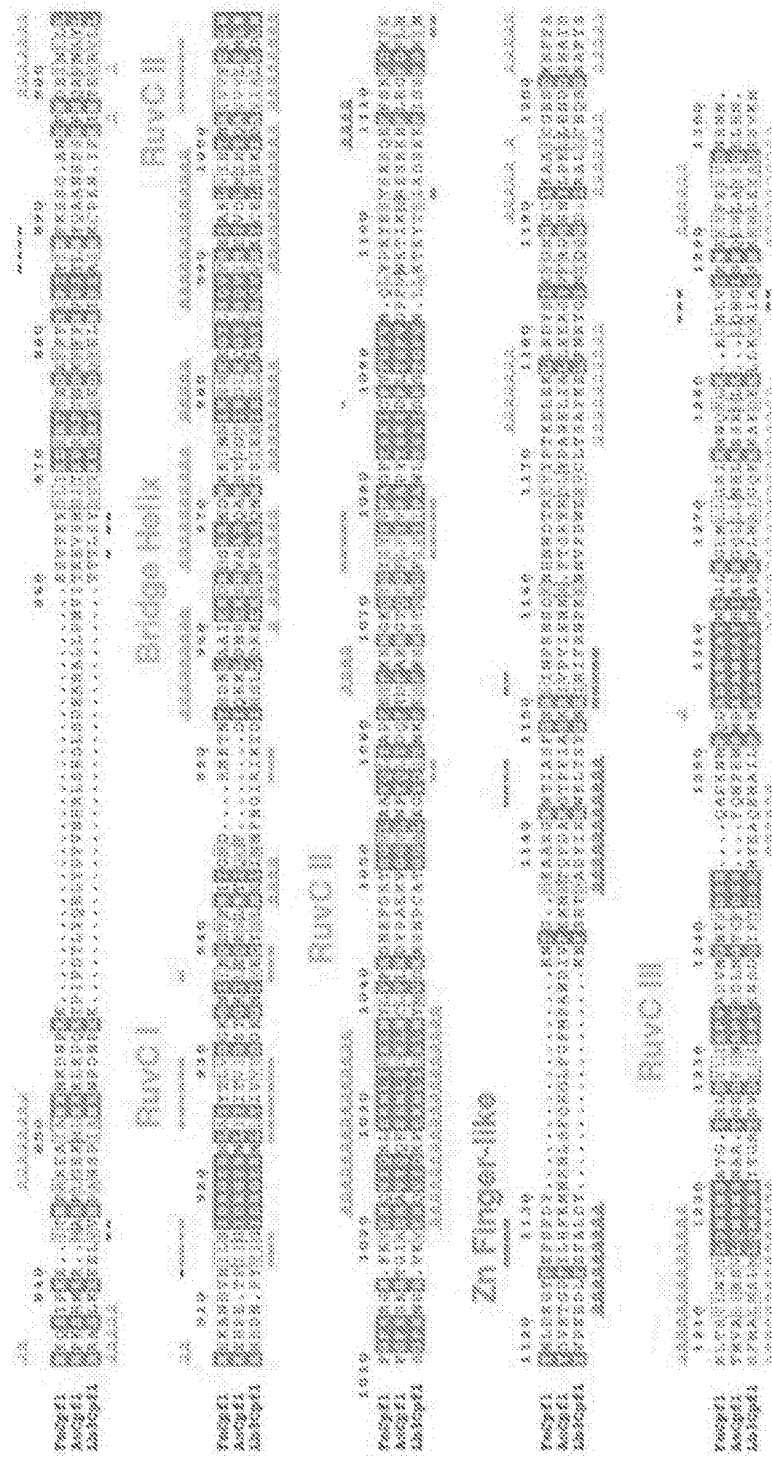
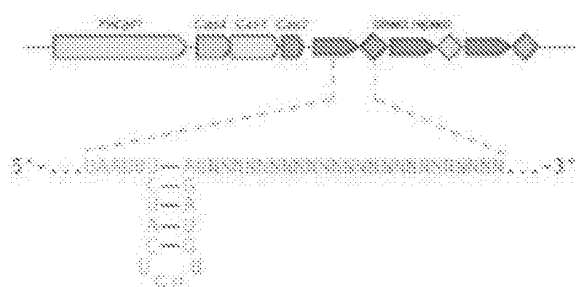
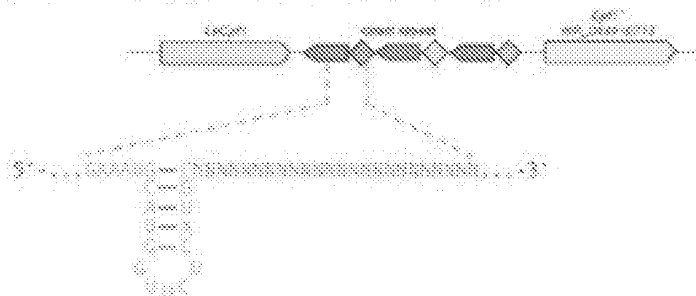


FIG. 105B

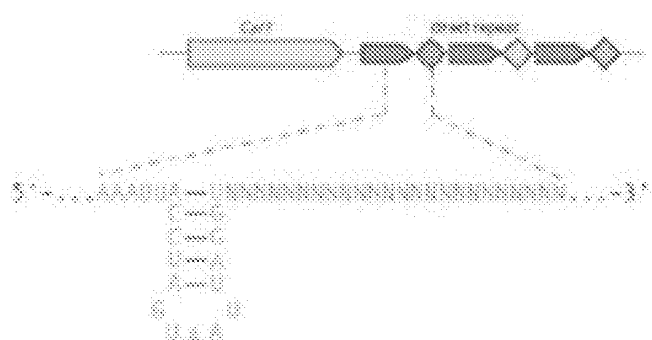
1. *F. tularensis* subsp. *novicida* U112 (NC_003601)



2. L. *hackeri* in MC2017 (12/12/2007)



3.3 polyelectrolytes (AC_01433)



4- *P. bacterium* GW2011_GWA_33_10 (LS0001000015)

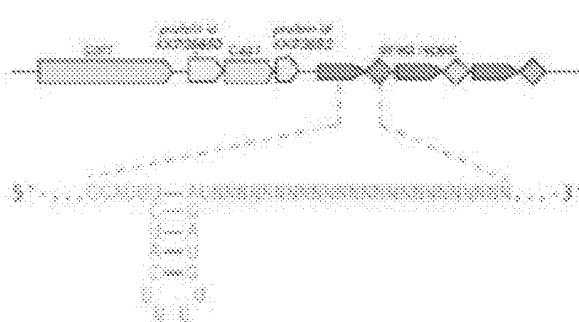


FIG. 106A

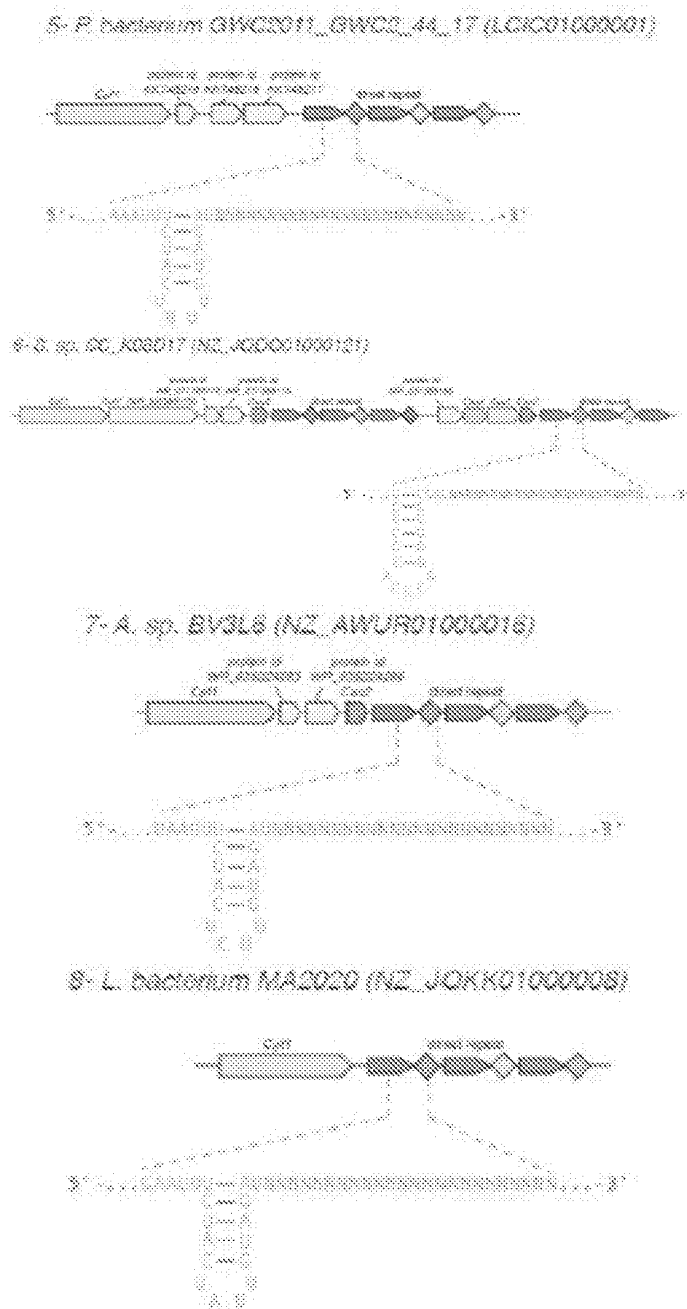


FIG. 106B

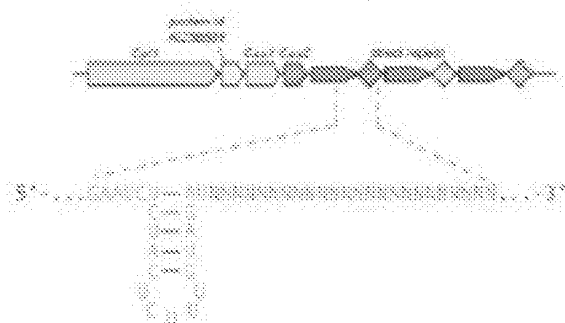
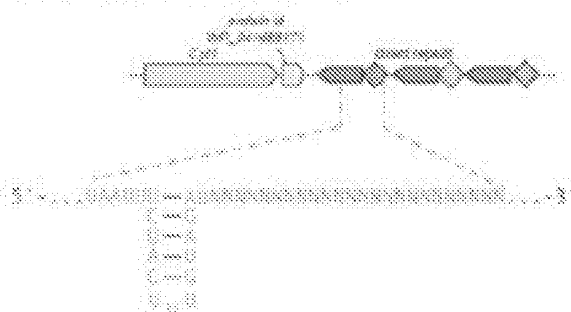
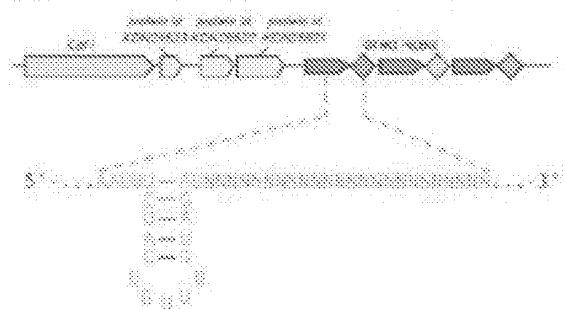
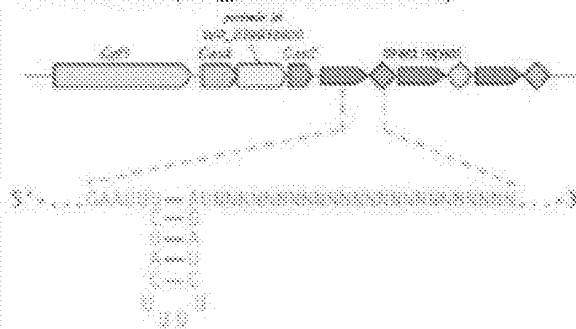
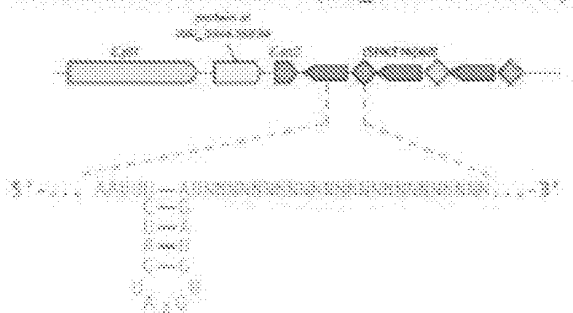
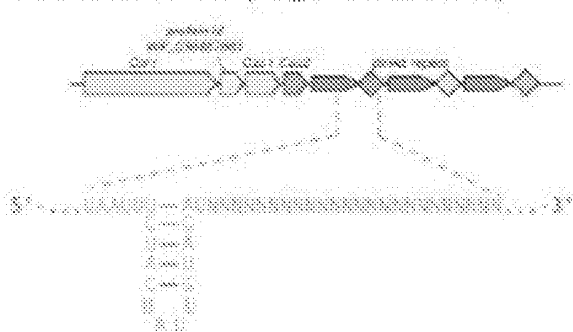
9- *C. Methanoplasma termitum* (CP010070)10- *E. eligens* (NC_012778)11- *M. boydii* 237 (ACMT01000011)12- *L. inadae* (NZ_AHMM02000017)

FIG. 106C

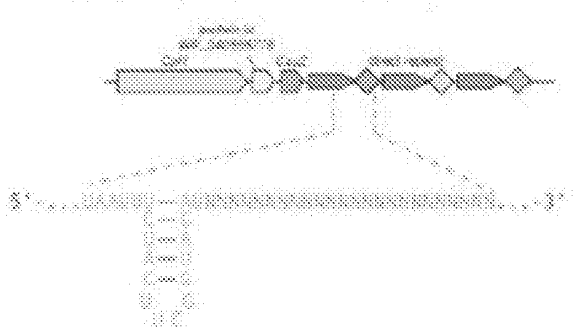
13- *L. bacterium* HD2006 (NZ_JNKG01000011)



14- *P. crevioricanis* (NZ_JQJC01000021)



15- *P. disiens* (NZ_AEDC01000001)



16- *P. macacae* (NZ_BAKC01000001)

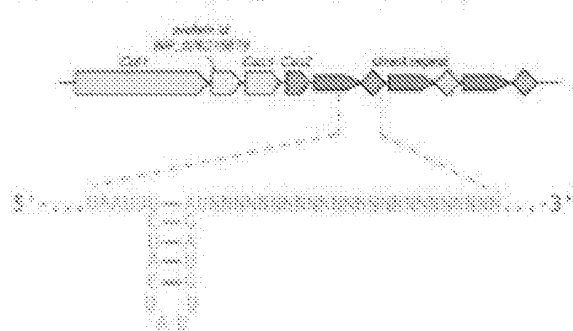


FIG. 106D

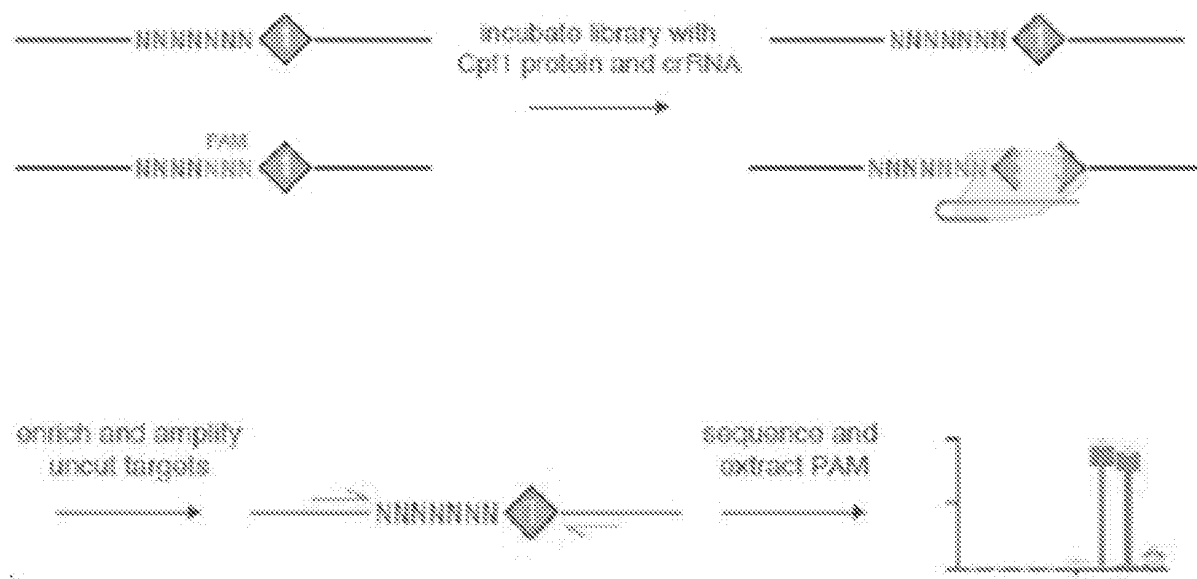


FIG. 107A

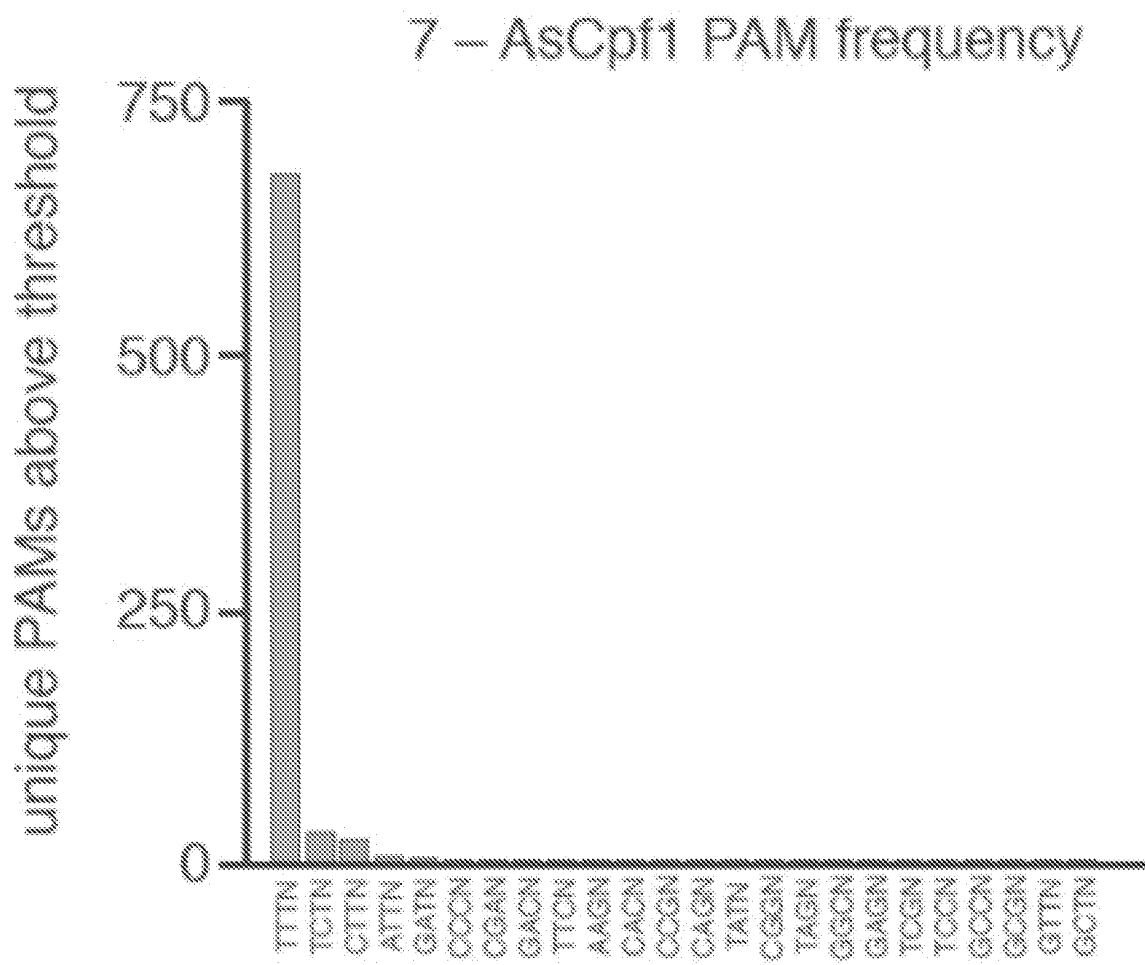


FIG. 107B

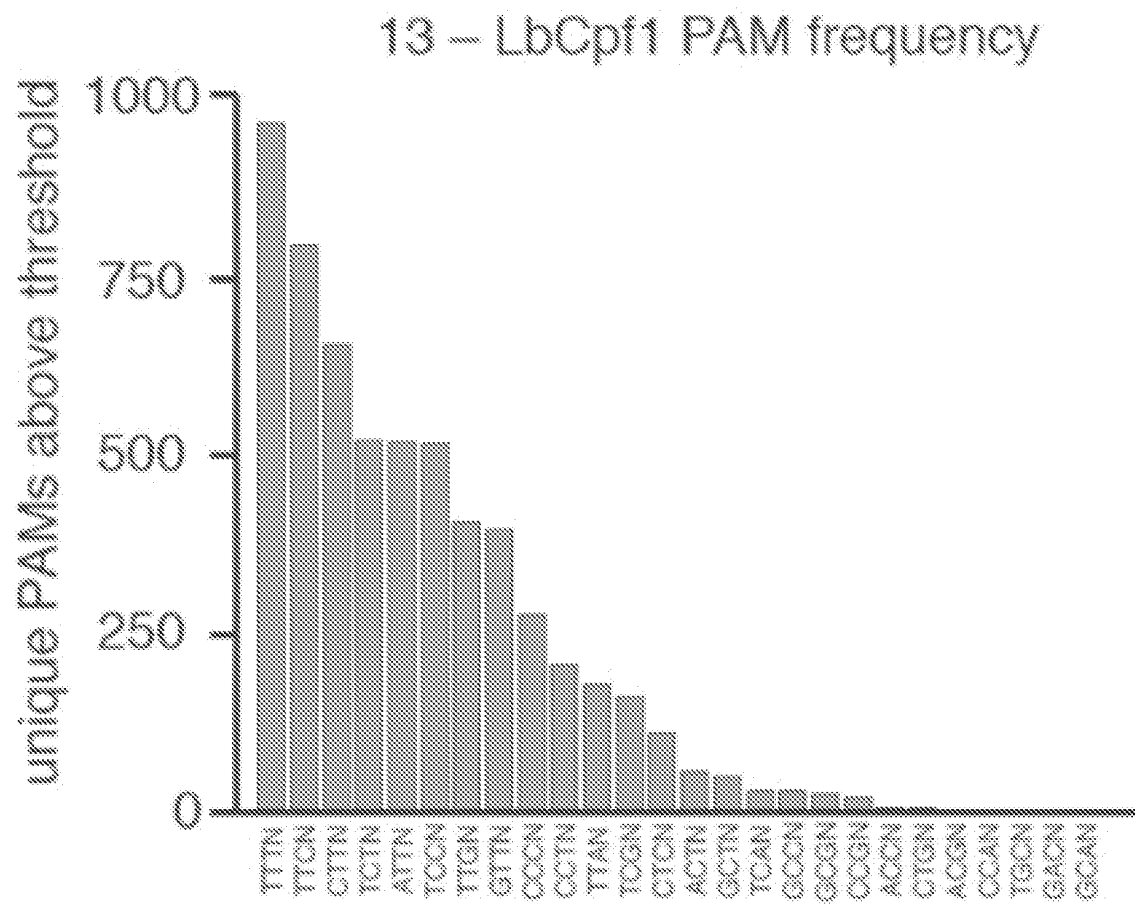


FIG. 107C

7 – AsCpf1, protospacer (1)

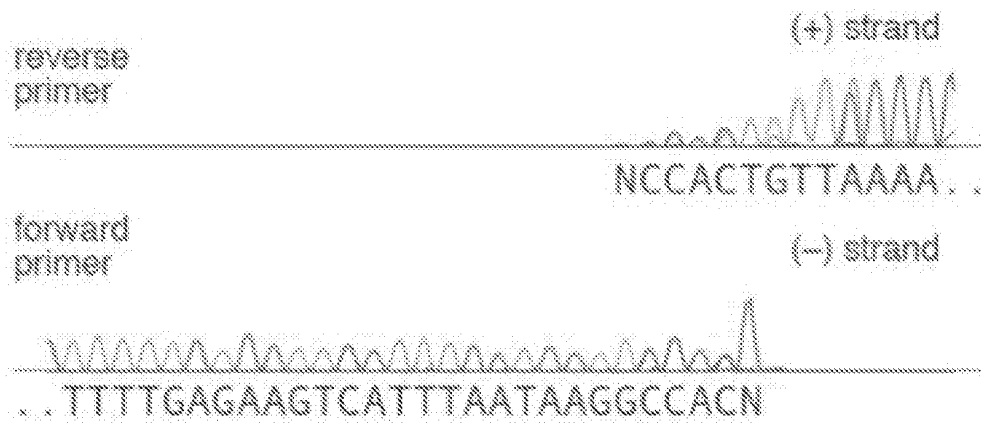
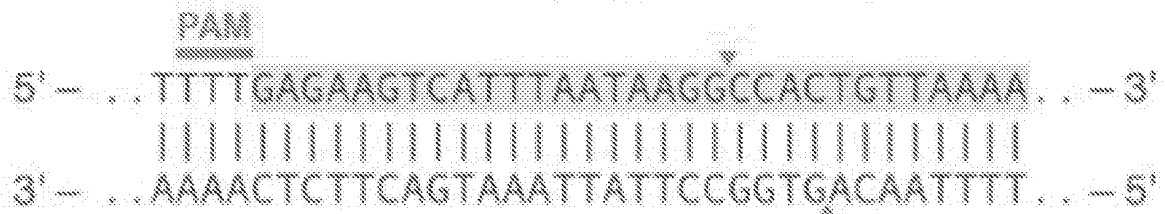


FIG. 107D

13 – LbCpf1, protospacer (1)

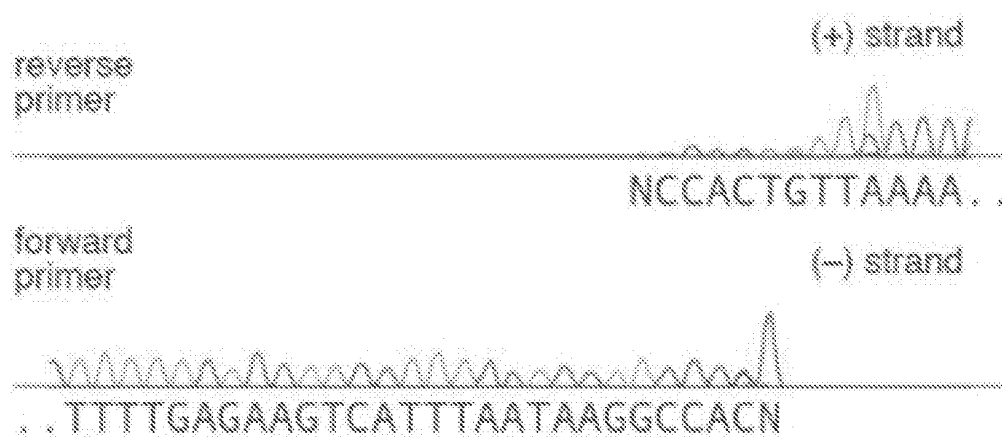
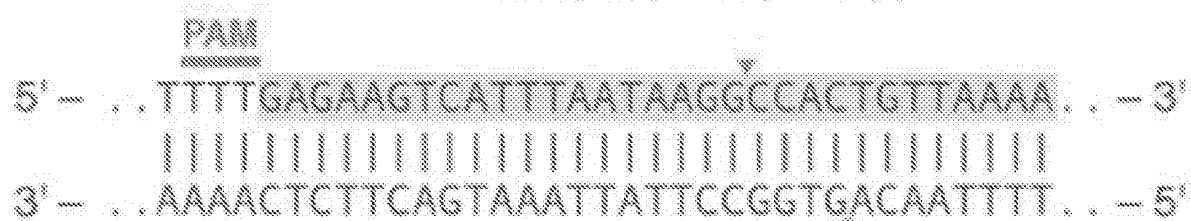


FIG. 107E

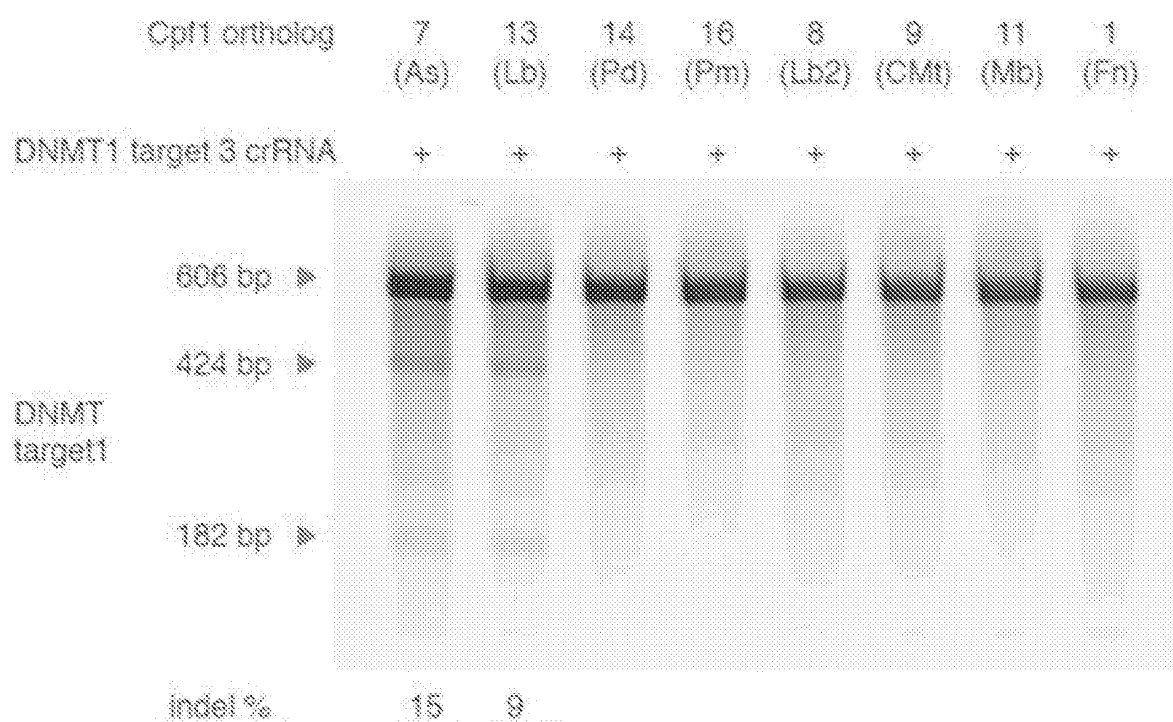


FIG. 108A

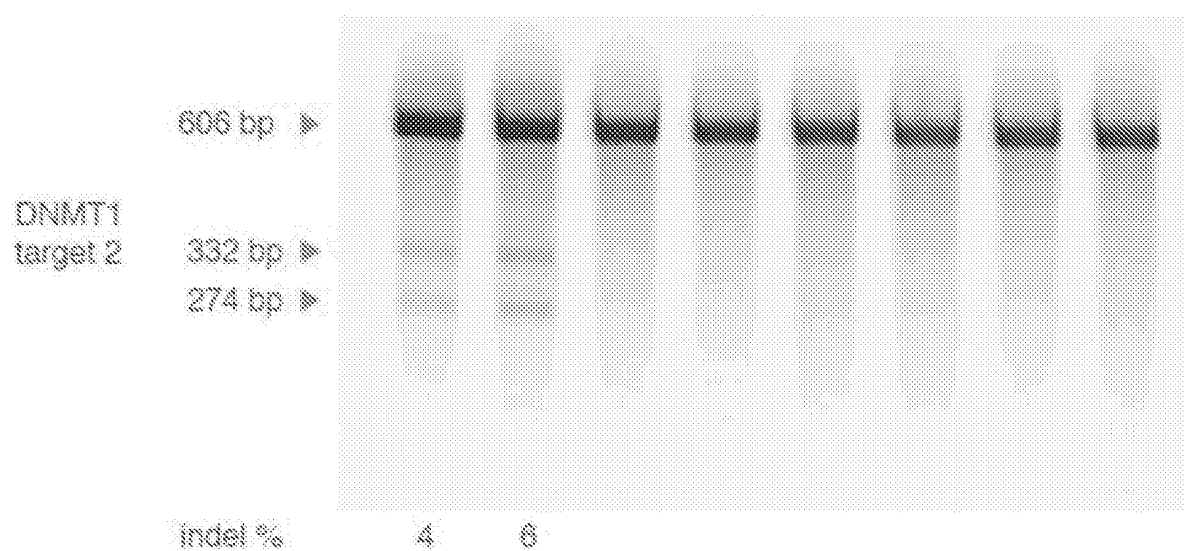


FIG. 108B

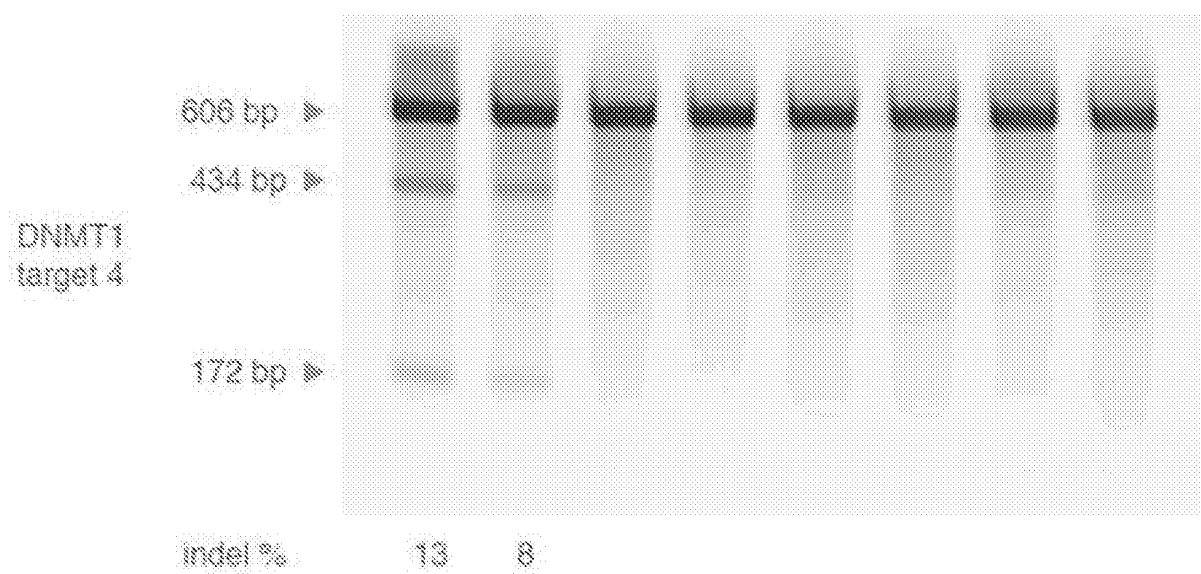


FIG. 108C

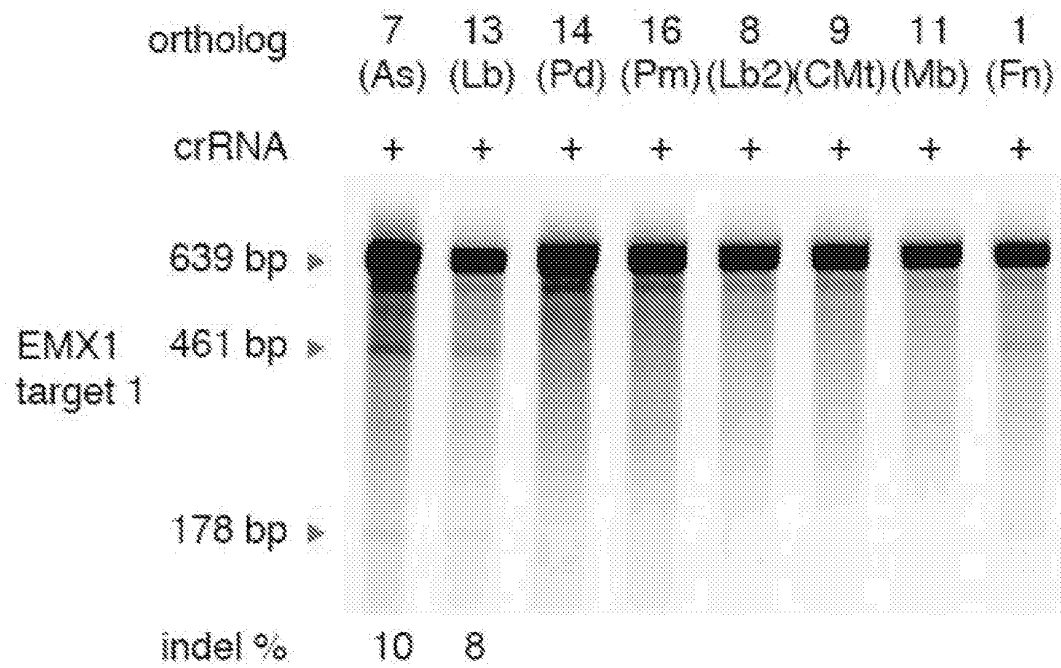


FIG 108D

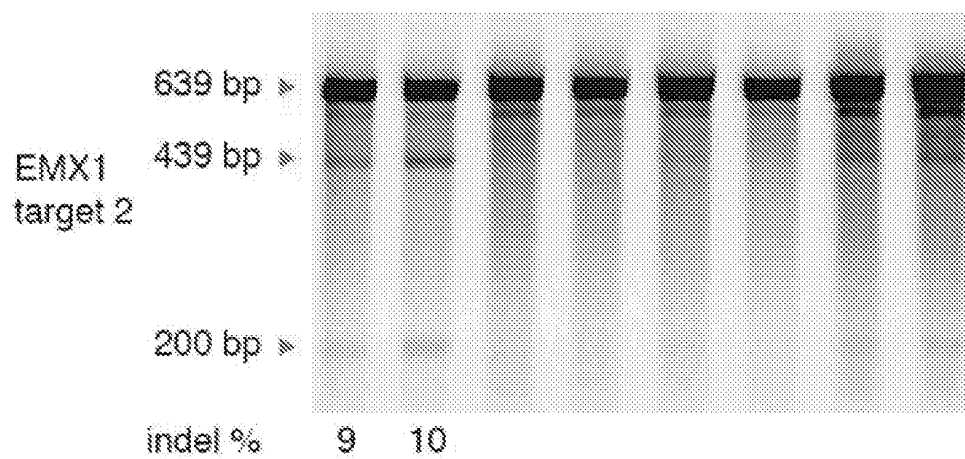


FIG 108E

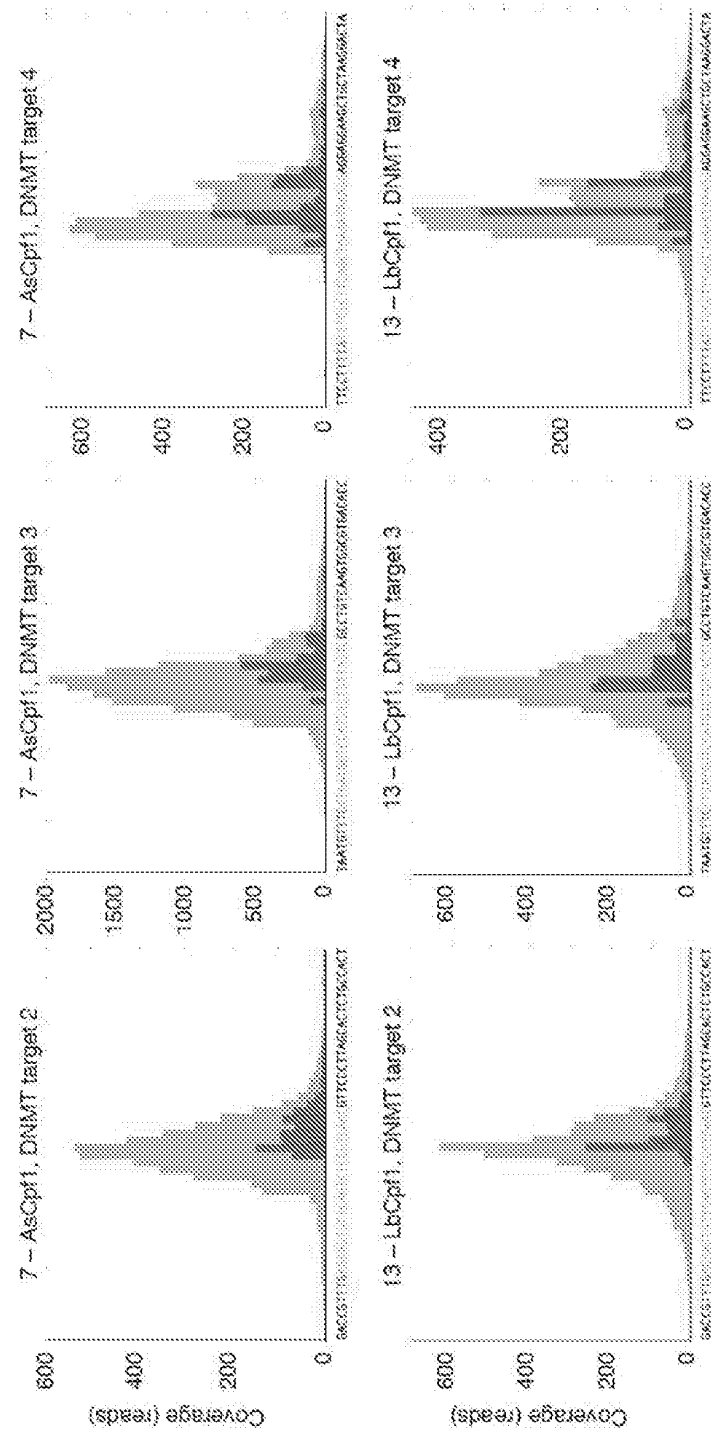


FIG. 108F

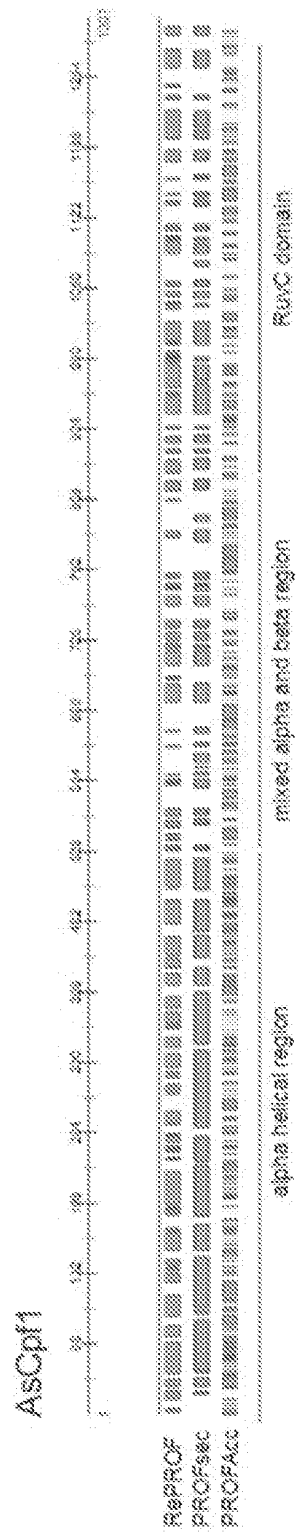


FIG. 109B

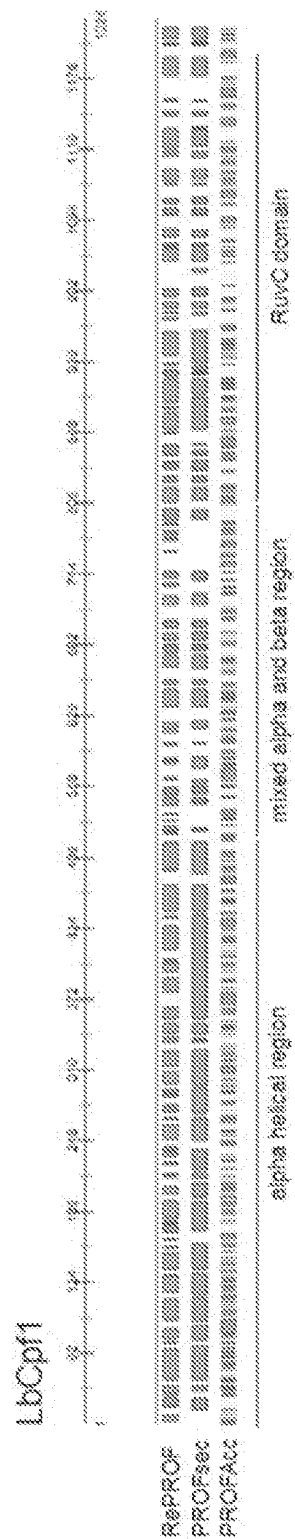


FIG. 109C

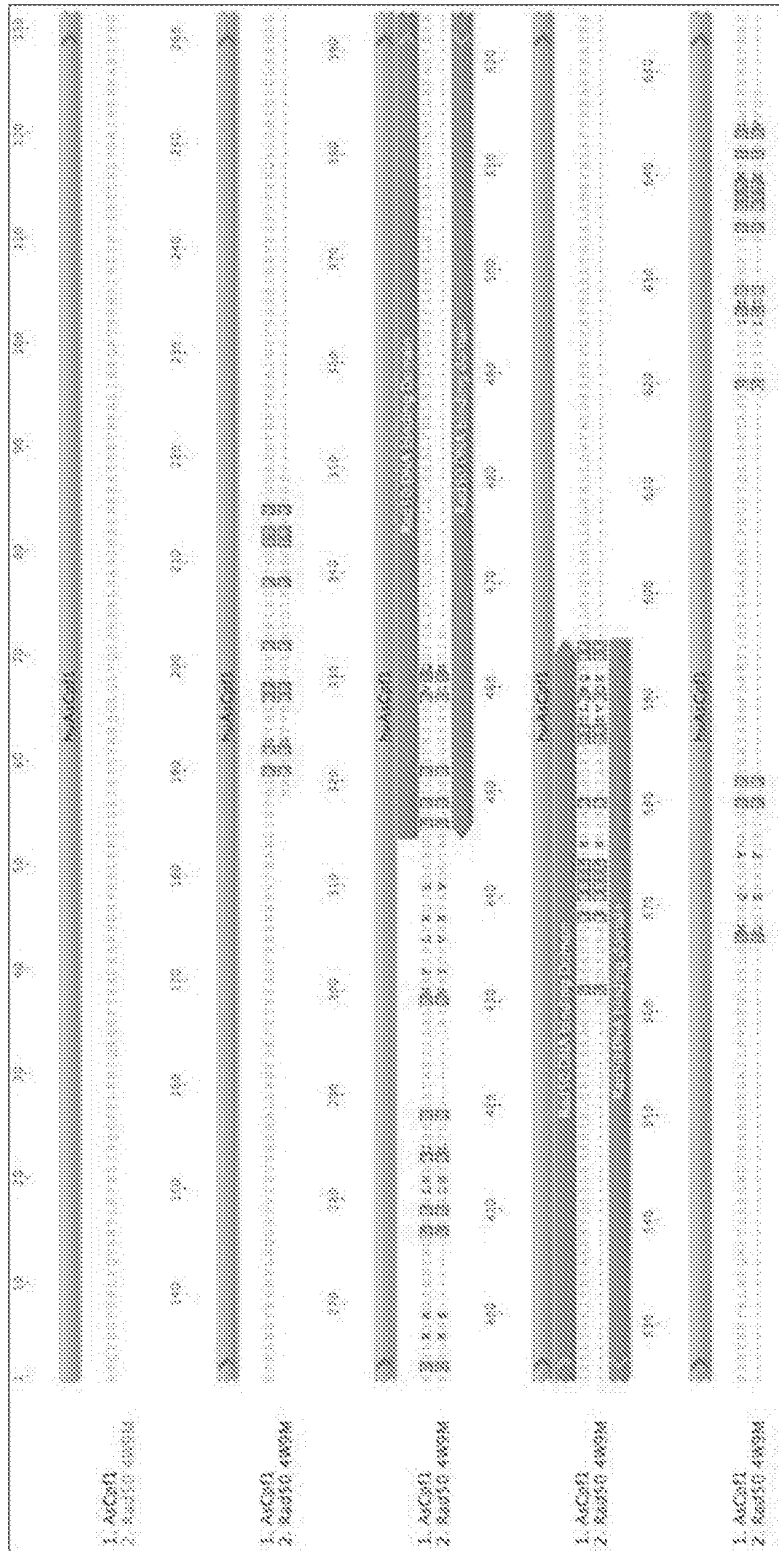


FIG. 110A

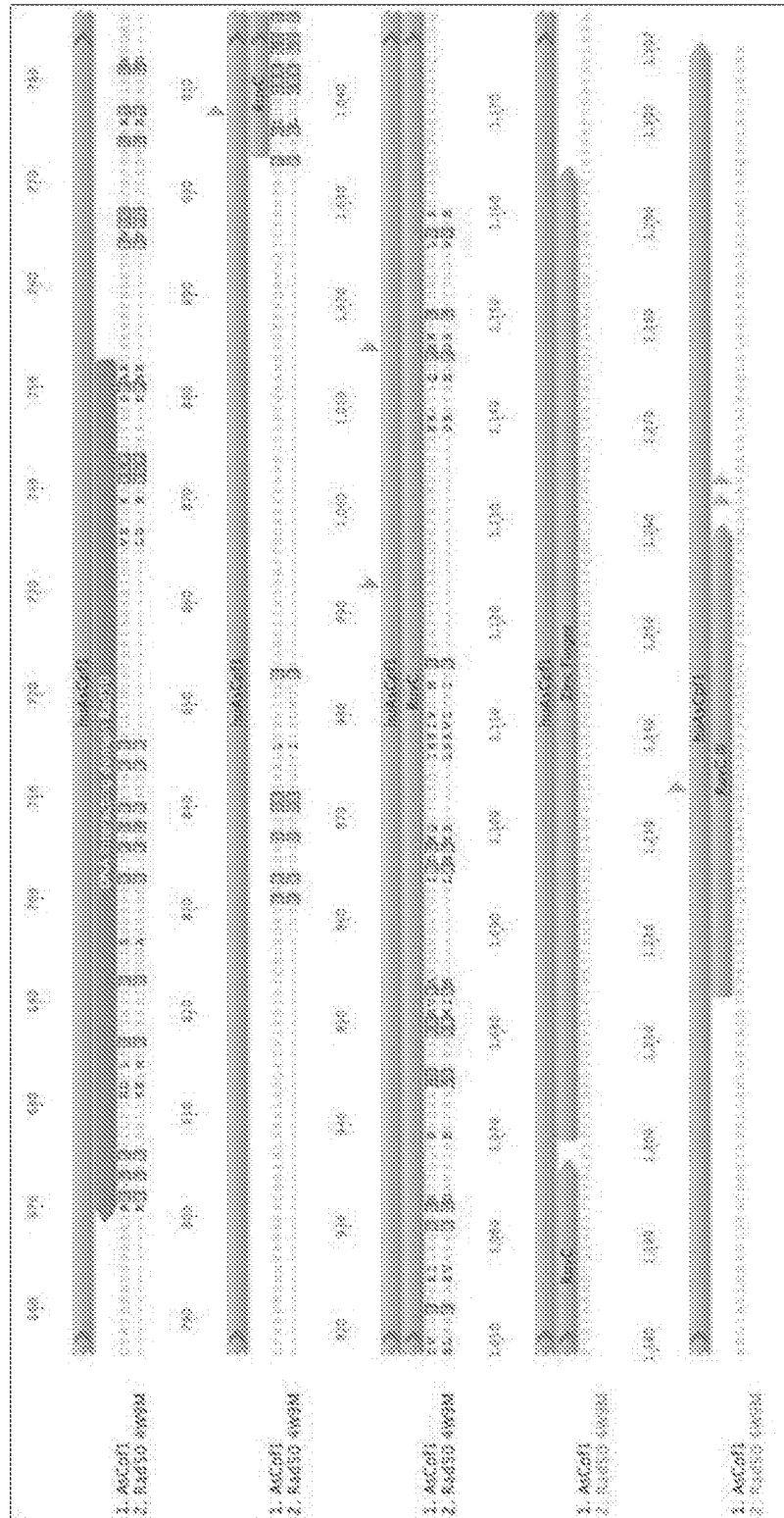


FIG. 110B



FIG. 110C

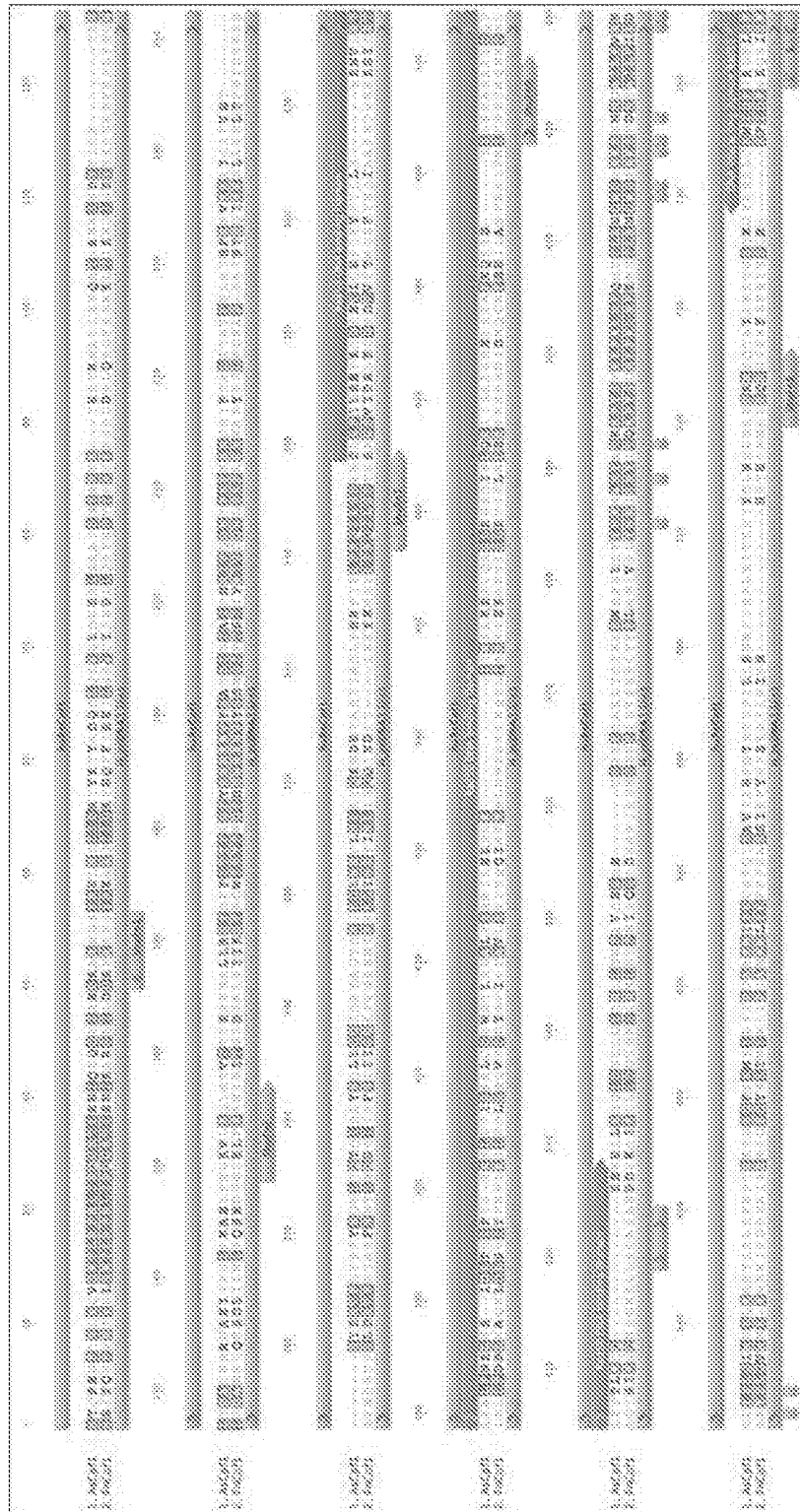


FIG. 110D

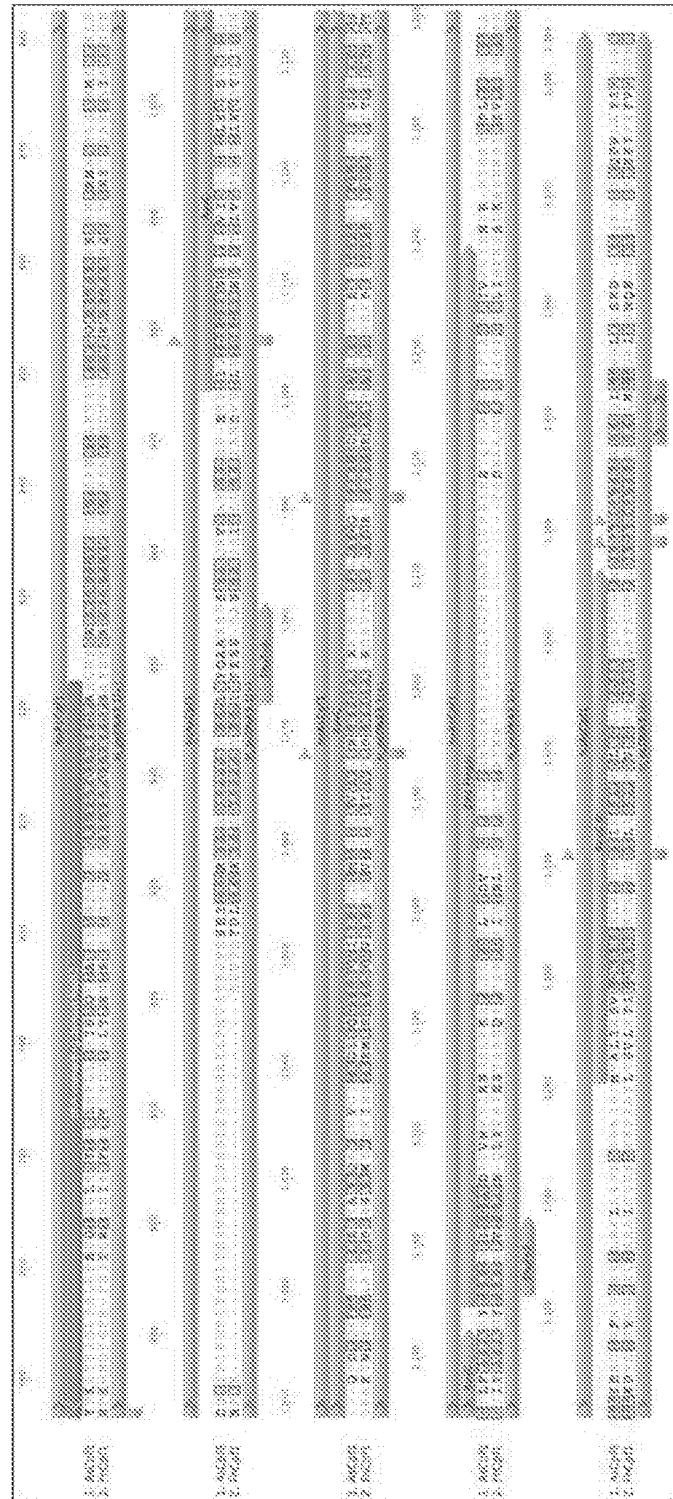


FIG. 110E

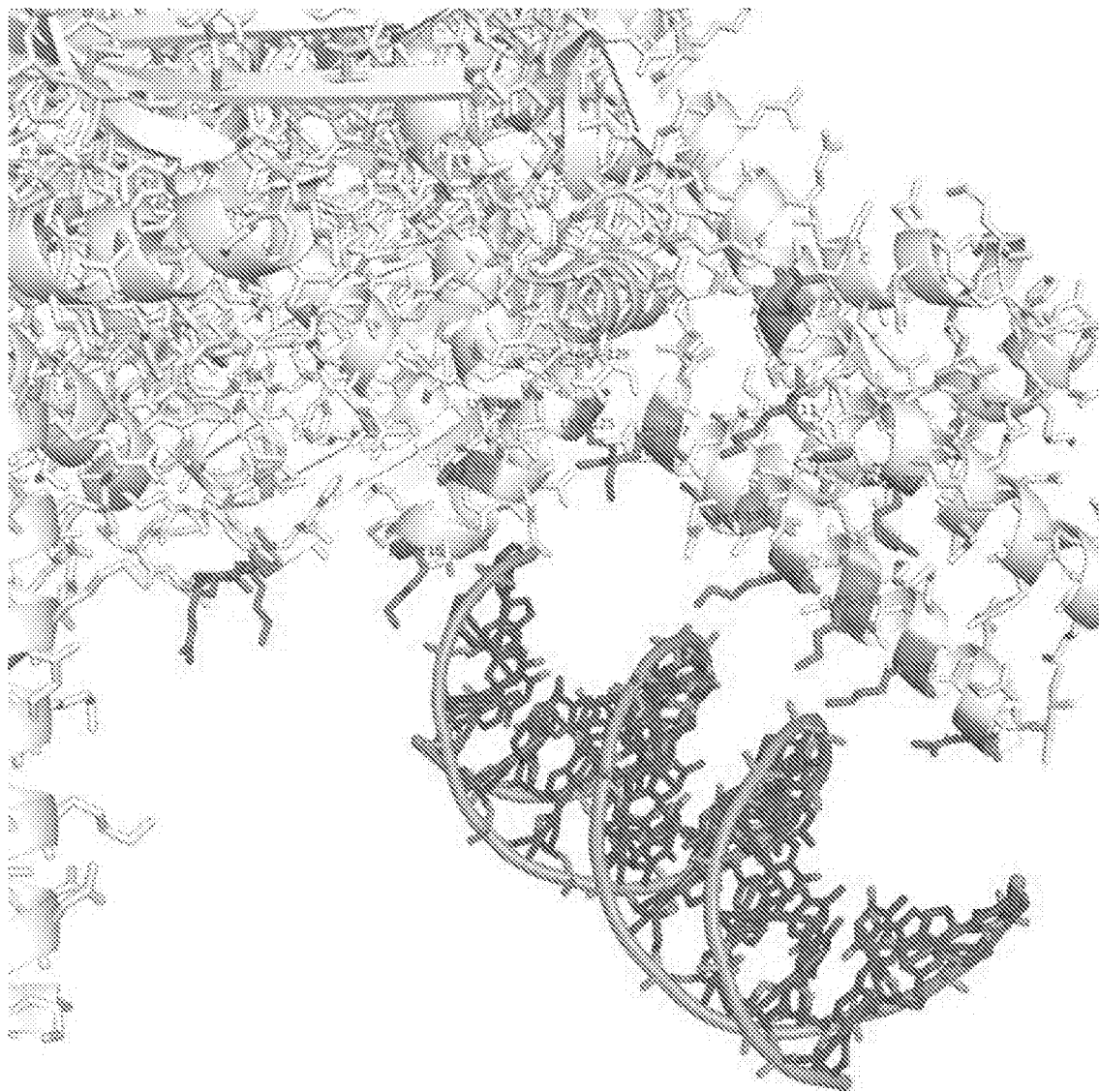


FIG. 111

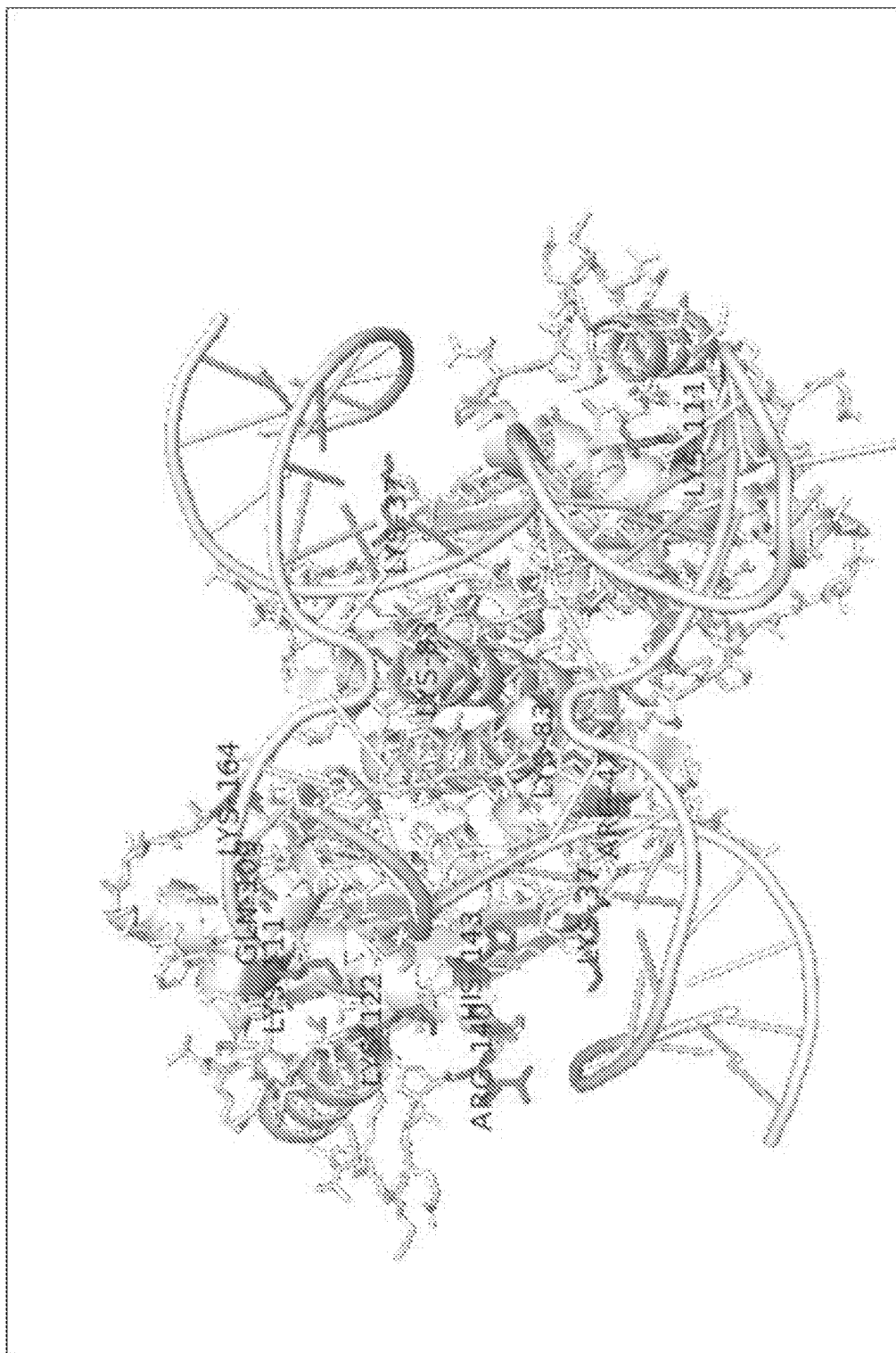


FIG. 112

Site-specific recombinase XerD [Roseburia intestinalis M50/1]

Sequence ID: [emb|CHL10059.1](#) Length: 416 Number of Matches: 1

Range 1: 175 to 292 Genes: Graphics

More Match Previous Match

Score	Expect	Method	Identities		Positives	Gaps
46.2 bits(108) 0.14 Compositional matrix adjust. 42/127(33%) 59/127(46%) 17/127(13%)						
Query	750	FAKCHCKP-----NLHTLYWTCLFSPENLAKTSIKLNQQAELFYRPKSRMKRMARICEK				805
		F +C KP LYWTC+ E LA + ++ +L + R R+C K				
Sbjct	175	FRCVADKPLSYICFEVLYWTCNRECELLALSPADIDIDNKLI-----SINRTYQRIQCK				229
Query	806	NLKKRLKQCKT-----PIPDLYQELDYVYVNHRLSHDLSDPEARALLPNVITKEVSHEIIND				861
		+ K +K+ PIPD L QEL DY+ R D +DE L P V +SHE+I+				
Sbjct	230	DVFTSPKTRKSKRTIPIPDFLCOELSDYIQSRYNLD--ADE---RLFP-VTKSYLSHENIRG				285
Query	862	RRFTSDK 868				
		+ T K				
Sbjct	286	CRITGAK 292				

FIG. 113

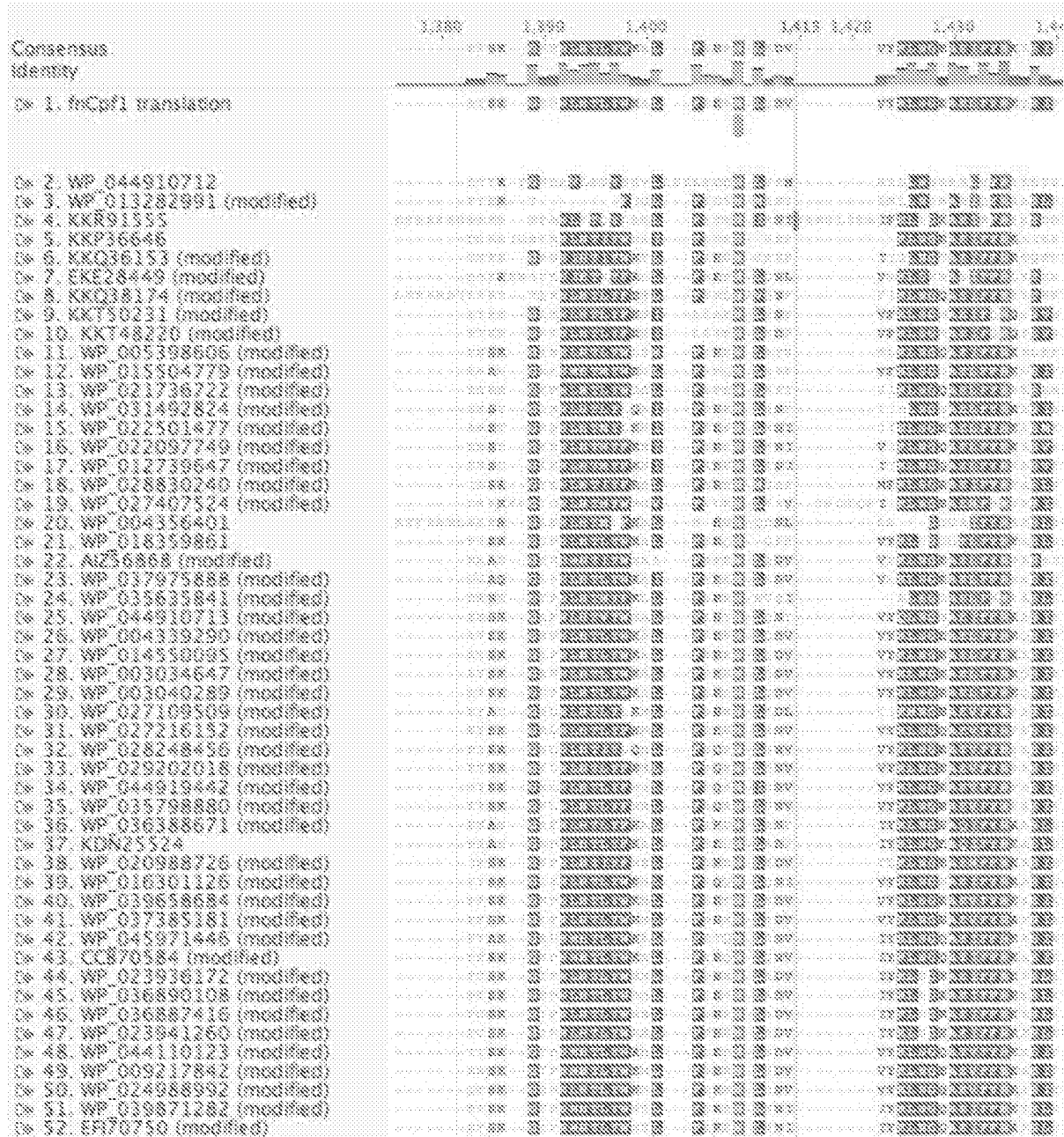


FIG. 114

Plasmids carrying tandem guides for AsCpf1

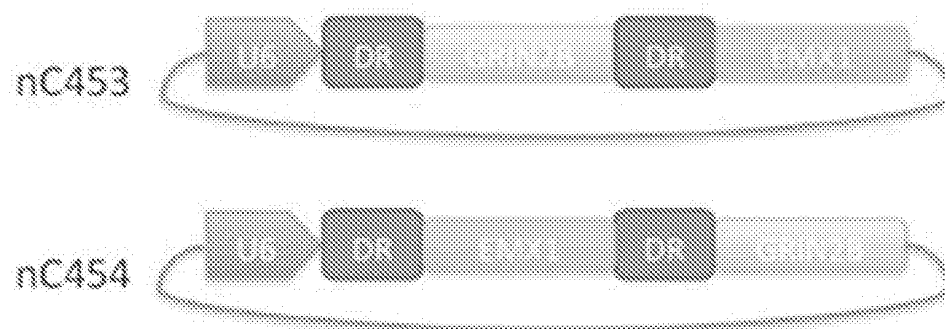


FIG. 115A

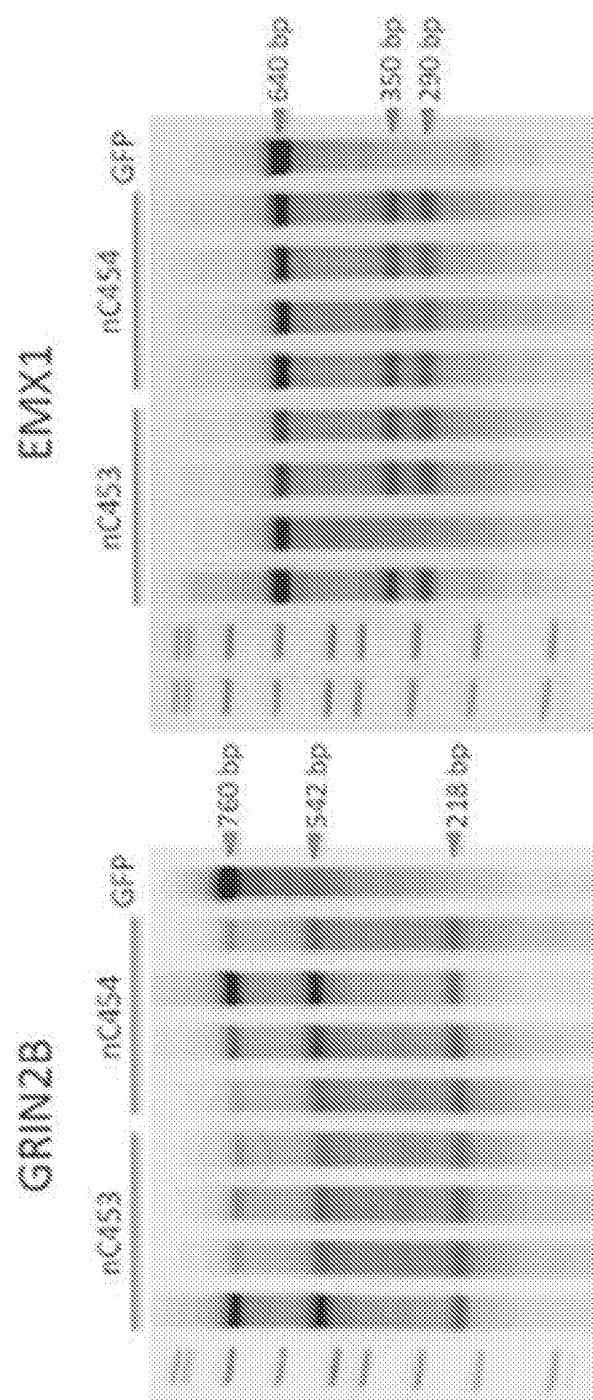


FIG. 115B

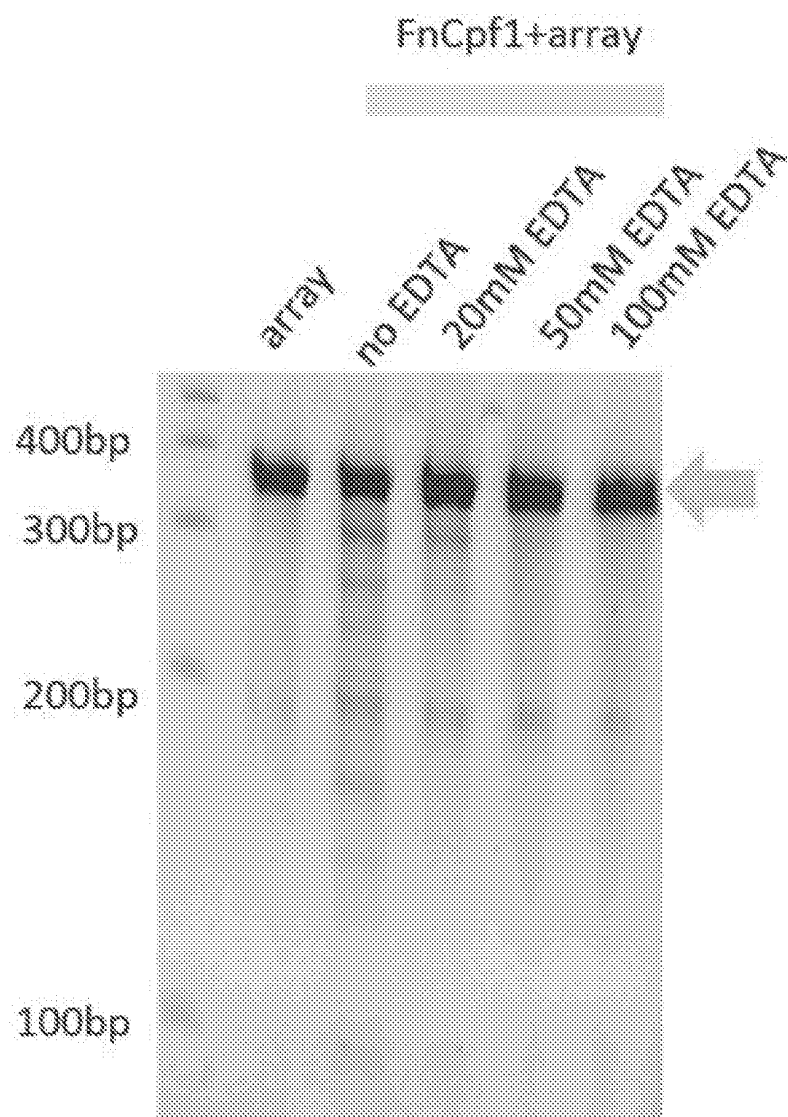


FIG. 116