



US 20250250630A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2025/0250630 A1**

Mozhayskiy et al.

(43) **Pub. Date:** **Aug. 7, 2025**

(54) **CALIBRATION PANELS AND METHODS  
FOR DESIGNING THE SAME**

(71) Applicant: **Life Technologies Corporation,**  
Carlsbad, CA (US)

(72) Inventors: **Vadim Mozhayskiy**, San Diego, CA  
(US); **Yutao Fu**, San Marcos, CA (US);  
**Earl Hubbell**, Palo Alto, CA (US)

(21) Appl. No.: **19/015,117**

(22) Filed: **Jan. 9, 2025**

**Related U.S. Application Data**

(63) Continuation of application No. 17/811,192, filed on Jul. 7, 2022, now abandoned, which is a continuation of application No. 14/975,001, filed on Dec. 18, 2015, now abandoned.

(60) Provisional application No. 62/093,754, filed on Dec. 18, 2014.

**Publication Classification**

(51) **Int. Cl.**

**C12Q 1/6874** (2018.01)

**C12N 15/10** (2006.01)

**C12Q 1/6825** (2018.01)

(52) **U.S. Cl.**

CPC ..... **C12Q 1/6874** (2013.01); **C12N 15/1058** (2013.01); **C12Q 1/6825** (2013.01)

(57) **ABSTRACT**

A method for preparing a homopolymer recalibration panel includes: extracting, from a set of amplicons used in sequencing-by-synthesis, a set of candidate amplicons satisfying a first set of criteria, wherein the first set of criteria includes amplicons known to belong to high-confidence regions of a reference genome with no variants; and selecting, from the set of candidate amplicons, a reduced set of amplicons satisfying a second set of criteria, wherein the second set of criteria includes amplicons that together comprise at least a minimal threshold number of homopolymers of each homopolymer length between a predetermined minimal homopolymer length and a predetermined maximal homopolymer length for one or more of homopolymer types A, T, C, and G.

**Specification includes a Sequence Listing.**

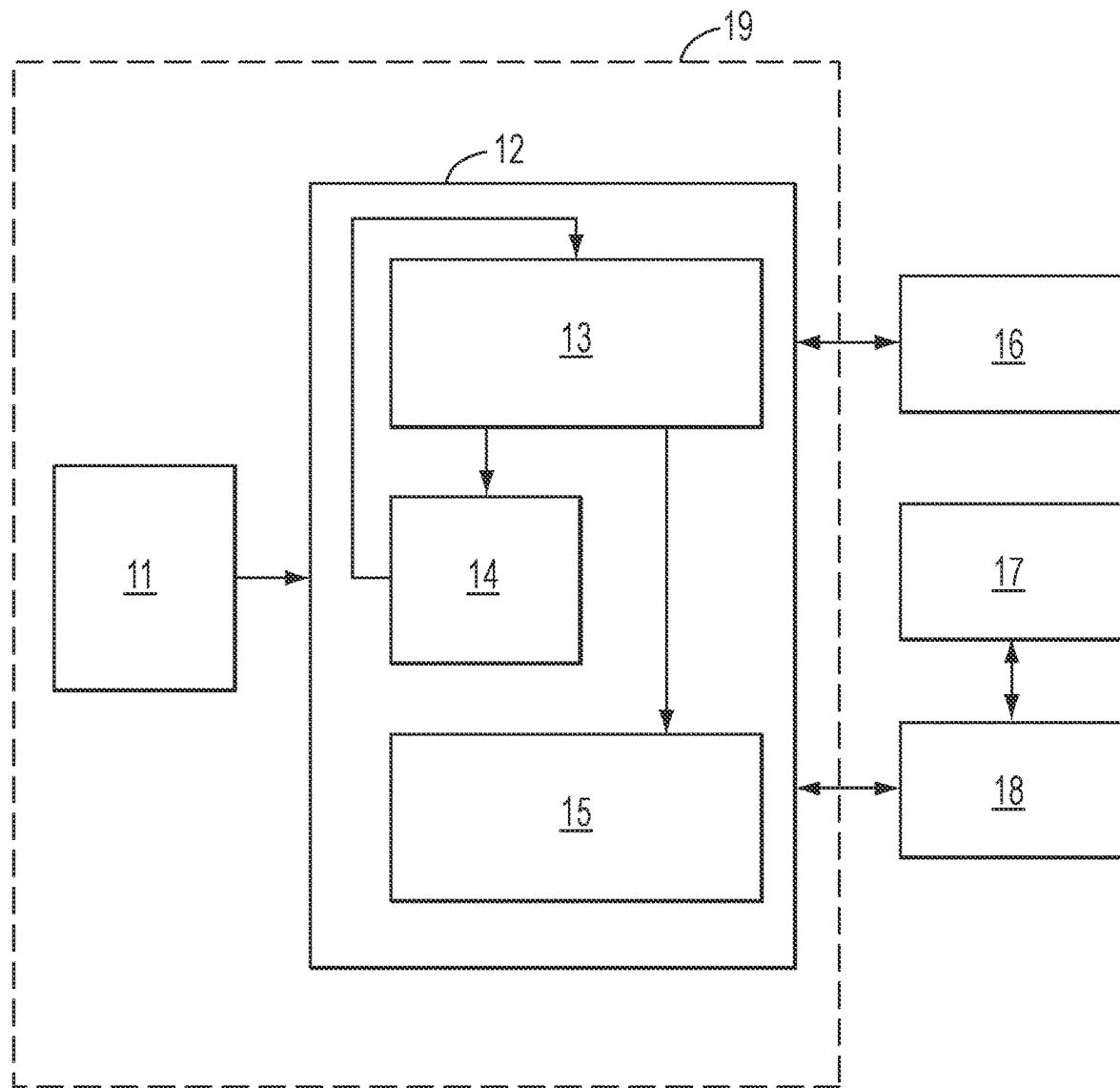


FIG. 1

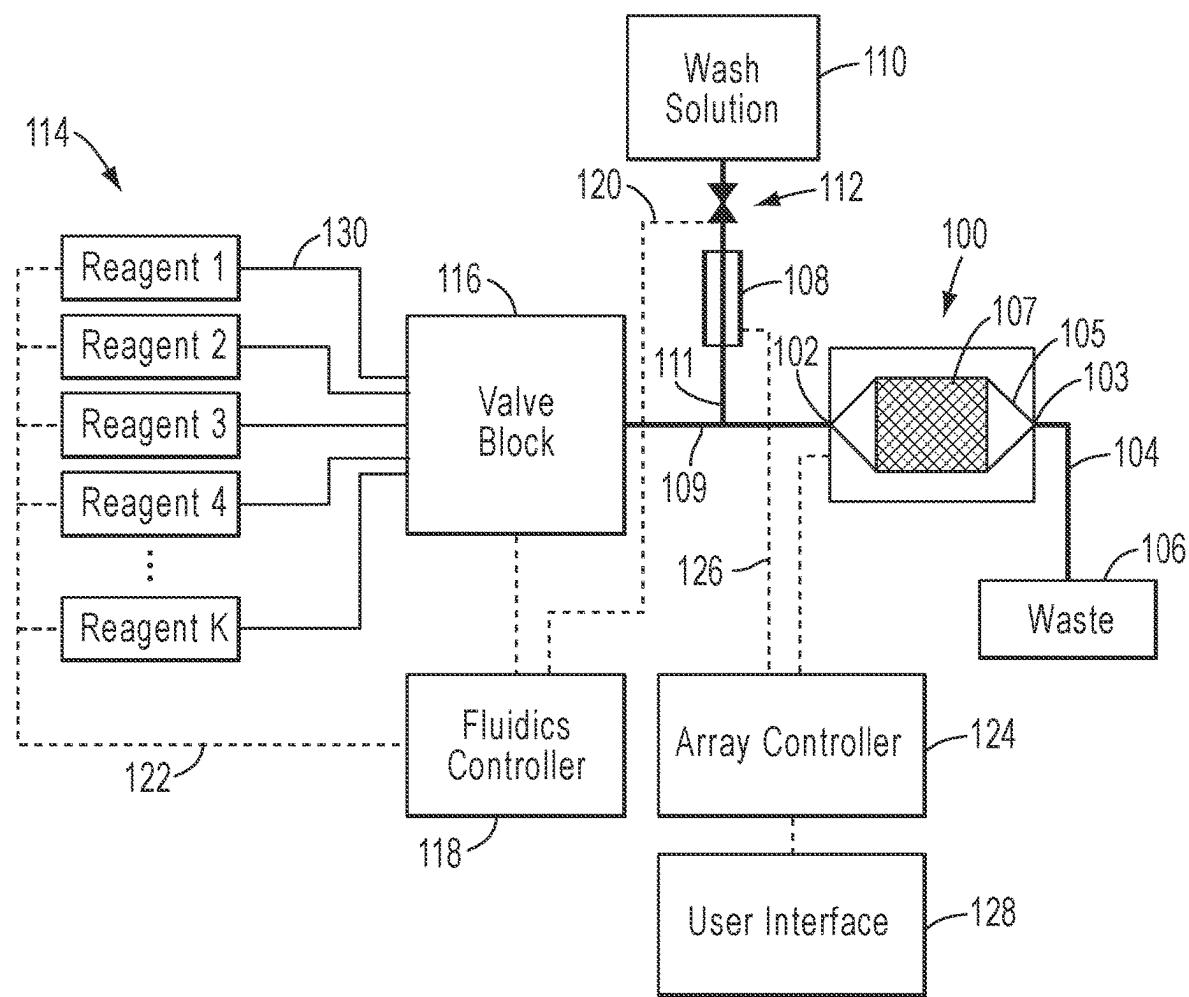


FIG. 2

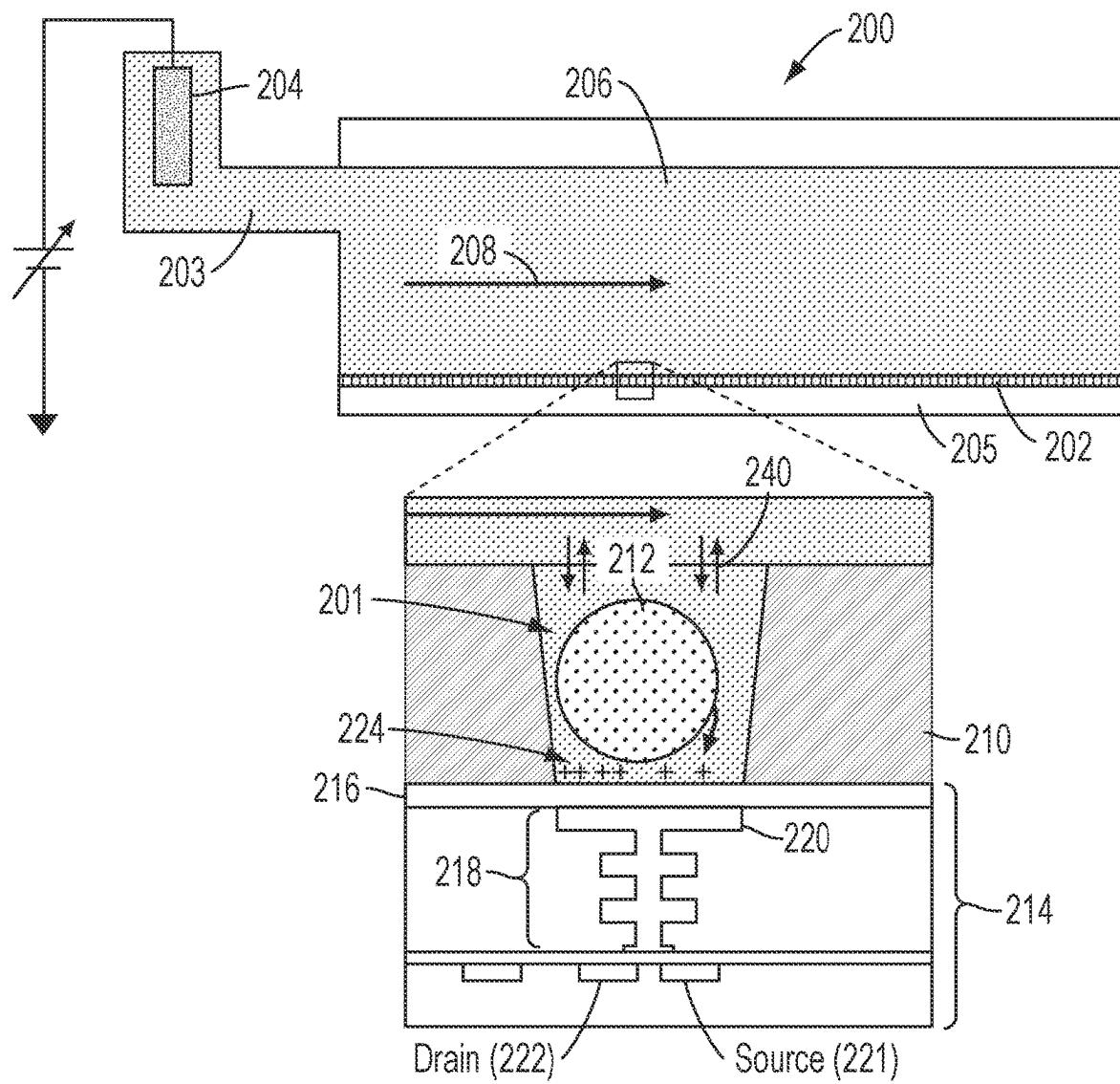


FIG. 3

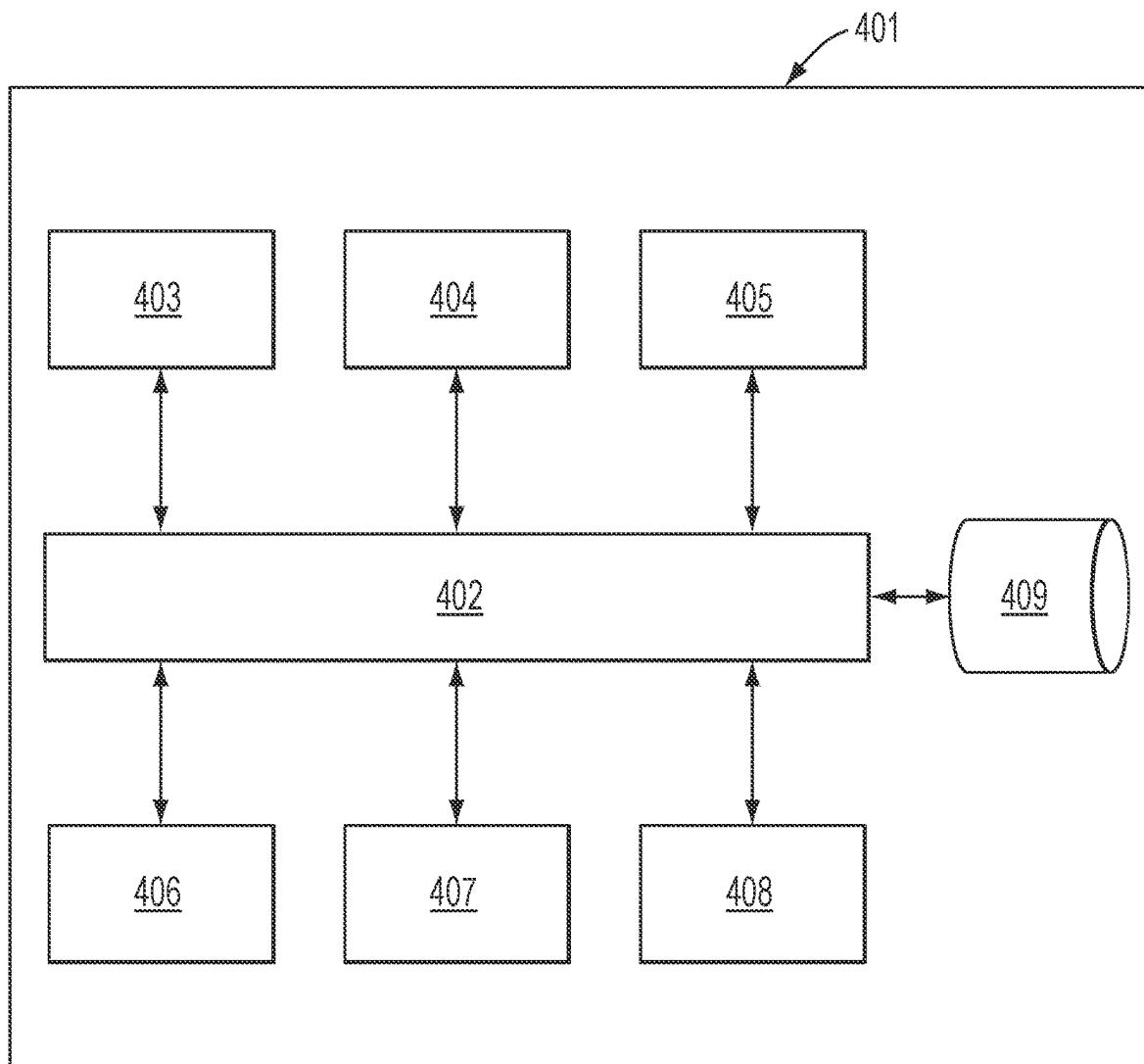


FIG. 4

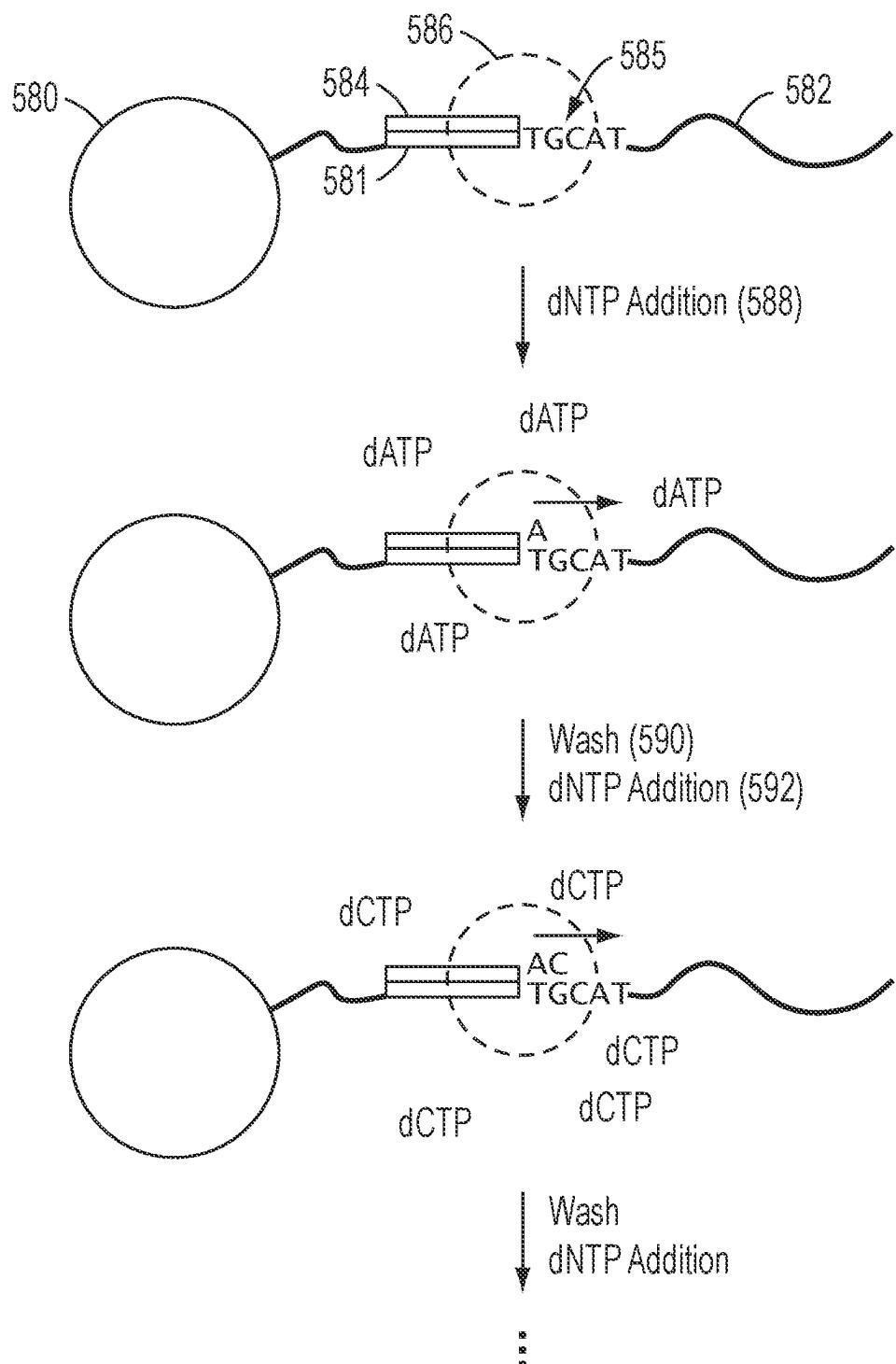


FIG. 5

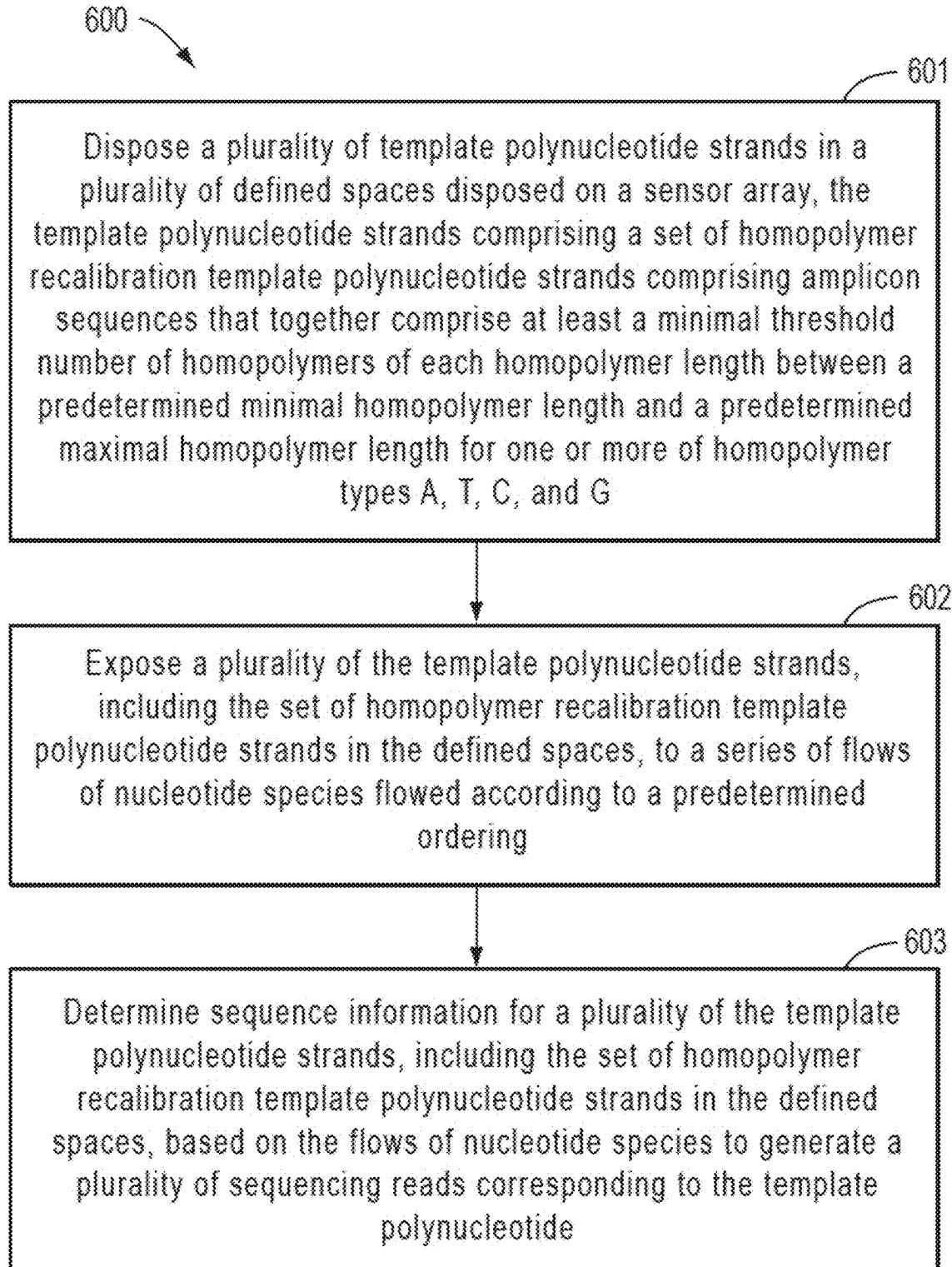


FIG. 6

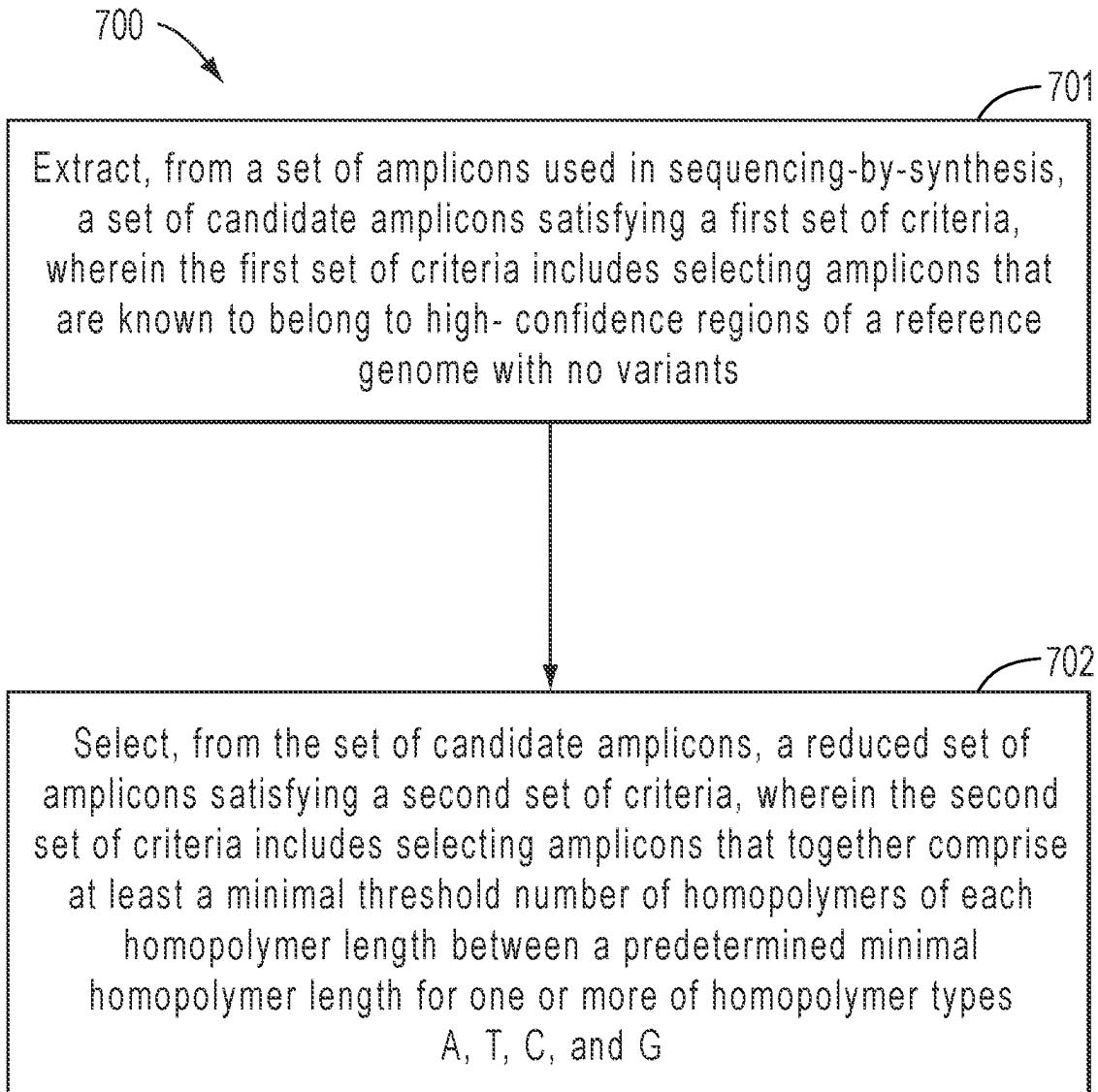


FIG. 7

800 ↗

Homeopolymer composition of selected amplicons													
Subset	Amplicon Count	5A, 5T	6A, 6T	7A, 7T	8A, 8T	9A, 9T	10A, 10T	5C, 5G	6C, 6G	7C, 7G	8C, 8G	9C, 9G	10C, 10G
base	288	65	51	50	49	49	49	66	33	43	37	31	17
backup	59	28	5	3	1	1	1	24	21	14	16	13	3
extra 9GC 10Gc	37	8	4	0	0	0	0	16	1	1	0	7	30
Total	384	101	60	53	50	50	50	106	55	58	53	51	50
Estimation of sequencing quality based on a good 2 Exomes/Pl run:													
Quality Threshold	Amplicon Count	5A, 5T	6A, 6T	7A, 7T	8A, 8T	9A, 9T	10A, 10T	5C, 5G	6C, 6G	7C, 7G	8C, 8G	9C, 9G	10C, 10G
coverage > 10% AND (Fwd OR Rev end-to-end > 5%)	370	100	60	55	50	49	50	97	52	57	52	50	40
coverage > 20% AND (Fwd AND Rev end-to-end > 5%)	284	85	54	49	47	47	48	60	41	39	38	23	17
coverage > 50% AND (Fwd AND Rev end-to-end > 15%)	233	77	44	38	38	37	28	42	35	36	33	18	10
coverage > 50% AND (Fwd AND Rev end-to-end > 50%)	170	57	37	30	34	27	27	20	24	26	23	5	3

Coverage is in % of median amplicon coverage in a good full Exome Pl run.

FIG. 8

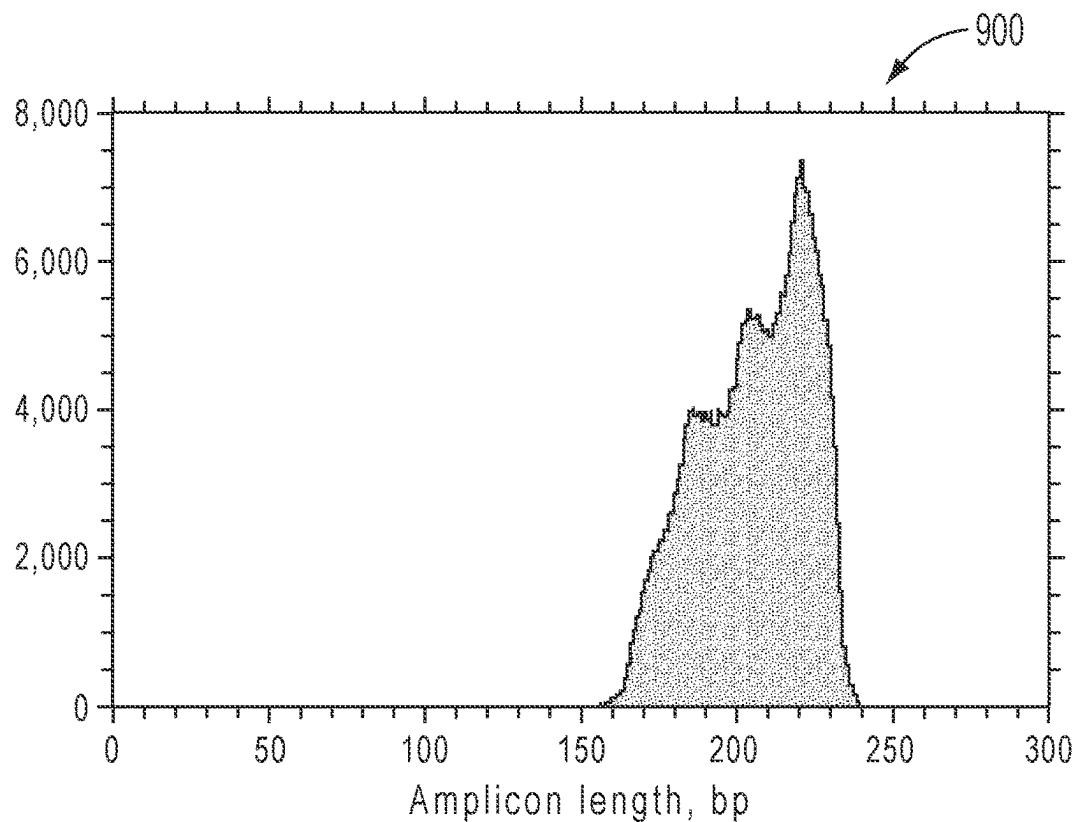


FIG. 9A

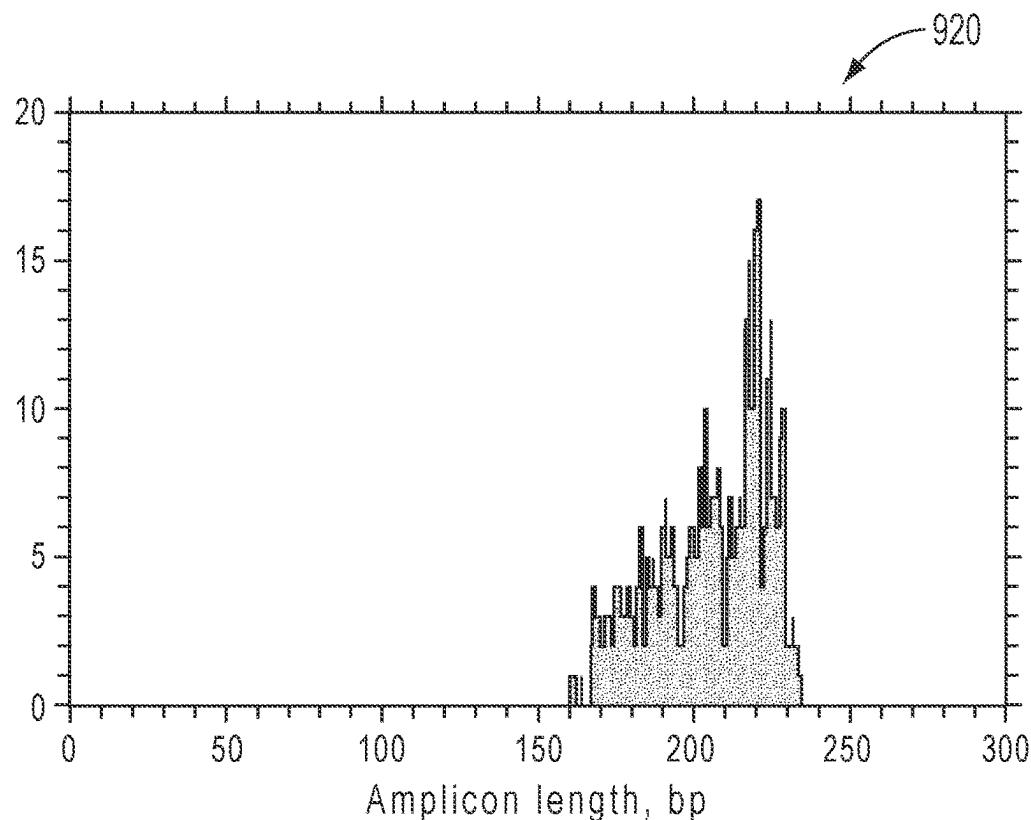


FIG. 9B

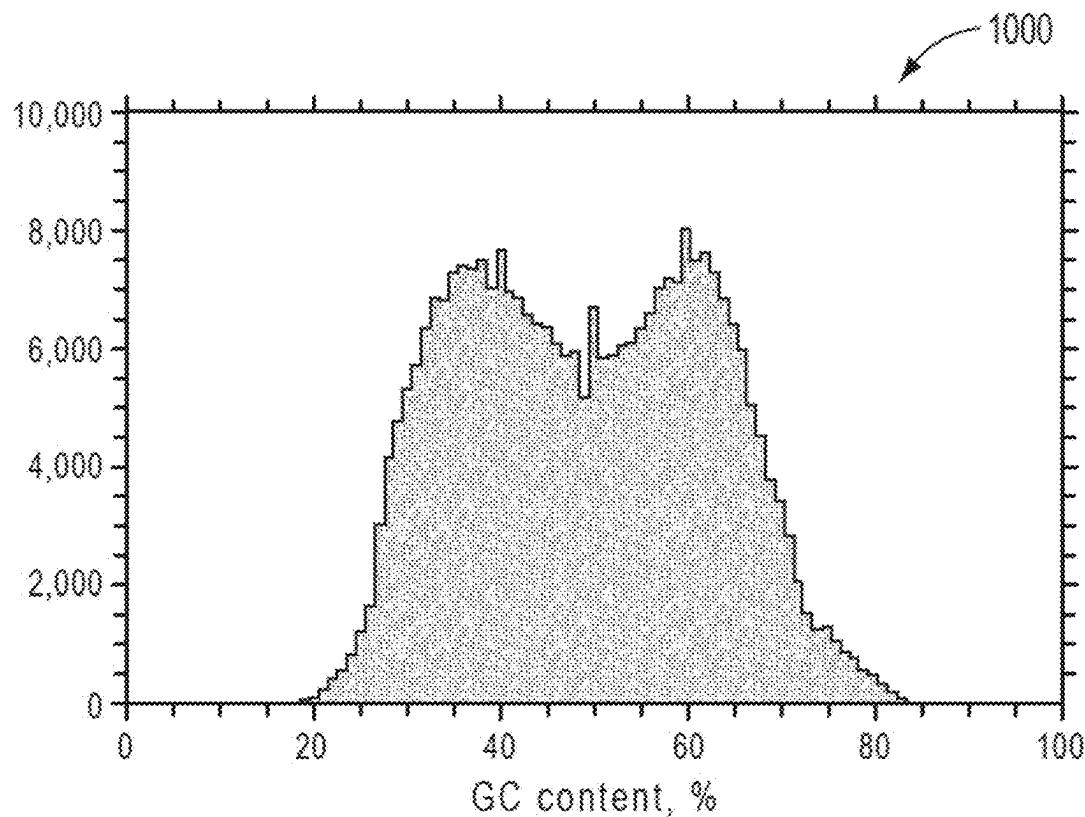


FIG. 10A

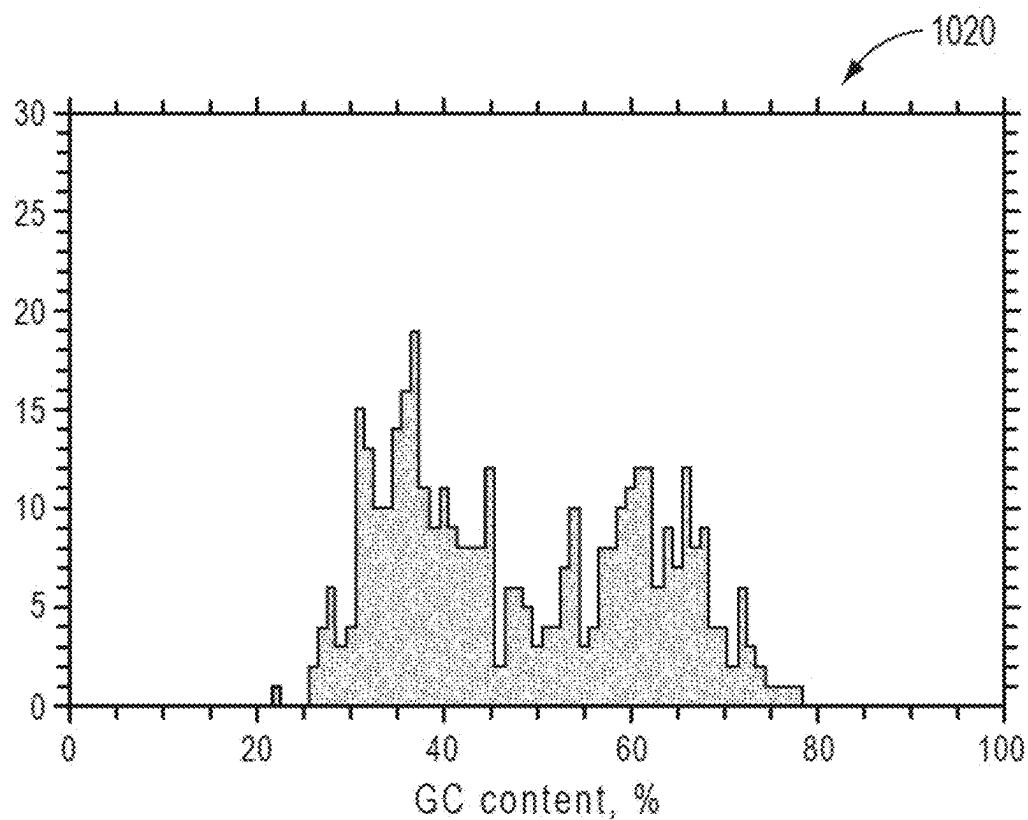


FIG. 10B

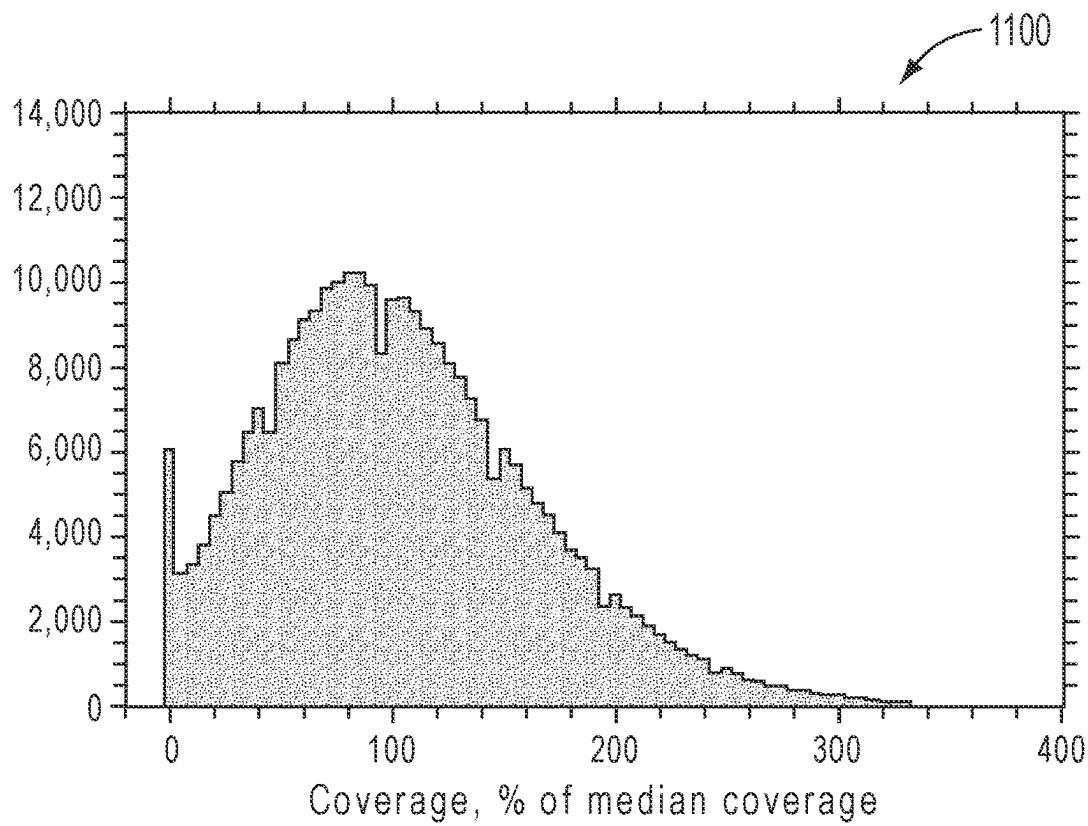


FIG. 11A

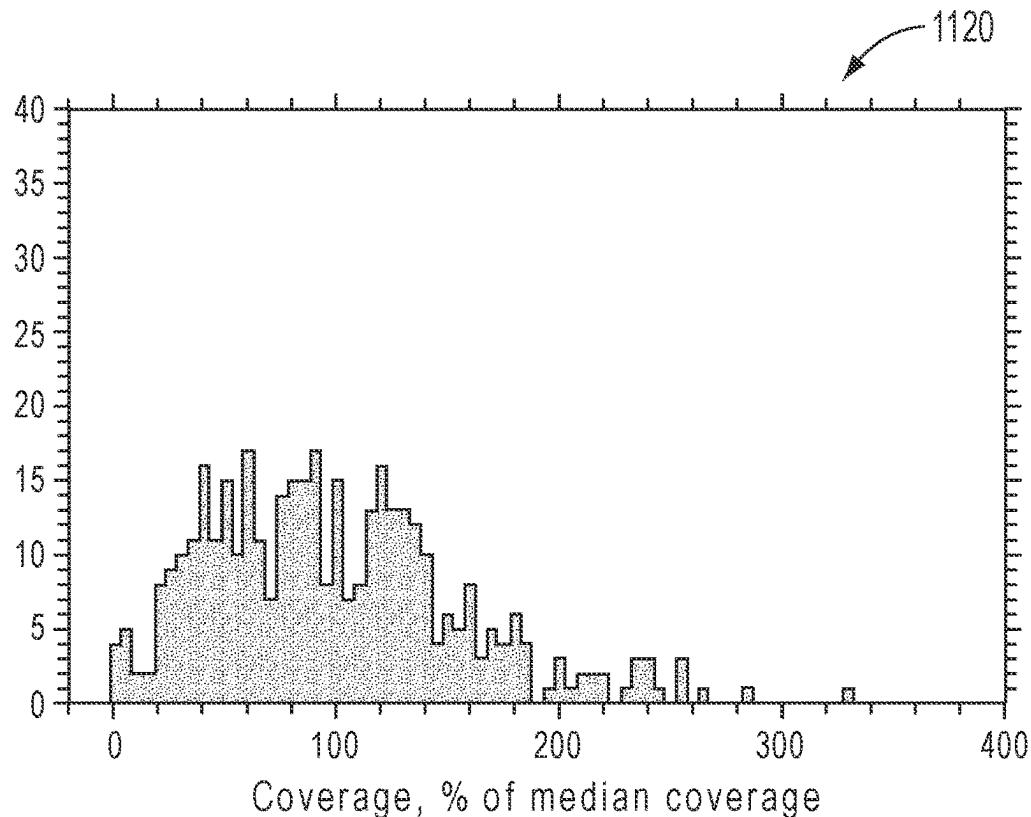


FIG. 11B

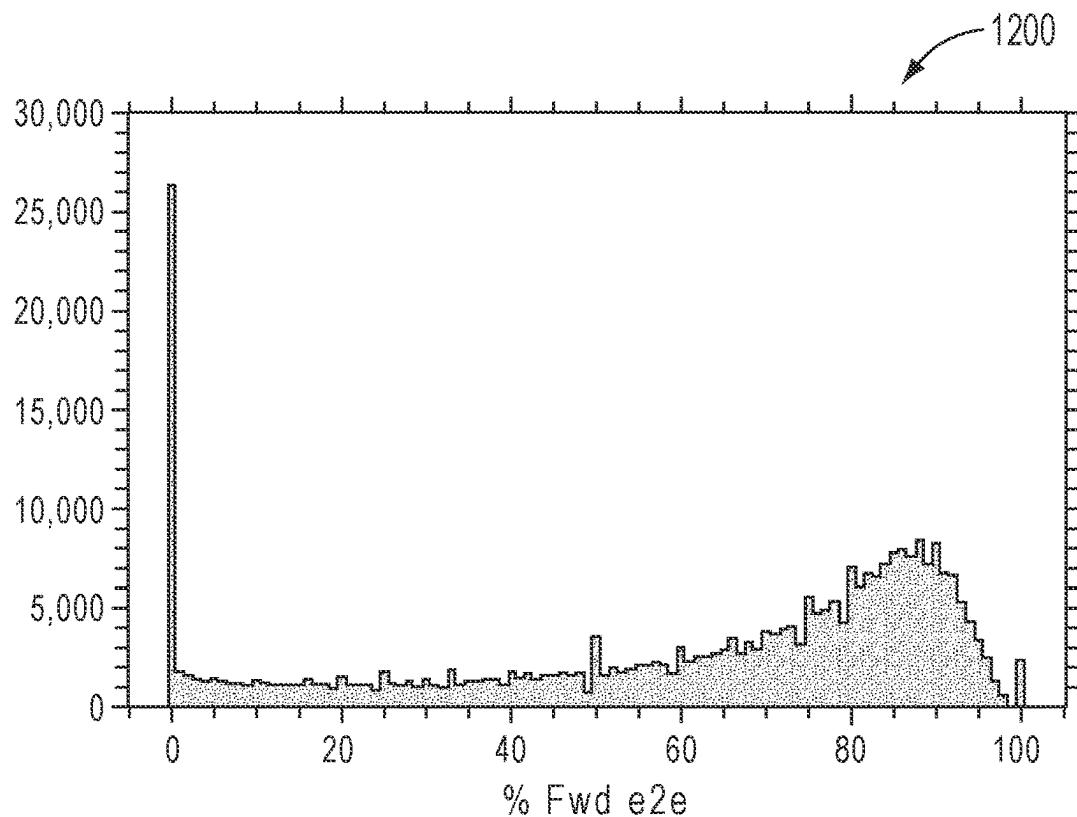


FIG. 12A

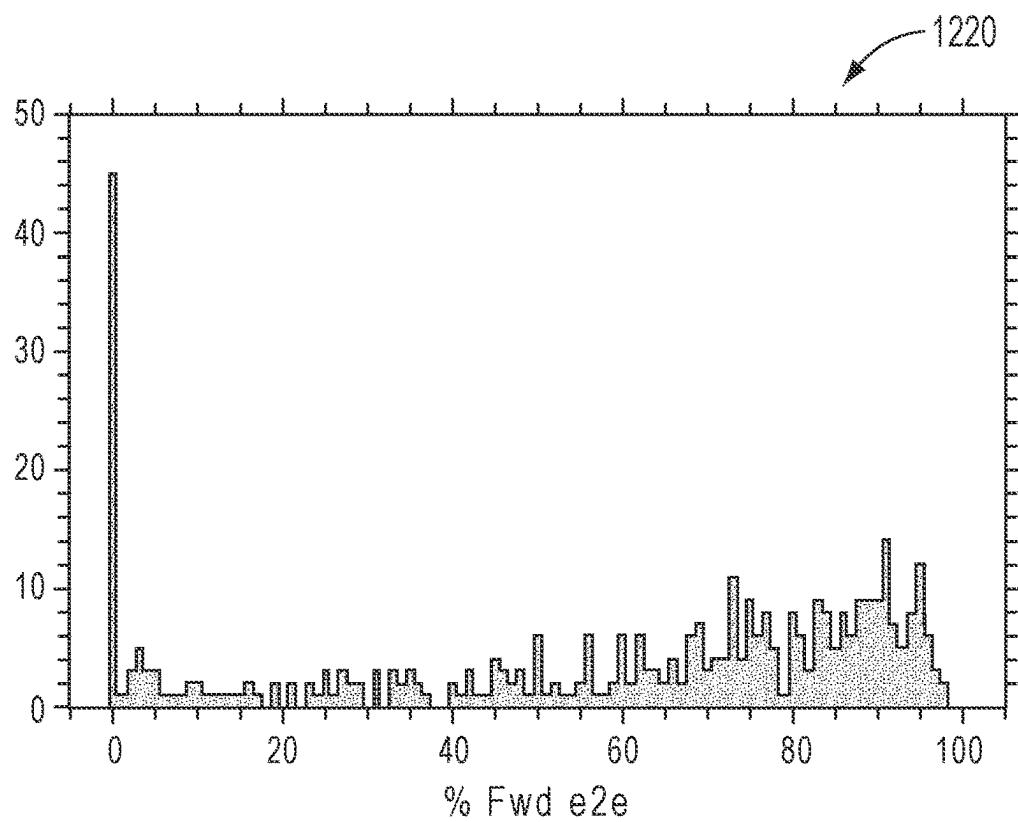


FIG. 12B

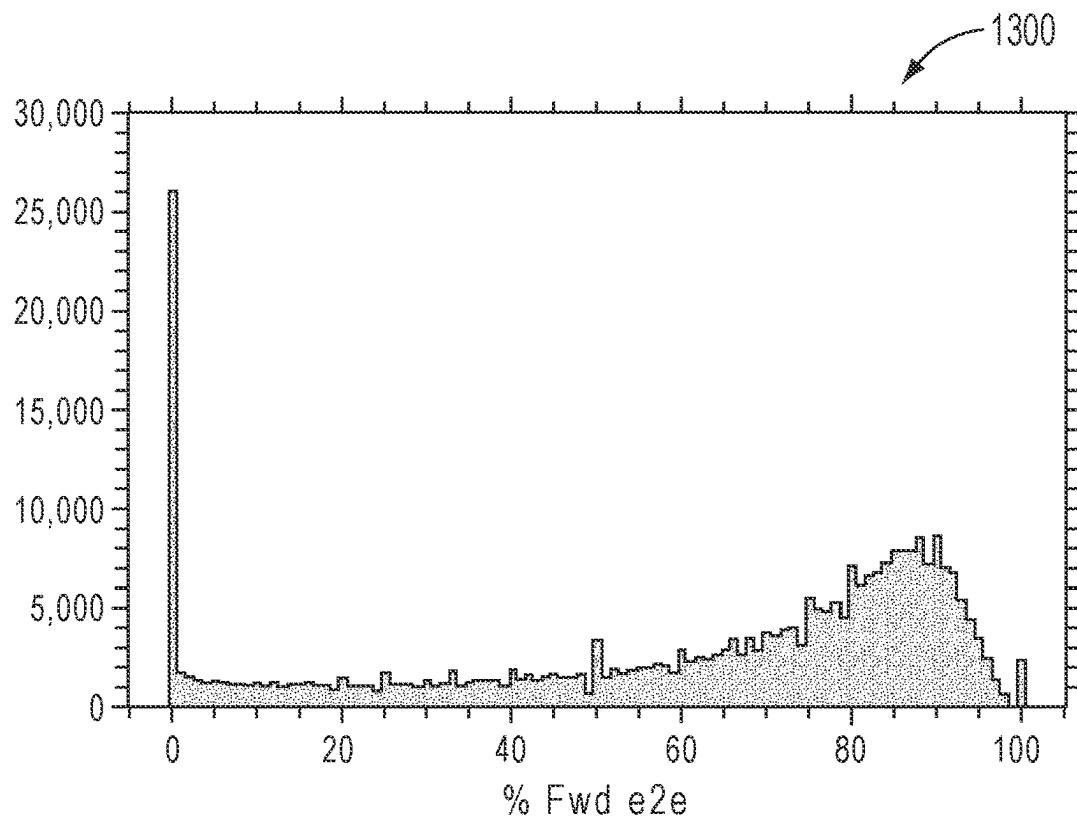


FIG. 13A

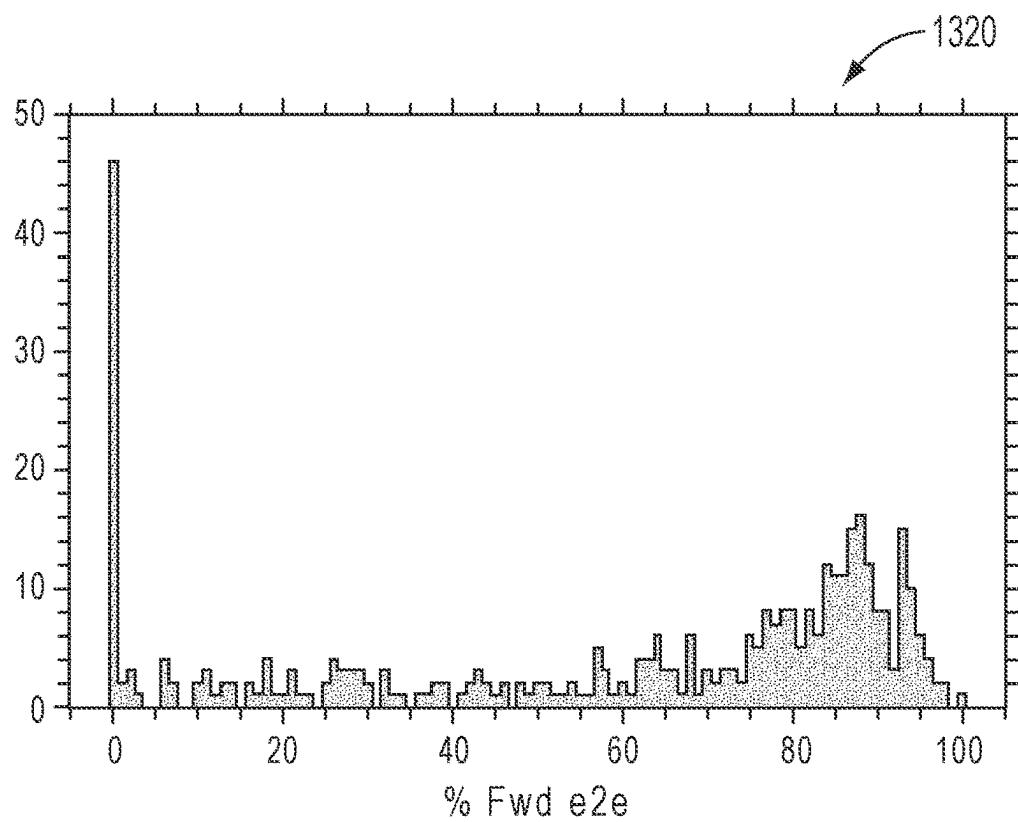


FIG. 13B

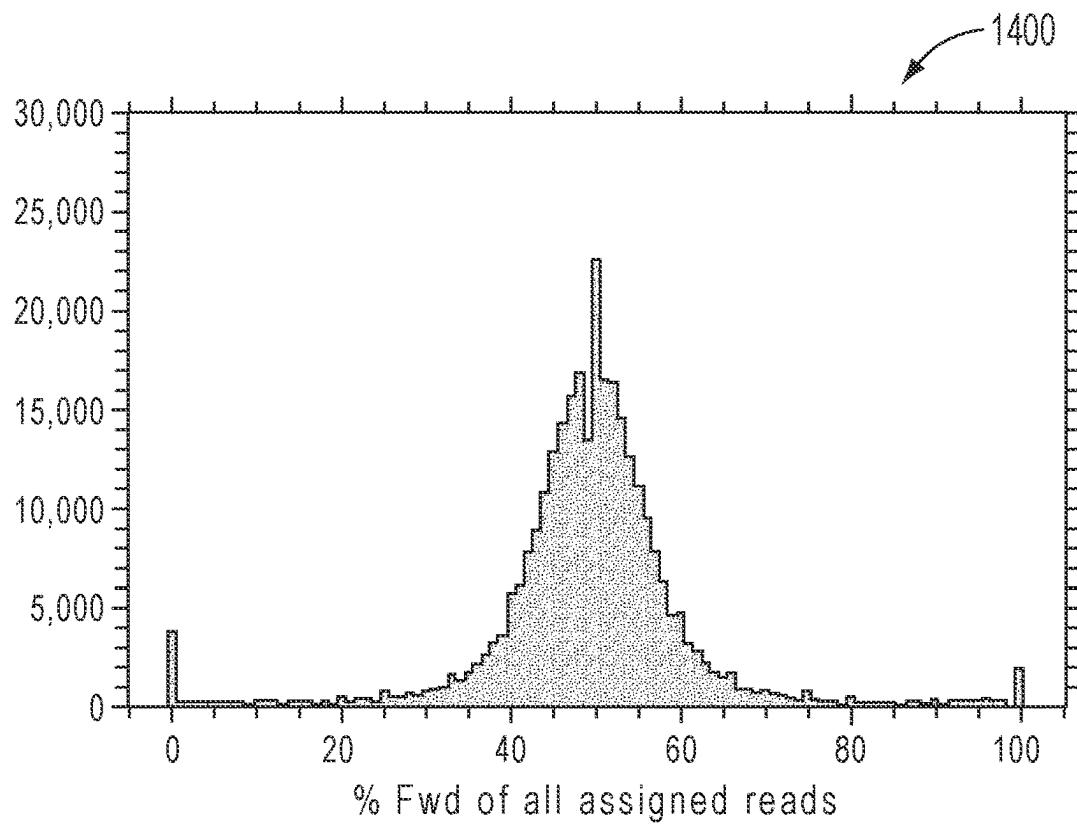


FIG. 14A

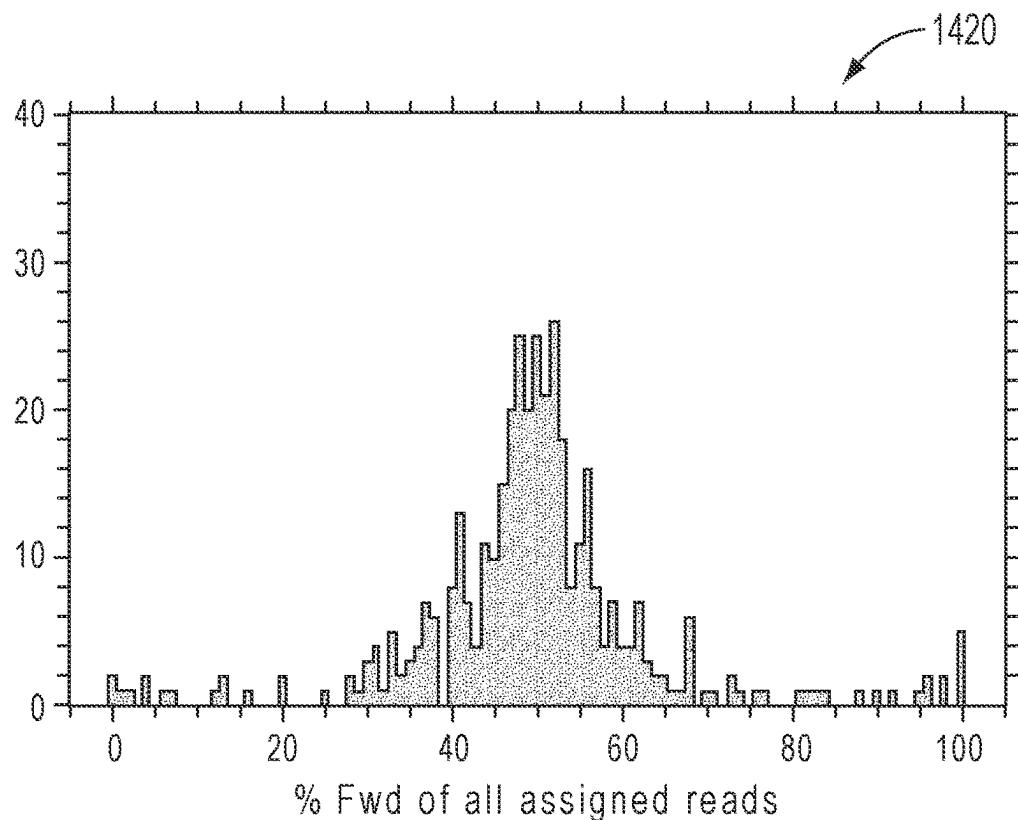


FIG. 14B

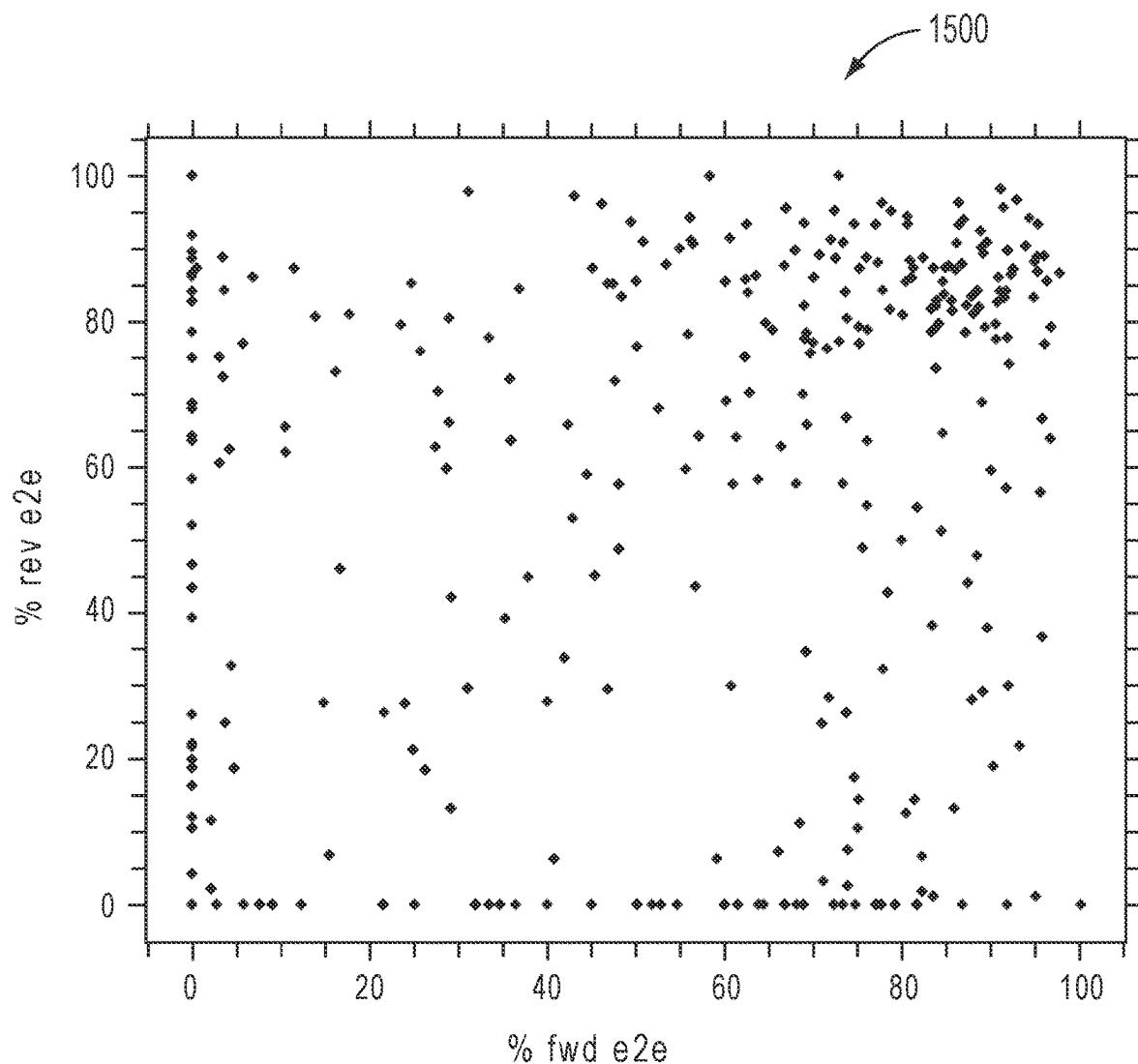


FIG. 15

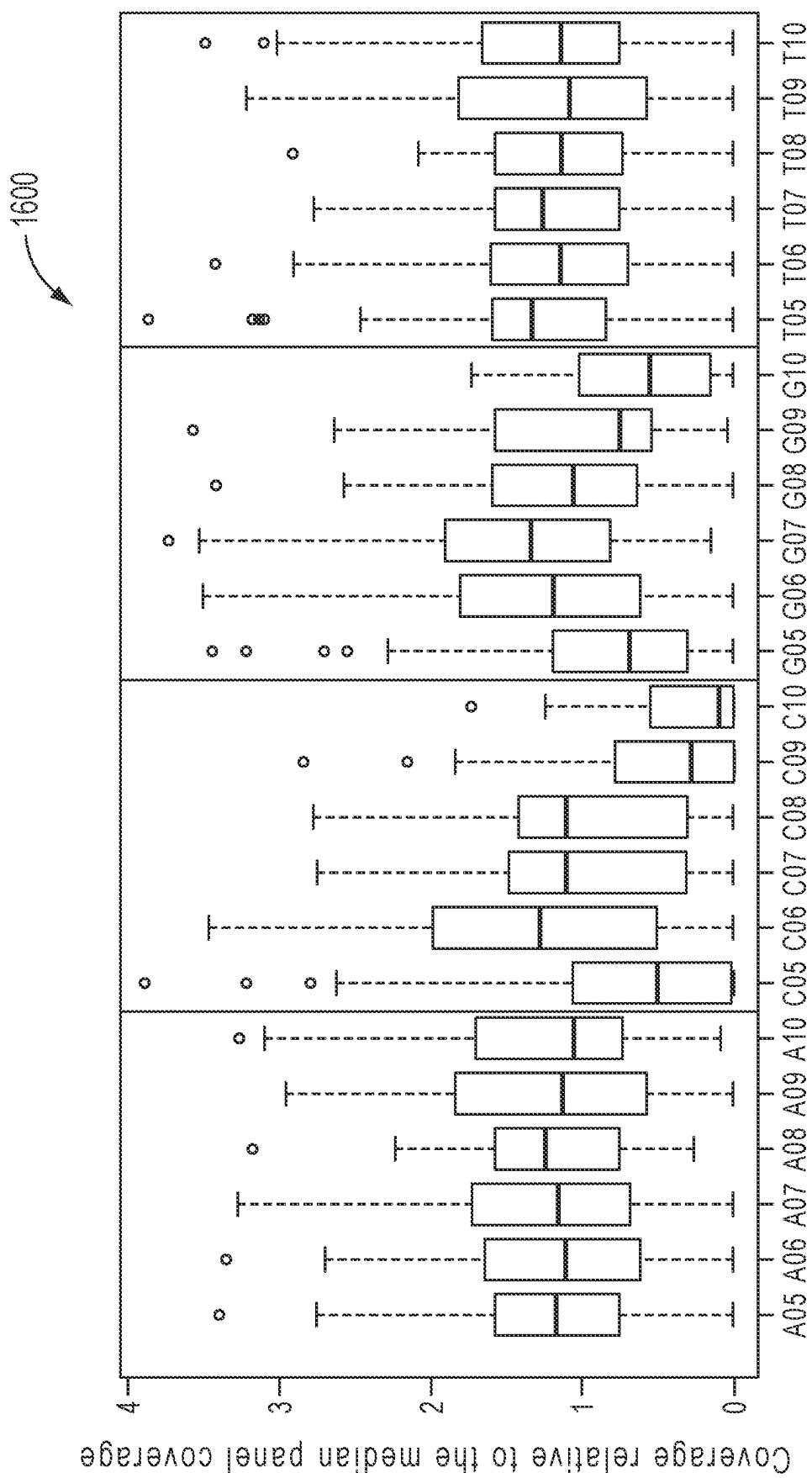


FIG. 16

1700 ↙

no.	chr hg19	insert	start	insert	end	insert	length
1	chr1	1247489	1247712	CCGTCTGGAGGTGCTGCACACAGTGGTCCTCAGGACGCT GTGGGGAGGCTCGGTGAGACCTGCCCTGCCTCCAGGGCC CAAGGGGAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG TGAGGTGGCTGTAGACGGGCAATGCCGTCCTGCTCCCTGC GTGTCATGCAGGTGACGGGAGGGTAGAAGCCAGCTGG TGCTCAGGCCAGACCCAGCTC			224
2	chr1	15772046	15772263	CTCTGGGGGGGGCTGGTGGCTTATGCCCTCCCCGGTCTGG TGCGAGGGGGACTCCGGGACTGAACCTGCCAGTTGGAG AACGGTTCTGGGGGGTTGGCATCGTCAGCTTGGCTCC CGGGGGGGCTGCCAACACCCGCCAAGGAAGCCGGTAGTCTACAC CGGGGTGTCGGCTACATCGACTGGATCAACGAGGGGGTG CTGCCCTCCACAG			218
3	chr1	23761077	23761267	TTGACAGGCCAAATGCAAGGAGGTCTTCAGGTGCTGAAAA CACATCCACACCTCTGAGATTCCCCCCCAGTGCACCTGGCT CCAACCCAGTATTCCCTCTTGAACCTCTTGTCTAACCCCTTAATCC CCCATAGGAAGCTGCCAGAACCCCTGGAGAGATGAGACTGCA AGGGAGAGGCCCTCATGTCTGG			191
4	chr1	27107052	27107279	ACCCTTGAGCCAACTAGTGTGGACATGATGCCGGGGCTGC CCGGCGGCTGCTGCCCTGGCCAAGGTGGACGAGAACACT CAGAGTTTACTCTGACGAAATCACGGGCTGTGGACATCTCGGT ATCACCGTIGATGAACCTCAATTGGTTTACAAGTCATTGTGATG TACTGTTTTGATTGGCCAGTCATGACAGGCCGTGGGACACCTC CCCCCCCCGTGTGTG			228
5	chr1	47034035	47034252	TGCTAACATCCATCCCTGAAAGCAAGGGTGGCTCCGGAA GTATAAACATAACATTCCAATCAGGCCAAAGTACCTTTCT GGAGATTGACACAAATATATTTCCTGGTTCAAGATCACAGATGAGC AATGCCCTGACATGACAAAGAGCAAAATGGTTAACATATGC AGATAATTCAAACATCCACACACAAATTAAATTTCCTCAAAT			218

FIG. 17A

1700

6	chr1	47881774	47882000	CTGATCCAAAGGCAGAAAGGGAGGGCGGAAATGGAGGG TACCGTGAAACAGAGTAAGGGGGCTGGGAGGAAATTAGAG GGGAGGGCCTGGCCAGGCAGCAAGGCCCTGAGCCCTCTCCC CCGGGGGTGGTTGGGGGGGGGGCTGTCATATAAGCGGG TCGGGGCCCTGGTCCCCGGAGCCGGGGCAGGGC TCCCGCAGGCCATGGCGGGGCGCA	227
7	chr1	59004866	59005047	GTTGGATGCTTAATGTTACATTTTCAGTTAGACAGTGAATT AAATCGGAAGAAAACATGGTTCTAGCGAGACTGCAATCCA CAGATGAAGGCTCATTTTCACTTCACTACCCCTGAACAAAAAA AAACTATAAAATATCTGTGGAAAAATATTACTTACCAAACGGGAC TATAAC	182
8	chr1	89434370	89434569	GATTTCAGGGCTATCGATTGCTGCAATCTTTGATAAATTCTTCTTTA CATATGTAGGGGGATATATCTGGAAAGCCTGGCCAAAGATT CACAAACAGAAGGGGTCTGCAGCCAATTGGTAATTCAATCCTA AGATTGAAAAAAATACTAATATTAAATAAGTTCTACAATTATAAAGT TCGGCTTTTATACTTATACTGTC	200
9	chr1	100545857	100546073	CATAGGACAGCAGCTTACTACACTTTGTAAGGCTGATAAGCTCTAA TATTAATAGTTTAAATGACATTGCTATAAAATTGCTAGTGTCTGAT AAGAAGGCTTGATTTAGATTAAAGTGAATTCTATTTGTAAGTGGTT CTAAGGGTAGGGAAAAAAACCAGGTAGTTACTAAGTGATAATT TGGTTAAACAAATGTAAGGTGTCGTTCAAGGAATGATAAC	217
10	chr1	101456160	101456376	CCAAACACAAAGTGTCTCCTCGGTAACTGCTATTAAAAAA GATACTGCTTCAATTCAAGGAAATGCATCATCCATCAACAGT CAAACAAACATTAGTACCAAGTACATTCAAGAAGGCCAA ATATCTAAATCTCTTTGCACATTGGACCAAGTTATCCATTAA TTCATCCAAAGGGATTTCIAAAATCATTTGTATACCTAAAT	217

FIG. 17B

1700

11	chr1	118494978	118495197	CATAAGATAAACAGTGGGATGCAGACAAATTGAAACACATACA GACTCTGGAGGTACCAACATGCCTCTGTGCTTGGTACT TAAGAAATACTTTAAGAATATTACATTACCCACTAAGTAGTGG TCATGAATGAGGTATTCTTATTTCCTTCTCAGGGTC ATCACCGAGGAATATGGTGTGGCTGTAAGCCCCAGTGGAG A	220
12	chr1	149907029	149907220	CAAGAGGTACAGAACAGTGTAAACCATTGGACAGAGTCTCC AGAAGGGACTCTGGAGCCCATCCCACTCCACTCCATCCCCTGCTGT TTCTGAGAACAGAGGAAGGATGAGGAGGGAGTCCAACAAACC CCAGTGAACCCCTTCACAAACCCCCCAACTTACCAAGCC TCTAATCGTCCAAATGGTGAACCAAG GAGCTGAAGGTGAAGAGGAACCGAACACTATACAGAATCTCA ACTAAAAATATCTTACTTGTAGGTCCCCTGGCTTACGGAAAAAAAT ATCTTAGCTTGACCTATGTGCTTTTATAAGCTTCTTTTATTTCTT TATTCTCAGGCTATAGATAATGGGGTTGAAGGAAGGGAGACAA AACTGCAAAGGCCAGAGCAATGGCTATATCCAGAACAAAGA GIAAGGTG	192
13	chr1	158669425	158669648	GAGCTGAAGGTGAAGAGGAACCGAACACTATACAGAATCTCA ACTAAAAATATCTTACTTGTAGGTCCCCTGGCTTACGGAAAAAAAT ATCTTAGCTTGACCTATGTGCTTTTATAAGCTTCTTTTATTTCTT TATTCTCAGGCTATAGATAATGGGGTTGAAGGAAGGGAGACAA AACTGCAAAGGCCAGAGCAATGGCTATATCCAGAACAAAGA GIAAGGTG	224
14	chr1	160302191	160302392	TTTGACCAAGAGATGACCAAGGGTTGCTCTGAACCTGGCAG CTCACCAGAAATATCCAAACGCCACAGCTGGTCCAGGCT GGCTGATACTACCAAGTCTCTGTGGGGTGGAAACTGAGCACA CATCACATAATGGTTGTGCCCTGTAAACACACTGCAAGAAAAAA AAAAGACAAATACCAAAATTAGAGGTGGAAGTCA	202
15	chr1	160589490	160589716	CAACAGGGCAAGGTAGACTGGCCAGGAATACTCTAGGAAACAT GACAGGTTCATCTCAAACAAATATAAGAACTTACACCTGGTTTC TGGAACCTGGGCAATAGATCGTAAGGCCTTTTCCACTGTTG TCTGGTAATGGTTCTGGTAAAGGAAACCATATAATG GTATCTCCCTCTGGTCCAGCAGGAGCTCCTTACTCAGAC CTGGTCTTACTG	227

FIG. 17C

1700 ↗

16	chr1	167190175	167190393	CGGGGGCAGCAGGTAATCAATTACAGCATTAC ATATTCAATTCAACTCAACCCCCGGCTCCCGCTGCC CGCGACTTAGCATTAATTATTAGTACTCAGGATTATTAA CGGGGGAGATGGGGGGGGAGGCATTGAGCCCCGC CGGGCGCTGAGGGGGAAAGGGCTCGGAACGGCCA GGGGAGGG	219
17	chr1	169521791	169522001	TTACTTTGAGTGTGTCGACCTGGCTCTGATAATGGACC CAAATCCCATCTTCAATTGGGATTCACTGTATGTTGG TGAAGGACTCATCTCGTACTGTGTGTACATAACTTATAAT GTTTCCCAATTGGTTGAGAAATTATCCAAATGCTGAGACCTG TATTTCTTAAAGTGAGTAAAAAAAATTGAAGATAAAC 221	211
18	chr1	174362934	174363154	CCAATAAGAAGATGAGATCAATAAAAAATTGAAGATAAAC TGAGCTTAAGCAGTACTATTGTTTAATGCCCCCCACAAACTG TTTCTCTGACCCAAGCTACACTTTATAAACTAGTATAACT TTTCATGGCAACAGCAGTTCAAGGAGTTGTAAGGAAGTGTGAA TGATTATTCTCTGCTTCTCAGGCACAGTAACCTGGTGCACGA CGA	221
19	chr1	185277864	185278024	CAAAATTCAAGCATTGCAATTCAACACAAAATAAGA TTTAAATTACTCTGATACTCTACCTGCTTACTCGATCCATCT CAGCTTTTTGCTGAGAATAAACATCTTACCAATTCTTATC TGCTTTCACCTATTAAAAAAAGT	161
20	chr1	222802588	222802806	AGGCTCTGAGGGAGAAGGACTCCTGGATGAA GAGTTTTTCATCACAAAGGCAATGCAAGGGCACAGAGGTTAGGA CAGACAGACAAACTGACAGCACAGGAGGACCTTCCCT TCTAAAGTAGAAGAGGATGATTATCCCTCTGAAGAACACTAGA GGATGAAAACGCTATAATGCAAAACGGCTAAAGAAAAAC CCTGGG	219

FIG. 17D

1700

21	chr1	225155064	225155230	CCATTTGTTAATTAGTTAATTAAATTAAATTGTTTATTATTTTTAATCTCAGAACCTTGGAAAGATGATGGAGAATTGTTTATTGCCCTTCCTCGAAAAGTCCTAAATCCCTTACA	167
22	chr10	3201061	3201280	ATCCATATGATCTTCAGGTAGTATCGGCTCA CCCCAACATCAGGAGAAAAGTCTG GCCAAGGCCAGATTCAAT CAAGGCTTTGTAAAAGGGAGAATTGGGGCAGGAAGTCAAGAG TGAAAGACAGAACCTTAATGTGAAGGGCTCAAATGTGTCGGTG ATGCTGCATTGAAAAAAAGGAACGGGGAGGTAAAGAAAAGA GTAAGG	220
23	chr10	27431202	27431405	TAGGTCCCTGTGAATTGCAAATGAAATTGTTAGAATTACCAATT AAGAAAAGACAAATCAGACTATAGTAATAATTATTTTTCCA CAATTATGGTTGTGCTTCATACTGGCCAGTACTGAAGATCT GAACAAATGCTTGAAGAGCTCTTGAATGATGTCGATAACAGC AAGAGGAACGTTAATGTACTAAATAATAT	204
24	chr10	32573447	32573660	ACATACAAACTCCTATAGAACAGGATTCAAATAATTACTTAGA CCAAGATTTTTTAACTGTTAAGAGAAAAGGACTTTAAC CATAAATGTAACAGACCACATAGTCGAACAAATGGAAATGGT TGACACTGACATTCTATAGAAACGTCATCATTAAACAGTCAC ATACCAAAGTGTGCTGAACCACTGCTACTTTTACCTTGC 214	
25	chr10	35930224	35930443	CCTTACACAGCGGCCAGGTGATCTCTGGCATGGCAGCTGCAGCA TTGGCCGAGGGCGCCGAGCGCCGCTAGAGCGSCTGCAGCA GCGCCAAGGGGGCAGCGAGGTCACTTCCAACAGGTAA CCCCCACTCCATGCTGTGCCCTCGCCCCGGGGGGGGGG TCAGGGGGCGCAGGGGGGGGGGGGGGGGGGGGGGG CGAGAGGAGGCCAGACG	220

FIG. 17E

1700

26	chr10	64022411	64022609	TCTCTTAATTCCTGAGCTCCGATGTTAGGCCATTAGTCGAGC ATTGCACACCATGAGATTCTTAACACTGCATGAAACCTGATCTT TCTGAGTGCTCAGAGAAAGGAGTTCCATATTCCCTCTCGCAT TCGAATTCTAAGTCTATTTTCTTGAATGGTGCAGTCCTAAAAA AAAAAAATGCATCTTGAAAATAA	199
27	chr10	65024499	65024666	TTTCTTCAACCAGAGGGTTGAAAGTCTGCCAAGAAAAAAA ATTCTTAATGATAATTCTGGTAGTATCCTGAAATTATGAAGG AACACACACCAATAACAATGTTCAATGTCATTACACTACAGAG TCTTAGGGTACCTAATAATTTCATAATTATAAA	168
28	chr10	70243174	70243364	AACCATAATTCTTGTAGGTGAAAAACTGCTTCATAA GTCGTATGTTGTAAGCCTGCTTGAGAGGGAAATACAGCG AATGTAATTAAAGATAAAACCAGATAAGAGTCCTTTCGAATTC CATGGTGTCCATAAGACATACTTCATAAGTATCCGGATGGTACTG AAAGACAAATGATTAA	191
29	chr10	76861568	76861785	ACCCAGCATGCTCAGGGACACACATCCTACCCCCCAAATTAA GGAGCACTCACCCACCTGTGGAGATCTGGCCAGGCTGT GCACTTCTCTGACCGGGCAGTCGCTGCTGGCTTCAGCCCG CCTGTCTCCGACACTTGATGCTGTTCTATTTTGATTGAG TTAATTATAAGCATGCTGTACTCCCCCTTCCAGCCTCCC CTGCC	218
30	chr10	91477205	91477387	ATIACAGATTGAAGATTCTGAAATGTCGTTGAATTCGAGTCA GTGAGTAAGTTGAATATTCTAAATTAAATTGTAATTGTTT TGAAGAACCTTTAAAGTACCTTTTTCTTAATTCAAGATTATCTT TATGTGATCTTGCTGGTTAGAACGAACTATGAAGACACAGAA TGAA	183

FIG. 17F

1700

31	chr10	96100091	96100273	ATTCAAGAACATCACCTTTGAAATGACAACACCAATTAGGT ATGTTATAGCATTCAAAGGTAAGGCAAACAAAACCTGGATTAGA AAVATCATACCTTGTCCAGAAAGAATATGAAAAGCAGTCGG CACAGTGAAAGACTTCTTGATAGCTTAGGCTTAGGTTAAAAAAA AGGGGG	183
32	chr10	100186886	100187076	CCTGGGGCGAAACCCCTAGAGGCCCAAAATACTCACTGCC CATCTCAGATCAGCTGGGGCCAGGCAGGGACAGGAGAAC TCACCTGAAGGGCATCCATGGGGGGTCCCCCTCCAGC CAGATGGTGTACCTGCAGGCCACAGGTATAACATAACGATG TAAGTACCATGAGATGTAAGSAAATG AATACAAGAAATAACAGATGTTGGTAGCTTAGCATTAAA CAATTCTACAAATTTCACCTTGAAAGGTATAAGCTGGTAACAA CGCTTGCTCCAGGGAAAGTAGAGCAGACACCAGGTTCTGACT TTGGCTTCTAAATAATTCCCTGGAGAAAAACACAACATGATCA TTTACCC	191
33	chr10	113924221	113924400	AATACAAGAAATAACAGATGTTGGTAGCTTAGCATTAAA CAATTCTACAAATTTCACCTTGAAAGGTATAAGCTGGTAACAA CGCTTGCTCCAGGGAAAGTAGAGCAGACACCAGGTTCTGACT TTGGCTTCTAAATAATTCCCTGGAGAAAAACACAACATGATCA TTTACCC	180
34	chr10	114053473	114053693	CCAGTGTCAATTCCCAAGGATGCCAAGTTCTCCATCAA GAAAGAAGCTCCCTTGTCTGGGCAATCGAAATCGGTTC AGATCTGTTGCAGGGAGTGGAAGGCCAAGGGTTAACGATTAG GTAAGTACATTCCCTCCAAGTTATATGTTAAATGCCAAGAAA AAAAGACTAGCTGACATACCTTTATCATCAGATGTTGTAGG CATAA	221
35	chr10	116919778	116919949	TTGTTTCAATTAAATAGGTCTGATTGTTCTGAATGT TCCTCTACTGAGCTTACTGGATTCTGCCAAACGTTAACCC TTCACTGCCCTCTGTAGGTGGGCTCACATAAGCAGTTTAC ACGGGAAATTATGTGGGTGATTGGGATAACTTTAAC	172

FIG. 17G

36	chr11	194324	194548	CCCCCCCCCACTGAGGGCCTTGCCTGATGGAGCTGTG GCTCTCCACTTAATTGAGCACAGCCCTCTTAGTCACATGTGTC TGCCTCCAGGTGCAAGTGCTGGCAGTCAGGACGGTGCT AAGTGGACCTCAGACATGGCTCAGCCATAGGACCTGCCACAC AAGCAGCGTGGACACAACGCCACTACCACCTCCCACATGG AAATGTTATCCTCAAAC	225
37	chr11	821740	821958	GGTCCCTGCCCTGCTGATAGCCATGAGGCATGCCAGTGGGGCCT GGGCATCTCCCTGACCCGGCTGAGCAGGGCGAGAACATGTCT TATATCCCACTTCAACTTCAACTCCAAGGACGAGCTCATCCAGGTGGG GCCTGGTGGAGGCCATGCTGGTGGGGGGGGGGGGGGGGGGGGGGGG GGGAACCTCAAGGCCCTGCTCATTCTCTCCACTCTGTCCCC TGCCCCCTGAAGG	219
38	chr11	836364	836555	GGGGGCACTGTCGTATGGTGAACGGGGGCTTGGGCTGCT GGGCCACCTTCAAGGAGCGCTGGGAACCTGCTGCCCTGGTC AGGAGGGGGCAGGGGCCACGGGGTGGGGTGGCAGATGG GCCAAGGGAGGTGTCGCTGCAGGGCTTGAACCTGTGCCTTG CTGTGCTCACCTGGGGGGGGGTCAACGG	192
39	chr11	1090308	1090506	GGACTTGTAGCCATGTGCTTGCAGGGTGTACCAACT CCTCCCAAGTCTGCTGGGGGGAGGAAGGTAAGGCTGCC TCTGCTGCCAGCCCTGGGTGGGGGGCCATCCTGGGAA GCCTGTGGGCCCTGGATCGGTGGGGTGTGGTCTCCTC CTGGGCTCTGCCCTTGGCCCCCAGCTC	199
40	chr11	4622797	4623013	TGTGCACCGAGTCTGTGCTCCCTACTTCTGGCTTAGTT AGGCCTCACTATCCCTACAGGGCTGCTTAGGGCTCAAGG GTGAGAACATAACCTGTGTTAACACTTCCCTGAATACTCCTGGAGG AGAAAAAAAATGCAAGCATCTGAGTAAGTGCTATTCCCTCAGAG GGAACAGACATCGAACTGGCAGGGAGTTCTAGGGGAAG GGAGA	217

EIGHTEEN

41	chr11	15133990	15134214	GAGTCGCTGAAGCCAAGGCTGGAGGCCATGAGACGGCCCC GGCAATGGAGGGCGCCAGCGAAGGTCCAGGTGGCTGGG GTCTATGGGAGTCCAGGAGTCCAGGAGCTGTGCTGTGCC GTATGTACTGGAAAGGTGGATGTGAGTGTGTGTGT GTGTGTGTGTGTGTGTGTGTGTGTGTGTGGAAAGTGG GGGGGGGGCAATTATGCCA	225
42	chr11	18317499	18317702	GAAATAAGATCAGTAATGGAAATTAAATGACAAGACTAAATTGGTT TACCTTCTTTGACAGCCTGAAGAGACAAGAGGAGATGGA GAACCGAAATGGTAATGGAATGCCGAACCGGGAGAAAAGTTTCAT TCTTAACTGTCTCAAACATCACTGGAGCATGAGAAACATTGTCT AATGAAATGGAAATCAGGAAAAAAAGAAACAAA	204
43	chr11	26742851	26743061	CTTCAACTACCAGGGTGTATCATCAAGAGGGATGAGAAAAAAA ATTAGAGAAGAATGGGAAATCATCTTACCTCATATAAGTGCCTGGT GATACCAGATCTGTAGAAACACAGGGAGAAAGAGCTCTGATGTT AAGAGGATGACAATAGGTAAAGCAATGAAGAAGACTAGGAAAG GATGCCCAAAAGCGGTAGACTTCAGAAGGGGCCAGG	211
44	chr11	28057886	28057860	GTTTAGAAATAACCTCAACTAATAGTTGAAAGTCATGTTAAA AAAAAGAAATTGGCAAATATTATTTTGACCGTTGAAATG TCTTTGTTATCATTTGGTAGTGTGATTCTTGTGGGTAATTACAC TTCGAGCTCTTGTGTCATCATATAGTAATGTAGAGT	175
45	chr11	47310440	47310657	TCTGGCTTACTGCTCCCTCGGAGGGTAGAGGGGGTTGGGGTAT TGTGGTAGCACCTACTAATTGTGTATTGTGTCCTCAGTATTGG GCTAAATACTCTAAATGGAGGATTGTACTGAAGGCCGGGG AGTGGTGAAGGTGGGCTTGCCTGTGAGCTGTGATGATCTT TTTTTTCTCTAGCATCTTGTGCCTGCCAGGGCATTCC TTCT	218

EIG. 171

46	chr11	62380934	62381123	GTCCTGCCAGGCTGGCAGCGGGCTGGCTGGCTC TGGGCACAGGACTAGTGGGTAGGAGCCAGCCCCCAAGT CTGAATGCAGCTCATACCTCCCTGGCAGGGTCTGGC CCGCTGCTGGTGGCTGGCACGGCTGGTGGGGGCTCC TGGTCGTTGGCTGGGCTAGCCCTGG	190
47	chr11	64507564	64507771	AGGGACATCTGGCTTTCACCTGGTAGGTAGGCTGGGGCA GTTGACCTGGGCTTGGGACATGGCATGACTCGTGCG CCCCCACTCACACAGACACACCATTACGCTTACTGAGAACCT ACATAATGGGAGGGACTGTCTAGGCCACGAATGTTCCCGAG	208
48	chr11	70172306	70172527	TATCTTACTTATTCTCTTATGTAAGGTATTACACTCATATTC AATTTTTGCTTCAGCTTAAGTCAATGATTCTTAAGAACAGAAATAATCA GAAAAAAACTCTAACAGATGGAGTGGCTGGACATAAACCATGAA CAAGAAAATACACCAAGGACAGGTGGAAAGGCAAGTCTGTAG GCCTTGTCTTTATTCTGTGTTGATTCTACAGTATGTGT C	222
49	chr11	73806157	73806326	CC CAGTAAGACTAGAAATATTCAACAAACACAAAATTCACTCAGA AAAAAAAACCTCTTAAGTATAATCTATAAAAGAACCATATAAGA ATAAGAAAGGCCATAATATCTAAACCTACCCATTTCAGGAAACTCA GGTCCCTTCAGCACACTGGGAGTGGTGAACTG	170
50	chr11	111941885	111942051	ACTGGAAATTCTCCCTCAATGCATTTCATGGCCATCTGAATTAAAC TGATTTTTTCACTGGTCCTTCTGCTGCATGAAGAACATC AGGATTGAGGCACACATCATGACTGTGTAAGCATCTGTGT GTAATTGAAAGGGATGAAATGAAGAAAAATTGTGTTG	167
51	chr11	112042394	112042595	CCATTAAATGAAATAGTTAACAAATTATTGTGTAATGAAGAACAA CATAAATGAAATGTTGTTTATCTACTTAGTTAACATACAACTTTT TTGCTTATTAGTGAAGAGCAGCAGTAGATGCATCTACATTGAT GAGATAGATGAACCTCTCAAGAAGCCAATGCTATTGAAACTT TCTAATACAAAAAGAGGTTCC	202

FIG. 17J

1700 ↘

52	chr11	114398481	114398688	CAGGAGAACTACATTATTAGGTTAATGCTGATAAAAGTTCC TCTGAGCATCACCTAACTCAAGCAGGCCAACTTACTGTAC AATTAGTGACATCAATGTTACGATCCTTCATCATTCAACTC CCACTTTGGACCTTCAAAAAAAATAAGAACATAATTAGCT GCTCTCTGCCATGTAAACACAGGTGAATCATC	208
53	chr12	416805	417021	AACACTGCAGGGCCTCTCCTTCAAGGCAACCGTACGGGCAACT TCTGAAGGGATAACAGGAGTGACAGAAATAGTCTCTAGCCTGG GCCTTCGAGACCGATACAAAGAGGGCAAAGGAATTCTACTTC TTTAGCTTGCCAGCTGGATCCTTTTTGGAAACTTGATTAG GAAGAGGAACACAGCTGTATGGAAACCAGCTTGCAGAGCT CACA	217
54	chr12	1909462	1909680	TGAGGACCTTGCAGGCCTCTGCTCAAGGACACCCCTCCTAAGC CACCCCTCGCATGCAGGCCCTCCACAGCCCCGGTGTCTACCCC ACTCACAGCACCAGCTCCTGCAGCAGGCCACCTGGTCGCCGT CAAGAAGGCAGAAATTGGCTGGAAAGGAGAGTGCACACC GCTCGCACCTAGGTCCAAGGGTGAAGGGTGCAGGGTCCCCCAC CACCAAGCACCAC	219
55	chr12	8807944	8808114	CTGGAATAATATTGTAAATTAACAAAGAACCATGCCATTTT TTTAACTGCGGTTAAAATTACCTGTTCATCTGTAGCGGGATC ATTCCACCGAGATCTGAAAGAACACATAACAAGGAATGAAAAAAT GCAGTCAGGAAAAGTAGCTTAAGGAAGCAGGATAAGGGGG	171
56	chr12	21644476	21644694	GAAGAATCATATTCAATTGCTGCCCGGCATCAGAAATCCTCAA ACACTGCCTTATTCTTTCAGGACTTTTTCTGAATAAG CTCTTGTGCTTCCGTAAGTTCTGAATTGAATTCTACTG CATGTAGCTCACTGGTTAGAAATCCAGTTCTCAGTTAGAGCT ATGGAGGGCAGGCCGGATAACAATGATTAGACAGAGTAATTAA	219

FIG. 17K

1700

57	chr12	26943052	26943213	T TAGCTCACACAATTAAAAATAATGTATTCTTACCTCTTGAA CTTCCTGGGAGGGTTGGCAAGGTCCCCGGCCTCTGGGTGCA CCACACATCTGTCATCCACTAACCTAGAAGAAAAGCATGTT ACATTGTTAACATCCTGTCATATTCAAAACC	162
58	chr12	48110493	48110716	ACAGAAACCCCTCCCAACTGGATATCCACAGTCCCTCAAGATG GCCTTGTCTCTGGAGAACTATTTTTTCTGCTCTCCCTTTGAT CGTGCCTGACTTGAAGATAAAAGTCGAGACATCGAAAGAAG CAAGGATAAGGGGGCAAGATTAGGGTAGAAAGAGGAATTGT GGAAACTCACGGCCTTGGGCAGGGATCCACTTGGTTCTGCTG CTGTGACGGGA	224
59	chr12	56558997	56559217	GAGGCTAACCTCATCTCCACACTCACTGGATTCCCTAGCCACAAAGT GAGAACTGGGACATTACCTGGGACTTACCATGGGTCCAGGGGGGAACC TCTGGAAACATAACACTAACCTGGGGGGCTCCAGCTGGTG CCTGGGTGGGACTGCCCAAGGGCTGGGGGGCTGGCTGCTGGCC CTGCTGCTGTGGCCCTGCAAGTTGACCCCTGCCTGCCAAATCTG TTCAAGAGGG	221
60	chr12	56575817	56576037	CTTCCACAGAGATGGCTCAATTCACTGGGATCCACGGTGTG GTAACACTGGCATGGGAAATTGAGCACACAGTAGAATCAATGGGA TAGAATTGTGGAAAAGTAAGTGCAGATCTTACCCCTTCCAT TTACAGAAAGTATTGGTTCTGGCTTATTTTTGTATCTAT TCATTATAAAGTAGTAAAGCACATTACGAACATTCAAATAAATAA G	221
61	chr12	58129105	58129322	GGTGGGAGGGTCAAGGCCTTCCCTATAGGAAGGGACTGGGTTA CCTACCAAGGGCAGTTCAAGGAATGGAGCGGCTCAAAGTCCAT TCCTGGCTATTGATCACAGGCCTCTGTAACGGAGAAGGGCAG TAAGAGGGAGGGCTCCTGGCTCAAGGGACCCCCAACACTG AGCCACCCCTCCCTGCAGCTTCCAAACCCCCAA CCCTGCCTG	218

FIG. 17L

1700

62	chr12	62147364	62147566	CAGAGACATTCTGTTTCCCAAGCTTGCTGTCCACTTACCC GTGTTTGACTTATCCAGAGAACAGCTCCATCCTTCC GATCCGGAAACTTACATTCTCCCTCTAGACATGGCTG CATATGGCACCACTTCGTTCCACTATTGAAGCTATAAGAG AGAAAAAAATTGACTCAACATTTTTCT	203
63	chr12	62895230	62895402	TTTCTAAGGAGCAAAAACATGTATTATAAGAAGATAAGATT TTAAAGAATTGAAACTGCGGAGTTGGAAAAGCACATTAAATT CTAGCATGTGAATTGTTGCAATTATTCCAGGATCTTGT AGTTGGTTAATGCTGATGCTTCCTTATTGGCTTAGTGGCAT	173
64	chr12	111923037	111923208	CTACAAGCAAAACACGGCAGGATTCTCACCTTACCTGCT CTATATGCTTGGCTTGATTCACTGGCATGGCGTCATAGGTAT TGGGTATAAAGGCCCTGAGAGAATTAAAAAAATTAAACATTAG GCACCTATGATAATAACAGGTAACTCTGICACCTIAGTT TCTTACCCAGTGAAGGCCTGGGGCTGGCTGTCAAGATGTTG CAAAAAAGGTCACTGAAGCCAAGGGCAACGGGGCCCTATCTCATAGACAC CCACGACCCAGACGGGAACCGAGGGCCCTATCTCATAGACAC CATCTGCAAAGGGAGAGAAAGGGAGCTGATAACCTCCCCAGCA CCTCAACACCA	172
65	chr12	113747991	113748211	GACATCAGCATTCCCCCCCCATCTCGGAGAACCCGGGCT TCTTACCCAGTGAAGGCCTGGGGCTGGCTGTCAAGATGTTG CAAAAAAGGTCACTGAAGCCAAGGGCAACGGGGCCCTATCTCATAGACAC CCACGACCCAGACGGGAACCGAGGGCCCTATCTCATAGACAC CATCTGCAAAGGGAGAGAAAGGGAGCTGATAACCTCCCCAGCA CCTCAACACCA	221
66	chr12	123285560	123285788	GACAACAGGACTCATTGGGATTAGCTCAAGTATACTTAAACCTCCA TGAATGTTAGAATGATGCACTGAGACATGTGTATTCA TTTTTTCTAATGCCACCATAGCAGAGACATGTGTATTCA CCCTTGAAATAACCACCCAAAGTCGATATTAAAGGGGAAAAAA AATCAGAAAGTCACGTGTTAAGGACCAAAATTGAAGGCCATGT TGTTCAAGCAA	229

FIG. 17M

67	chr12	125455794	125456014	GAGCTGGGGGCTAGGGCTGGGGCTCCCCAGCACCC GCCTCTGGCTAACCTTAGGCCGGAGTCACAATGGGGAC AGGAGGCCGATGAGGATGCCGTGTACACGCTCCTCGTGG GCCTCGTGATGATCACCAACCTTGACCGAGCAGGAAG TCTGGGGGGAGTCCGGGGTGAAGGCCACCCCTCCAGCT AAGCCCCAAGGGCC	221
68	chr13	36124971	36125153	AATAAAATCATGATAAGGCACTTGAAAGATGACTTGTA GCTAACTTATCCTGCCGTTTCTAAACTCTGTGTTTCTCT ACCACAGCCTCAGTTATGTTAATTCCCTGATCAAGCAACAGT AAAAAAAGTTGTCTATAGCTTGCCTCGGGTTGGAGTAGGGAC CAGCTATG	183
69	chr14	21899812	2189997	GTAATGGAGGGTACTCTCAAGGTAGGGCTAGGGAGGGGA GGGGGGGTTACTGGCTCTCCCTCCCTCCCTATTAAAGAAA AAATGTAACACAATGTAAGAGGACTACTCTCAAGTAAAGGCA AGAAACAAAGTGCATGTCAAGTTGACCTTGCCTCATGGAGCAA GATGGCTTACGGTCTTC	186
70	chr14	23993275	23993502	GGGGCTTCCCAGTGGCATCATCTACCTGGCTAATAAGA GGTTGAACCTGGCAAGGATGCAGGGGGCTGGCTGGCAGGC GGTTTCACCTGCTGGACTAGGGCACCCCCAGGGGTGCTTCG AAACTGGCCTTCTCTGCCGGGCTGGTATTCTGGAAACCA GACCTGTACCACTGCCTTTTGAGCCCCACCTCTGGAGAT GCAGTCGAGCATCTTGCCT	228
71	chr14	24040425	24040631	CTGTGGCCCTGGCCCCCCCCCACTCCTCAGGCCAAGGCTTTGG GGCTGGAGGAAGGGACCCCTGGTCCAGGGCAGTGG GCTCCGGCAGGGCAGGGGTGCCAGGGTGGGGAG GCAGGACTGCTGGGAGTTCAAGGGATCCCTCTGAGGGGGT CAGTCCGGTTCTCATATAACCCAGGGCTGTCTCATCCAGGGG CTC	207

FIG. 17N

1700 ↗

72	chr14	29237919	29238088	AAATCAGGGGCTCTTCCAACCCTTAATACATTAAACATCCCT GGGACCAACTGTAAGTGAACGTTTACACACATTGCATTGT AAATGATAATTAAAAAAATAGTCAGGTATTTTAAGCCC CCCCCTCCCATTTCTGTAGTTTCAAGCCACTGCAG	170
73	chr14	35782138	35782338	GTGATCCTGCAGGTACTACTGTGGGTTAAAGCCACTGCAG CGGGAGTTAACAAACTGAGTCACCCAGCTTCCTGAAAAAA AGTGAAGAAGAAATTGATTGACATTGACAGACAGTGGA GTAAGTCACACAAAGGGAGCTGACTTTTATGCTATAGAA CAGTGAGGATCTTGTATAATTCTATAC	201
74	chr14	68290256	68290475	GAGGCATGGTAGCAAGAACAAACGAGGTGGTTATCAC AAGAGGCTGTGTGACCCGCTCTGAGTAGACATCAGATCCTTACCTG TCAGGTAATTAAACATTTTATTGATAAGTTTATGCACA AGTTAACTTTATACTTAATACCTTAACCTAAACTTGAAAATGAAAT GTAGGCTTACAAAAAGATAAAATGATGGTGCTTTCCATAG	220
75	chr14	71511686	71511901	AAACTACATAATTAGAAATTAACTGTAGAAAAATGTTAGTT CTTATATTCTAAAATGGACGTTGTCATTACAGATGGAAATTG AACTGACAAATTAGAAATTCTACAATTGAGTCTCAAATCACTCC TAATTTTATCATTCTATTATTATTAGCCTGCCCTCAAGTATG TGTGTATACATTGTTGGCTTTGGGGTTTGTAA	216
76	chr14	94046519	94046727	ACGATGGGTTAGAGCACAGCTTACAAGGACATTTCAT TATAAACATGTCTTCCAAGGCCCTGGGGATCCACAC CTGCCAGAAGGACGAAAAGCAATCGAGTGCAACTTATGTCAG CTAGTATCCTCTGCTATCAGCTTGTGAACCTCCTGGAG AGACTAGCTCCTAAAGAAGAAAGCCGGCTGGGGTAAG	209

FIG. 170

1700

77	chr14	96783355	96783582	ACTGGTCTCATCAGCCCCGTGAAACAACTTCAAAATATCAAATA CTTCATTCAACTAGATATGTTACCTTGGAGATTATTAAAGTAAG ATACATTGAAAAAACACTTCGGATTGTAAACCTCACCTTCAAC ATCTGTGAACACTGCATTAACTCCATGATTAAATTCAAGAAGAAC TAGAGAGAAAACCTGAGAGTCTTGAGAGTCTAATTTC TTTTCC	228
78	chr14	97326817	97326991	ATAAACCATGAAATTATTTTTAGAGGTAGAGAGAAATAATAITG AAAGTTATAAAATGATTGTTATTCTTTAGGTGAAATTGCCAA ATACATGGAAACAGTGAAAATTACTAGACTACACTGAAAAACCTC TTTATGAAAAATTACCGTGACATTCTTTGCAAGGACTAAAA	175
79	chr15	23890791	23891011	GGGGGTCCGGGGCTGGGCCTGGCAAGGACTGCAGGGGGGTGCCT GCCAGGAAGGGCTGGAGGGCAGTGTTGGCACCTCCGGCTTGC GGACCCCGATGCCCTGGCCCTGCTGGGGGTAGCTGGATTG CACGGGCTTTGGGAGGGGGGGCTCCCTGAAAGGGCTGCT CCAGCTGGACCAAGGGGGGAGCCTGGCTGGCTCTGGCTGG GCAGGGCAGGGGGCTGCCA	221
80	chr15	33876543	33876762	CGCCCCATCATTATCCTTTAACCTCTATGATCAAAACATCTCCCT GTTTGCAGCTGCTCTCATCGGGAAACAGAAACATTGGGC TCAATTCTCCAATAACCTTGATTGGCTCATCAGTAATTGGACA GACTAGAAATCTCAGGTGAGAATTGAGAGAAACTATAAC TAGAAAAGATGAAGGGTTTTTCTGTAAAAAAAGAA AA	220
81	chr15	33915874	33916098	TTAACCTGAAGTCGTTTACCTCTCAACAAATTATGTTGAGGGT GGGAGGGTAGGGATTGATCTTAATTTCACCTTATTGGCTGCA TTGTTTTTTAAACCTCTGTTATTTCTGCTATCTCAAAATGACA GAGTACGTTCTGATGGGTGGACGTCATGGAGAGTTAAGTT CCGGCCTGGCTATGCCCTGGCTATGAAGGCTTACTT CCAAA	225

FIG. 17P

1700

82	chr15	42111683	42111899	GAGGGCATCACTGTGGCCACACAGGGTATGGGCTGATGGGGC CAGGGCTCCAGGGTTGGGGGGGGGGGGCAGGAGAC ACTTCCTCTGACTGCCTGTTCTGCTTAGCCTCGGTCCCCAG CCCAGCTTGGCCCGAAGCCGTGCCCACTGGGAGATGAGTC GGGTGAGTGCCTCATTTAAATGTTCTCAAGGTTAGAGGA GGCAAGGGGA	217
83	chr15	48427112	48427284	GAGGACTGTGGCAGGGTACACAATTAGTGCAAGGAGTTCTTG GTATAAATATGACAACCAAGTTACTGGTAAGCTTGAAGAAAATAT TCTTATGTAAAAATTAGAGATTATGAATTCTGATGGTTCAAGT AATTTTTTTCAGAAATGTTATTCAAGTCACCTTAATCTGCC	173
84	chr15	48766283	48766494	CTAATTGACCTGGTTCCAATTTTTTTCCACAGAGCTCTAG TGCCAAATGATGAATCTTCTTCTATAACTGACCCAGTCTCAA AAAGAATTGCTAGGCTGAGAAATGTTGAACACAGTTGTTCAAGC GACTAGTGTGACACAGTTGTTCAAGCCTGTAACACCTGTA CAAGTGAAGGCCATCACCTGTGTATCCTCCCTGGCACAG	212
85	chr15	56166145	56166315	CCTATACCAACCGGTAATGGATAAAGTGGAACATCCACTGAC CTAGGAAATCATCTCTTGCTATAAGAGAAGGAAGAAAAATAT AATTAGTAATAAGCTCAATTCTTAAAGAGTTGATGAGATA GTGCCACCATCACAGAAAATGTAGAGTTTTTTTATT	171
86	chr15	56386714	56386934	CCACTATAACCCATTATTCAATTCCATTCAAGCTTGGTTTGTCA GTGTGTGGCATGGTCTTGCATCGGTTTCACAGGACTACT GAGACAATGCTGGGTCAATGATGCCATACTGGAGCTTATTG GAGTGAATGCAACGGATTCTGCATTCAACAGGGCTGGGG GGACACTACTGCTGCAGTTAGAATGTGGAGTACCAATGGGTG ATGTCG	221

FIG. 17Q

1700

87	chr15	65477240	65477463	TTGGCAGAGCCCCATTCCACGGGAAAGTGGGGTGC GCGAAGGCCGCCGCAATTCCCTCACACTCCCCGGTGCAGCA GGCCTCGTGCCTCCCCGGGGGGGGCAACCATTGCCAGTC CACCCCCCCCCGACAGGCTGAGGGCTCAGGAGTGGCACTA TTTCGTTAACCTCTGCGGAGGTGAGGAGGTGATGA GCCGGACGGGGGGGGCGC	224
88	chr15	78587181	78587406	TGAGTATACTTGAAAAC TGCA TGTGGTCAGCATGAC TCTTAACAGGGACAACACAGTTTACAGTTTACAGATGAATTGC TTAACACAGGCGAGGCTACTAACCTTGCTTGAAGAGAAT ACCGTACTAAAAAGAACAGAACGACATTAAACCATGGTC GATTTCAAGATGTCACATTGCCTCAGAACAGCACTCC GAGTAGACAGA	226
89	chr16	460646	460836	CCCCGACACCCGGGCTCACTGCCGGGCCTTGTTGTT GCAGGACCAACGGGAGGGTGAACATCACTGCCACCC TGGGGAAACCGGGGGCAGGGCTCCAGGTGCATGCAGGCTC CGCCAAGGCCGCTGTGGGTATGACCACCCCCCGCCAG GTTTGGCCACGTTGGGTCCCCAATGGGC	191
90	chr16	1250536	1250733	GACAACATCGGCTACGGCTTGGATTGCCATCTCCAGGTGGC GGCAAGATGGTGGGACGGGGACCCCTGGGGCACGGCAGGG AGGGGGTGCAGGACTGCCGCCAGGCCAGG GGCACGGGGAGGGAGGGAGGGAGGGTGTGGGGCAAGTC AAGCCACTGCCGCCAGGTGATCACGCTGGAAAGGC	198
91	chr16	2095983	2096207	GGAGGGATGGCTGGGGCACGGGGTGCATCCCAAG GTGCTGTCAGGGGGAGGGACCTTGCTAAGATGGGGTCATCTG TAGAAAGAAAACAGGACCTTGCTACCTTTGGGGGCACTGGAG GGCAGATGGGGCCCTGCTACATAGCAGTGGCCACAGGTGATCC TTTTGGTCTCATGGCA	226

FIG. 17R

92	chr16	30960586	30960791	CGCCTGGTCCCCCCCCAGGATGAAGGGGGGAG GGGACGGGGGGAGCAGCTGGCTGAACCCCTGAGGGG AGAGCCCTGCAGGCTCGGCCACGTACCGGAGTTGTGCA CGGGCTACCTGGACCTCATGGGGCCAGTCAGCACTCGCT GGGGGGCTCAGTGGCCGGCTAACCTCAGCGGGCCA AGC	206
93	chr16	53493271	53493448	CCCATTTGTGAAGTGAATAAGTCTTCTGGCCTAAAAATAATG TTCTTTAAAAAATGTTATTAAATTCTGTGTAATTTTTTCTTTA GTCCAAGGCACTTAGAATCTCCACACCCTAAACTGCTGTTAGG TACATTAAGGAGAAATAGCCCTTGTGTGACTCCAGTTCTACAG C	178
94	chr16	56377713	56377929	GCCCAGTACGGAGGCAAGAACAAAGTCAGCCCCACAAGAGATC TACACCCACCGTCACCTGGCCACGGACACAACAATCCAG TTTGCTTTGATGCTGTGACGGGACGTCACTCATCGCCAAAACC TGCGGGGCTGTGACTCTACTGAGCCCAGGCCCTGCCG GCACCCCTTGCCTGCCCTGCCGCCCTCCCCCTCCCTGG AACCAAGGC	217
95	chr16	58709797	58710005	AAGTGAGGCACTTAGTOCCAGGTCTGGGGGGAGGC ACTCACCTGTCCCCATGTAGACAGCGAGGGCTATGACCATGG CAGCTGTCACCACTCCACCCAGGTCTTCACTCAGGCTGCT GCATGCTGTTGAAGGACGGGCACACTGCTGACGTGGCACTGT CCAGGTGAAGGGCACCGTCATGGGGAAATGACCTAAC CC	209
96	chr16	66759707	66759882	CAAGTCTCCAGGGCCCTTGTACCTGACTCTTGCTGTGC TCTGCACCCACCAAGCAGACTCCAGGAGAGGCCTGCT TTTACTCAACAGACTGTTGAAGAAGCTGGCCAACACCCCTTC ACTTGCTGCATTATCTAAGAAAATTAAAGAAAAAGCAT CATAT	176

FIG. 17S

97	chr16	67469949	67470180	GAGGTGAATTCTTGGGGCTCGAGCTGACCAAGGGCCTC CTGCCCTGCTGCGCAGCTCAAGGGCCATCGTGAETGT GGGAGGCCAGGGGTGAGTGCCTCCCCACTGGAGCAA AAAGGAGCCCCCTGGGGTGGGGAGGGCTAGGAGCCCC TTGCCAAGCTGAGCTGCCCACTCCCAATCCATCCGAGGG GACATGCCATATCCGTGCTGGGGCC	232
98	chr16	67697657	67697861	GCCTCCGGGGCTCTTGGCAGCTCTGGCTCAGGACACAGCA ATGAAGTCCACCCCCAGGCCCTGGCTCCTGCAGGACACAGCA AGCGTGGTGGGGAGACAGCATTATGTTGTTGTCCTGAAAC ACCCCCCACACCCAGGGAGGCCCTGCCAACACCACATGCC CCCATACATCCTCACCTTAGCTTCCGGACCTGGTGGG	205
99	chr16	70714642	70714859	CAGCCTCCAGGGCCACCCCTGGGCAACCCAACCTCCAGGCC ACGCCCTCCCCACCTCGGGTGGTAGCCATGTCAGGCCAC TTCTCGAAGGGCATCCAGGAAGGCCACAGCAGGCCAG TGGTCTGGAGAGGGACAACGGCATGGATGGTCGGTGGT TGGGGGGGGGGGGCCTCTGCCTGCCCTCCCTGCTGG GGTGGCCCTGGGGCC	218
100	chr16	85682150	85682378	CTCCACAGGGTCTCACTGAGCGAGTCGTCCTCCCGTGT CCTCTCCGGGCCACACACAGCTCCCCGCCAGCACACCC AAGCGCGTGGCCATGGGCCCTATCATGTCCTCCCTGGGG CCACAGCGTGGCCAGCACCCCCCCCCTGGTGACCATCGCTC CAACCAAACCGTGAATGGTGTCTGGAGGAGTGAGAGCCGG CAGGTGAGTGGGGGGGGCAGGGGA	229
101	chr17	1964719	1964933	GCTACACTGGGGCCCTGGTGGCCCTTCAAGAACGGTCCAC GGGGGGGGGGCCAGTGTGGCTCCCTCAGGCCACCTGGG CCACACGTGAGGAAGGGCTGGATGTCCTGGTCAAGGAACATTC TTGTGAGGGCTCACACAGCAAGGTTGGTCAATCCGTCAAACA GCACACACCTCCGGAGTAGCAGGGATTGGCTCTGTGGGCA TGAGAGAG	215

FIG. 17

1700

102	chr17	4805073	4805259	CTCCGGCGGCTGGAGGCCACGTGCCAACGCCCTG GCTCCGCGCTCCAGGCCGAAGCCCCCGGGGGTGC TTCCCGGTTGGCCCCATAAGCCCCCAACATA ATCCTCCGGGCTCCGAGTAGCTCTCCACGGAAATAAG CGAACAGTCTGCCAATCGAAG	187
103	chr17	4875522	4875747	AGTACCGGGATCACATCCACAGGCCCTGEGGCTGTAGCGTCC TCTGGAGGAGCCCCATCATCTGAAGGCTGGGGAGGGCAGG AAGGGAGGAGGGCCCCCTGGAGGTTGTAGCCTCATAGT CCATGAGGAGTAGGGGGGCTCTGGGACACCAGAGGAAGC TGCCCTGGGGCATGTCCTCCATAGTCATCTCAGAGGCTGGC AGAGGTGCAGGGGGGGAC	226
104	chr17	5238397	5238565	ATATAACTTTAAAAATTAGTGTGCCAGGTTATTTTTTTCAATA ATGCTTAGATGTGAAAATATGTGGGTGATGTTTGTAGAACAGTT CGTGACTATGAGCACCACTTCCACCTTAGGCTGGAGCAGGAG CGAACACAGTGGGGCACAGTATAAGAATCCGCAGAGA	169
105	chr17	8387374	8387585	AATATAATTAACTTTTTGATCTTGAGAAATTCAAGGGAAA TATATTGTTGAACAAATGGATAAGTAACAAACCTCTGC AATTGAAGGATTCTGCTTCCAGACTCTTCAAATTCTTCACT CTCTTGGATTGAGCAAAATCTCACTCTGGATGCACGAGCT TCTTCTAAATTCACTGGTAATCCTTCATCTGAGCC	212
106	chr17	8397017	8397232	GAGGCCACCGGGCCCTCTAAAGAGGACGCCACACCTGGAC TGAGGGGGCAGCAGCTGGTTCTCCAGGTCTGGCTGGCTCC TCCTCCTCTCTGCTGCTGGATCCGACTGCTCAGGTTAGTTCTGGGTG CCAGCTGGGGATCCGACTGCTCAGGTTAGTTCTGGGTG TCTCCTCTGAAGAAGGTCCTGTGTTTTAAAGGGCAA CACTT	216

FIG. 17U

1700 ↘

107	chr17	10555703	10555906	CCCCAGGGAAACCAACGTGCGAAGGCCCTCGCTGACATAGATCA TCCAAGATGTGTAACGGTCCTTCAGGGTGTACAGCACGGCTG GCTCATTCAGGTGGTCAGCATGGCATGGCTTGATCCTGCTG CGAACCTGGGGGGTTCATGGGTACACATCCCTGGTTGA CCACCAGGGTCTAAAAGGAAGAGGCACAGTGCTC	204
108	chr17	17041117	17041318	GGAGGTGAGAAGGGCCCTGCTTGGATCCTACATGGGCTTAA AAAAAAAGTCCCAAAGCAAAGCTGCCAGGGCTGAGTC CGTCTCCCTCCACTGTCTTCCAGAGGAGGCCAT GAGTAGCCACCGGATGGACTGTGGCCGAAAGTCGGGTGG AGAGCGGGTACTTCTCTGGAGAAGACCAAAACAG	202
109	chr17	37340957	37341145	GGCCCTCTAGGGATCTCTGGGGTCTGGGGTCTGAAGACITTCATAA GGCCCTCTGCATCCACTCCCTCTAGGAAGTCCCTACCTCGT AGCCCTTGGCGACCGGTCCCACCAAGGATGATGGGCCTCATG GAAGGCACCGTCA TAGGGGGCACATGCTCTGGGG GGGAAGCAGGGAGGGAAACC	189
110	chr17	37873516	37873740	CTACTGGCCTGGGGTGTCA GTGCCAGCCCCCACAAATCTT TTCTGGCCCCAGGAGGGCTGACCAAGTGTGGCTGTGC CCACTATAAGGACCCCTCCCTCTGGCTGGCCGCTGCCAG CGGTGTGAAACCTGACCTCCCTACATGCCCATCTGGAAAGTT CCAGATGAGGGGGCATGCCAGCCCCATCAACTG CACCAACTCGTGAGTC	225
111	chr17	39976564	39976759	TGTCATTGACTCCACAACCTGGGGATGTGGGGAAATCAG GACACTGCCCCGGCCATCTGAGACCTGCAATGAGACCAAA GCCTTGGGGACTTTGTTCGATACCATTACAACCTGTTCTTGCTG GACGGCACCCAGCTGTTCAACCTGTTGGGTCCCCGGGGGG CCGGGACTGGGCAAGGTGGTGGCACAGG	196

FIG. 17V

1700 ↗

112	chr17	46655079	46655267	AGGGGACCGTAAATCCGGCAATGGCGAGTTATAAGGGGG ACAATTGGCTTCCCAGCTAACCCCCCAACCCATGCCT CCGAAGTCCCTTGGTGTAAAGCTCCAGGGTGGAGGGGG AAGGGGTGCCAACGCACTCACCGTGCTCACGTGAACTTTG CGCATCCAGGGTAGACGGGG	189
113	chr17	49067892	49068094	CAACCACTGTGATGCCCTAACAGGGCTGTCTGTCCTGCTG AGCTGTTGCTTGTACACTCCACATAAAAGATGCTTGTCTACC TGACCAATTCTGAAAGATCTCTCCTGCAGGGTAGTCTGT CTCGTGCACCTGAAAAATAAAATCTCTTGTAAACAGCAAAGTAAA CTCTTATAAAATTAACATTTTTTTACTAAAA	203
114	chr17	55334119	55334306	AGCCAAGGCACCTCGGGCAAGCAACGACTCCCAGCACGA CCCCGGGTAAAGTTCCAGGCCGCTGCCAACCGGGCGCCTTG GGCTCGGTCTCCTTGACGGGGGGGACGGGGGGGG AGCCGGGCATCTCCCGGCCGCCCCCCTCTCCCCGCC CCCCCGTGGAAAGTTACACACCTT	188
115	chr17	71398001	71398208	AGGCTGGGCCACCCCTCTGCCTGGGGGCCAGGCTGGGCT TGGGGGGGGGGCAGAAAGAGCAGGGCTGGGGGTGCG CTACCATCTCATGGGTGCGCACCACTGGCTGGGTGCG TGCCCCGGTCCCCGGGGGTGGTGAACACAGCACTGAGGTGA AGTACTCGGTGAACCTCTTCAGGCCAGACACGAAGGCCACGT GGAT	208
116	chr17	79176004	79176210	GCTTCAGGGCCCCCCCCACCGCCCCCAAGATCGGTACCCCTC AGCCTCCTCCGTACCCCTCCCTGGCAAGCCCGT GCTGCCCGCAGTTGTTCCGGGCTGAGTGGTGGCACTGG AGACATTCTCGGCAGCTTCCCAAAGGCCGCTGCTCACTGC CCTCACAGGGTGGCTGCCTTGTGATGTTGCTGACAGGTAAG	207

FIG. 17W

1700 ↘

117	chr17	79990719	79990939	GCCAGCCCCGGAGCTGGGGTCCCTGAGATCCGCAC AAGGGAGGGAGCAGGGCCCTGGGAGCCCTGACCACACTCC ACCAAGGTCCCACCTTTTCCAAAGGACGTCCTCTGATCTGCT TCCTCTGTGAGGCCGCTCCTCGAGAAATGTTCTGCCA AGTAGGGCAAGGCTGGGGCCCTGTGGGAGAGGGCTT CCCCAGTACCTGCC	221
118	chr18	6588830	6589034	CAGAAGTGAATTCAATTGGCCAGACCTCAAAACTGTGGTGAGT ATCACCAACAAAGATCTGGGGGOAGTGTGTGGGGAGTGT TTAATAACCAGCTACCCAGGATAAAAAAAAGCCAAATTGGT GTATTGCCAATTCCAAGTGTAAATAACTCTGCTATGCTACC AATGCCAGTCACTGAACATGGAGTGGGAACCTG TGAATTGAATAAAACTAGGATTGAAAACAAATTCTGAAACAAT TTTCTCATTTTACCCAGATAAAAGTTGCTTATAGGTTTGAGT TAGAATCAGTGGCTTICACATAATATGGATAATCTTTCTTA TGGTCCTGTACATTAAAGAAAAAAAGATA	205
119	chr18	10675104	10675309	ATAACTAGAAAAGTGGTTCATGAAGGTCCAAGCACAACACTGT TGAATTGAATAAAACTAGGATTGAAACAAATTCTGAAACAAT TTTCTCATTTTACCCAGATAAAAGTTGCTTATAGGTTTGAGT TAGAATCAGTGGCTTICACATAATATGGATAATCTTTCTTA TGGTCCTGTACATTAAAGAAAAAAAGATA	206
120	chr18	13612629	13612848	CGGAAGCTGAAGGGATGCTTGAACGTGGGGGGCTGGGT CACAGTTGGACTCCCACTTGCAGAGGACCTGATTATGTCAG TGACCAACCTGAAACAAACAGCACACTGAAGGGAGGCTCAGTCAA AGACCTGTTCTTAAAAAAGGTACATTCCAAAGAACACTGCTATG GATGATGCAATGCTCTTCTGGGTTTGTGTGGTTCTGGAG TTGGTCT	220
121	chr18	13681596	13681823	AAACTTACCTCAAGGCTGCATAATAAAATGTTCCAAACCAACA CCAGAAGTTATTAGATGCACTGGAAATCAGAACCTTTCCATACTG TCTAAATGTCCTCTGAATGTTGATAAAGACTAATAGATTGTC TTGCAAGGGATCAGGCTCTCCCTGGAGTTCCCTGAG GCTGTGGCACTGGATGAAAAAACCTCCTGAATGAAACATGTT GCCTCCCTT	228

FIG. 17X

1700

122	chr18	21053254	21053426	AAAGCTAACAGGCTAAAAAAACAAGACTTGACACTGAGAGAG GTAGATCTAATTAGTTACAGAATCTTTTGTGAATTGTTG TACTTGTGAGCTGGCAACTGTACTGACTGATTTCCTCTC CACTTAGCATGGGGATGAAAAGGATAAGTAAAGTTATA 173	
123	chr18	34322638	34322861	AGAAGAACAAAGTAAACCCAAATTGTACACTCTTTATTGCTAA TTTTTTTTAACTGAAAGGAAGTAGCAGAACCACCTCTGGACC TGAAAGGAAGGAATAAGACCAGTTGGAGAACAAATAAACCTTGG GCCTTATCCTGTCTACTCTTAGCCATTGGAACTTTCTAAAT GGAACTAATGTAAGTCATCCCCATCCCTCATCCTGTAGCCCC CCTTGTCA 224	
124	chr18	66377125	66377341	ATACAGCTCTAAATTATAAGTTAAATTGAATTTCACATGAGCAA AAAAGTCTTCCTCCCACCCCTGAATTAACCTGTGCATATAATGTT TTTGCTGTGAGAACAGTAAATAAAACTAAACTAAGTGAAGGCTTAC TAGAATAGGAAAGTAGCATCCATCTTCCAACCTTAACCTGGAGA ACCAATGCCTTTCAATCTCAAGACCAACTCATCCAAATG 217	
125	chr19	873563	873791	GATCATGACCTTGTCAATTGCTCACAGGAGACGTGGTGGGG TCAGCTCGGGCCTCTGTACACACAGGGTACCTTAACTGCAC CCCTCGGTCTGACCTGCACTGAGTGCCACCCAGTACCA GGACACAGGGGACCCACACGGGAACCGGAAGGGGGGA TGCGCTTGCGGGACGGAGGAACCTGGAGCCACTGCGCTG CCCCCCCCCCGGGGGGCGCTGTG 229	
126	chr19	1056145	1056357	GGGGCTGCTGAGTCCCCCTGCCTGGGGCCCTGAC TGTCTGAAAAACCTCACAGGCTGGCTCACAGCCTGGATGC TCAGGACAGTCTCAAGGTGGAAACTGGGGGGAGGTGGG CGTCCTGTCAAGCAAGGTCACAGCAACCCATTGCTCTGACCTA TGACCTTGACCCCCACCCAGATCTGGTTCAACAAAGGCT GGCACTC 213	

FIG. 17Y

127	chr19	6361560	6361773	GGAGCCGACCGGGGTGCAGGGATGTGGCCCCGAATA TTGGTAGGGGGGGGGGGTGGCTCATGCAGGTAACCCGC GCTGGGGCCTCGCCGCTCACTTCAGGCCAGGGC CGCCCGAGGGACACTCCAGAACGGCCTGGCCCTGCAGGC GTGCCCTGACGGGACGGGACCCCCCTCCCCTGCATTC CCATCGTGGT	214
128	chr19	9407949	9408176	TCTTTCTCACGAATGGCACTCACCTCAAATGTCATGGCT TTTGTTCGATAATCAATACACAGGAACCTCCGGTTTCATGTA AACTGGAGAACCGGAATCATTCCTTTGAATCTTACTATTTC TGTTCATGGATAATTCTCCACAAATATCTTGTGAAGCACT GATTCAATTAGTTAAGTTGAGTCACAAACCTGAAACGAAAAA AAGTGAAA	228
129	chr19	10102511	10102729	GAGGACCTCTGGGAGAAAGTGGGGTTTCAGTGAAAGCTAAGA GGGGTCATGGTGTGGGGGTTAGTGAAAATTGACTTGGGGG GGGTCACTGGAAATAGGGATGTGGACAATCGGGGT CAGAGTGTCAAGGGGTCAAGGGTCAAAGGGCACAGGGCAC ATACCACATTGCCTTGGCACCAAGGACCATCAATTCCAGT CACACCTAGG	219
130	chr19	14040360	14040544	GGCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG GTACCAAGGACATCATGCAACGCCAGGGGGGGGGGGGG AGCTGGAGCAGGGGGTGTGGGCACTCCGAGGGGGGGGG TTGGAGATGGCATCTGGGGAGGGGGGGGGGGGGGGGG GCATGGGGGGGGAGGGTGCAGG	185
131	chr19	17994510	17994526	CAAGGGAGATAGGCAGGAACAAGGGGGGGGGTCCTCC CTTTAGGGAAACTGAGACACAGAGGAGTCCTTGAGACCCACA TTGGAGTTCTGAGGTCTCGCTTCCAGGGGGTAAACTGAG GCCAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG ACCTACTCTCCAGGGCTCCTTCACCGTCATGGGAGTCATCA GGGGCCCC	217

FIG. 17Z

132	chr19	37975959	37976178	TTAGTTACTGTGGATCCCCAACATCAGAGATTCAAC TGAGAGAGACCCATGAAATGTAAGGAATGCAAAAAACCTTC AGGCAGCATGCACACCTTGCCTCATACCAGAATTCAATTC GGGAGTCACTGTCAACCACCCAACCCAGTCATCACCAAGTCC TATAGATCCTGAGTCTAAATGTTCTAGAAATTATACTGTTTTT AT	220
133	chr19	39989454	39989648	CTGTTGCTCCTGCACGGCTCCCCCTCCCGTGAACGCC CCCCCCCCGGCCCTTGGTCCTGGGGGGTAGCTGCCTG AAGGGGGGGCCATTCTCAGATAATAAGGCTTGAAGGCCA GCAGCTGGACTCCGAGACCCCCAACAGAGGCCATGG TCCTCCCCACGGATGTCGGGCTCCCTCTCCC	195
134	chr19	46393696	46393911	GGGTGGGGGGCATTACTGAGGGCCCCCAGAGAACATC AGGCACTTACTAGCACAGGGAGTGGGTGCCTCACATCCCTGGA GGAGGGAGCTAGGGTGCAGCTGGATGGCAGAGGCATCG CCCCCTCGACTGGCCAAAGCTGGGAGAGGGGTTGCAGG CAGGATCATAGATACTAGATTTCACGGAGATTCCATCGGGCG CGCCCTTC	216
135	chr19	48737622	48737838	AGCGTCCAAATAGTTTCAACGGAAATAGCTTTTACCTGGG AATGTCCCCCCAGATACTGACACTCAGGAACAGCACGGAA CAATAATGGCTCTGCCTCTGCTCATCTCTGAAAAAT GTGAGATGTCACAAGGGTCTCAGAAACACAGGGTAGCTCC TGTTACCTGGAAAACAAACAGAATTCTACTATGAATAAA G	217
136	chr19	49982146	49982354	TCTCCCTCCCCCTGCTCCAGCCCCCCCCAGCTGCTTCGCT TCGTCAGACCAACATCTCCCCCTGGATCACTCGCCAGAAC AGCAGCTGGTGGCCTGAAGGCCCTGGATCACTCGCCAGAAC TTCTCCCCGGTGCCTGGAGCTGCAGTGTCAAGGGTAAAGG CTTCAGGGACCCCCCACTCCTCCCCCTGTCCTCACGGCC G	209

FIG. 17AA

1700

137	chr19	50098758	50098989	ACAAAGACGGCAAAAGGTGTTATGGAGGCCTGCCGGGGT GCACCCAGGGCCCCCCCCAACCCCCAACCCCCCTGACCGCCACCCAA ATGTCAGAAGCCTGGGGCCCTATGCCACTG GGAGGGCCTCTGGGCTGGAGGGCAGGGCAGGGCTTA TTCCCCCGGTCAAGGGCTCAAGGGCTCTGGGACCCCAGGGCT ACGGCAAGGGTTGGAGGGGGCAGGC	232
138	chr19	51227540	51227737	GGAGTCCCCGGGTCCCCCCCCACCCGCCACCCGCCACCCCTCAG GAGCCCTAGATCCAGTTCTCAGTCCTAGGGCCCCGGCT GCATCTCCAGGCTCCCCCTCCAGCATGATCCTGTCTGTCCTG CAGTGGGGCGGCTGGGGCCTGGACGCCAGGGCCTGGCACCA GCTGCACCGTCCGTCATGCCGCGTTGGACACCCGC	198
139	chr19	51528815	51529026	GAGGATCCTGCCCTGCCCTGCCGTACCTGTTGCCA GAGCAAGCAGGATTAACTGCAGAAICCTCATGGCCTGGAGGG GGGAGGGAGCGGGCCCCAGGGTCCCTGGGAACAAAGGGAGGG ACATGGGGCGCATCACTTACGGGAAATGGGGGGGG GGCTGGCTCATGCCCTCTCTCCCTACCTGCTCCCC TCCCC	212
140	chr19	53517955	53518181	CACTGTCCCCACATGGCCAAAATCTACTGTTCCCTGGAGGG GCCTCTATCCCCCTGCATGCAATCAAACATTCCAGCAGGGTGG AAATCGCAGTTACACAAGTGGTCCGACAGCCACATCCCCGG TGGGCCTGTACCCCCCTCCCTGGCTATGTTATGTTGAAAGTCC AAAAAAATAAAACTGCCCTTGTATGGAAAGTCCAAAGATAACCTAT TCAACCCCCCTG	227
141	chr19	535533116	53553342	CACTGTCCCCACATGGCCAAAATCTACTGTTCCCTGGAGGG GCCTCTATCCCCCTGCATGCAATCAAACATTCCAGCAGGGTGG AAATCGCAGTTACACAAGTGGTCCGACAGCCACATCCCCGG TGGGCCTGTACCCCCCTCCCTGGCTATGTTATGTTGAAAGTCC AAAAAAATAAAACTGCCCTTGTATGGAAAGTCCAAAGATAACCTAT TCAACCCCCCTG	227

FIG. 17AB

1700 ↗

142	chr2	9666190	9666391	ACCCACCCAAATTTTTTTCAATTACTCGATGCAATATCAAATA CTACTAAGTAATTGTAGTTGACTCTCCCTCTGCCCAT GTATCTGAGAACGGATGATGCTGCTACCAATAATTACACG TGTTCCTCATGGGATCTGGGTCAAGCTCTTCACTCGATG AACAGCTCTAATATGAATTGTATG	202
143	chr2	21230989	21231213	GCCAGCCTTCAACAGTGTCTTGACTIONGCAGGTCAAGC CTGACTAGAGAACGTCAGTTGGGATGTTCAATTGTGAAAG TATTAGTGTGCTATCCAGGGTAAGGCTGATTGTTATCTTGAC AATCACTCATTACTAAGCTCCAGTGTATTGTTCTGTGTA AACTTGCCACTGTGTTGAIITTCCTCAATAGCATTTCCAAA AACAGCAT	225
144	chr2	27015940	27016163	TTTAGCCAGGTTCACAAAAAGCAGCCCCGGCGCCCTTCATCT GTGTCGCGTCTTTCTCTTCTGTCAGGCAGCAGAACGAT TTCTAGTTCACTCTTGAAGGACCCCTATCTCCTCACCTTACAT GCAGGGCCGAGTTACTCTCTCCAAAGGAATGTGCAACTGGCC CGAGGATCCGGGGCTTGAGGAGGACTCGGCTGAGCTCC TGACCCAGTG	224
145	chr2	27660058	27660244	TGAGGGGAGTGTGATGGGGAGCACTGAAGTTCTTAGGTGTC ACTTTTTTTTATTCTTCCCTCACTGAAGATGGCAGTGTGGAG ATTCAAGGGCAATGGAGAGTCCTCATATGTGCCACAGGAAGCC ATCAGCAGTGCCATCCAGCTCTAGAAGACCATTACAGAGG GTAAGGATCTTCAGGAT	187
146	chr2	27892084	27892290	CTGTGCTATAATTCCACAAATCATGAAGATTACCTGTGGAC GATTCAACTGGAAAACAACCTGGTGGCTGAGGCCATTCACCA GGAAAGAAAAAGAAGCAATGATCAGTCAGTGTCTCATTGGAAGCTT GTGGGATTCTTGACACTTGGGATTGCTTCCGGCAGGAAGCAG GTCAAGTATAGGAGCCCTTTCTAGATTAAA	207

FIG. 17AC

1700 ↘

147	chr2	43624945	43625123	AAAGTATAACTAGGAGATAACCTGATCTTCCGAAGTAGAAAAAA AAAAGTTCTACAGATAAAGTTAAGAGGGCTGTGTATCAAATCTTC ACCAGCTAACAGTACATAAGAAACCAAGTTCATCTTATGGTGT TTGACAGAAGTCACACAAAGCAACTTACTTACCTGGAGAAGT GTCC	179
148	chr2	71337070	71337257	TCAATTAAATTAGTTGCTTGCCTGCCAG TCCACAAGGACTCCACCAACAGTCTTAGGATGGAGAAAAATC ACTGGTTTCCATGTGCTCTATTGACCTCTTCACTTAGACT GGGGATCTCTCTTTTCAAAATCCATCACAGCTGCATTAATT ATCCACCTAA	188
149	chr2	72723548	72723751	ACAACCGAGAAATTAAAGTATTCCCTTACCTTTCCAGT TCTATATCTTGAATGGGAATTGTCTTACCCACCTTTGTACATC TCTTCACTTGTACTGGTATAGGACTGTAGTTGTCTAGAATCAAG TATGTTCTATAAAATGGAAAGGATAGACAAGGGAAAGGGAAAGGA TAGACAAAGTGGGGGTAGGGAGGG	204
150	chr2	97507756	97507959	CCCTGGTGTGCACTGCCCTAACCCCCAACCTCCACTCACC TTAACCGGGGCTCAGGTTCCCTGGGGACTCTGTGTC AGTCGCTTTTCCGTCTTGAACCAAGTTCCAAGTCAGCCA GGTTATCCAGATTAAATCTGGTACTGTAAATCTTCAAGCTGG GGCAGAAAGACAAGGACCATGCCAGCCATCAGCCCC	204
151	chr2	120005734	120005930	GCTGGCCGCCACCGCTACGACCTGGTCAACCTGGCAG TCAAGCAGGGTACCCACCCATGCCCTCCCTCTGGCAAC TCAGCACATGCTTGTCCAGCACCTCCCCCCCCACCAACCAAG GTGGCAGCCGATAACCCACGGGTGCAGCCTTACCAAGGTGCCAA CTGGCAGATCTGTCCAAATGGGGAGCACC	197

FIG. 17AD

1700 ↘

152	chr2	152684571	152684759	TTCGGAGACCCACCTAGGCCGGGAGCTGACCCCTCC CCAACCCGTACGCCGGGACCGAGCTCCCTCACCCCTGG TTGAACAGTCTCCATATTCTAGTGAAAGAGAATTCCCATTCGG GCAGGGACCCCCCCCCTCCAGACACCCGGGGCCTGGC TCCCCCGGCTCAGGCTGAGGGG	189
153	chr2	166747135	166747346	TGGITCTCAGTGTACCTGTTAGATGTTAAAAGAATTATTAT TTCAATCACTGACTACAATAATTATAGTTATAATCTGTC GTAATAAAACTGTTGAATTATTGAAATGTACATTTTTAC TACAAGAAAAGCTTATTCTACTTGATTACATGAGAAAAAA AATCATGTAGCATTATTGTTAAATAGATTAA CTCTACTCATTTGCTCTCTGAGGTGGGATCATTCAGCA TCCTCTGAAGGGAGATAGGAACCTAACACGCCCTTTCTTGTGA TTTGATTGGTCGTGCCCTAAAGAAATCAATTAAAGTGTCTAA GAACAGAAATCCTAAATAATTGTAGGTATAAGATAAAATCTTGCTT	212
154	chr2	167138216	167138392	A GGAAACCTATGGAAACTAAAGCTGGCTTATGAAATTACTTTT CCTCCATCAGGTATGAAGATGATAAAAAAGAATGGGCTGGGAT GTTGAAGGAAATACGTTATGCTTCAGGGAGCTAGTTGCCTAGCA ACACGTTAAATCTTCAAAACTGCTAAACTGTGAACAAGGT GACAAAAACATAATAGATTGGGAGGGGGTTGTTT CTTATTAATAAAGTTTTTCCTTCACCTCAATCAAAACCTACC CTTATTGTTAGTGTACCTAATTCAATTATGCTACTATTACCT	177
155	chr2	170382041	170382247	CTCCAGTCCTACTAGCCTCTGGTCAAGATAACACTGATGCAGA AATTGGTCTGTATGCAACTCTGGATCCAGCTGGATCTATTAA GAAAAAAATAAAGATTCAATATTAAAGAATAACAGGTAAG G	207
156	chr2	176044737	176044959	CTTATTAATAAAGTTTTTCCTTCACCTCAATCAAAACCTACC CTCCAGTCCTACTAGCCTCTGGTCAAGATAACACTGATGCAGA AATTGGTCTGTATGCAACTCTGGATCCAGCTGGATCTATTAA GAAAAAAATAAAGATTCAATATTAAAGAATAACAGGTAAG G	223

FIG. 17AE

157	chr2	207460712	207460930	ATTATCACTATTCTACAAAATAACAGTGAAATGCCATATCT GTTCTCTTTCTTCACAGGTGTTAGTAATGAAGGCCACTTGCA TTTGTGATTCAACCTGGCAGGGACAGATTGCAGTAGCCGGG ATCCAGTTAGGAACCTTCACCCCCAAGGATGAAGGACCCA AGGGTTTGTGATTGGTTCAATTCAATGGAAATACTGAATTGTT	219
158	chr2	227661547	227661780	GCTCATGGCATATAAGTCTCCACTGCCCTTGGGCCACTGGG CACTGGGCACCCCTGGGACATGGCATGTAGCCATCATC CGTGTGGAGGGTGGAGCTGTCTGGGGGTGTTGCCCCCCC CGACGGCTCCAAGGGGTGCATTCCAGACCCCTCTGGGTAG GAGCGGGTGGCACGAAGGGGGAGTGCCTGTGTCCCCGCA GTGGGCCTCCACTGCCACCTCTGGTGG	234
159	chr2	233388461	233388659	CCCGTGTCCCCGGCAGGCCTCCAGGGGGAGGGGGGGCCC AGCTGCAGGGAGCTTCTGGGCTGGGACCCCCCAGGGAGCT GCAGGCAGAGGCCGGGGGGGGGGGGGGGGGGGGGGGG TGTGGCCGGGGTCCAGGGGCCCTGTGGGGCTGGGGCA CCAGCAGTGCCTGCAGGGGGGGGGGGGGGGGGGG 160	199
160	chr2	241808231	241808439	GGCCAAAGGGCAGTGCAGCCCCAGGTCCCCAGGGGGAGGTT GGGTGGGACCATGGCCTCTACAAGSCTGCTGGTGAACCCCC CCAAGGGCCCTGCTCAAGGCCCTCCATCCCCAACCAGCTC CTGCTGGGGCCTGGCCTTCCAACCTGCCTCTCGCATCATG GCAGGGGGGGCTGCAGATGATGGGCCATGAGCAAGGATA	209
161	chr2	242062110	242062330	GGTGGCCGGCCACTCCCCCAGGCCAGGTTACCTCC TTGTTTTTCTGTCACAGGAGTCACAGGAAAGCCGAAG GCCCAACTGCCAGGGGGCTCATGGTACTGTACTTTGGAG TACTGCCCTCACAGGGCCCAACCCCTCCAGTTCCACAGCC TTGGGGGGCTCCTCAAACAGGGTCTGGCTGAGCACCTG CAATGGAGAAGAA	221

Figure 17A

1700

162	chr20	2596671	2596873	TGGGATGGAATGCCCTGGCTTGGCCAGTGGTTGGTAGTAG GGTTGGCTGAATTCAACAAACGTGCTTTTTCTCTCT CTCTCGCTCCCTCCCTCTAGCCTCATATGCTGAGTT TGATATTAGTGGAAATGTGCTGGTTGATCTCAACCAAGGA ATGATCTGGTGAGTTATCCATTCACTCTGGT	203
163	chr20	3127958	3128141	ATCGGGGTATGAGGCCAAATCATATGGCTGACTGGGG CCCCCCCAGCATGGTCTCCAGTAAAAGAACAGTTCAAGAAT TCAGGCTTGAATTCAATCCTCTGCTTAATAACCACAGCAA TTCAGACCCCTCTGCTGAAGGTGAGTACTAAGACGATTGCT CTGTAATCTGG	184
164	chr20	4850517	4850740	GGACCAATCAGGAACCTGGCTCTCAGGAACGTACCTGTTTAT TGCCTGGATGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG CAGGGCTAGTAGTCACCAATAGACTCGATGATGCTGGCGACC ACGGCACTGAGGATGCCGATGACACCCGGCGAGACACGGT GGCGAGTCCCCACTGAACCTGTTGGAGGAAACAAACCATGAAT GCTCCTAGAGGGAG	224
165	chr20	17581551	17581729	GGAGATGTTGGTAACTGATGCTTCAAGCATTGTT GGGAATGCTTCCTGAAAAAGATTGTCGCTATGCTTTGATGAT GCAAGCTTGAACAAAAGAATCCAGAAAAGAAGAGTTGATGTT TTTTTTGTGGTAAGCATGTTAGAAATATTGAGCCTCTGTAAA CTCAT	179
166	chr20	29965056	29965278	AGTCCTGGCCACTGCTTCTCATAGACCTGGGGCAGCA CTTGCCAGAAATGCTTACGCTCATGAGGACTGGGGCTGTTCTCTT GGACTGACAATTCTGCAATTAGAGGTAGGGCTGTTCTCTT TTGCAAGAGGGCCCTACAAATTCCACTGTTACCCATGCATCGAA GGATGTGGCTGGCTGGCTGGCAAAAGGAAAGAAATACAATAGT TAATTCCAC	223

FIG. 17AG

1700

167	chr20	31671142	31671356	AATGACTACAGACCCATTGAATTCCCCGAGGGTGTGGTGAIA TTCCCTACAATGACTCCATGTCGAGGACCCCCCAGATA TACCAACGGCAAACAACTTGATGGTATTACAGTATGGTCACA TTGAGACCAACGACAACACTGCTCAGCTGGGGCAAAATACC GATATGGTGAGATCCTTGAGTCCGAGGGAAAGCATCAGGGACC T	215
168	chr20	62193098	62193324	CCTGAGGGTAGTTGGGGACGTGAGTCCTGCCCGCGTGC AGCTCCCCAAAGGGCCCCACCCGACCCCCCGCAGCTGC CAGGACATCCACCGACTTGTGGAGGGGGCCAGTACAAGAT GCAGGGACCCCCAGCCCTCTCCCCACGGGGGGCCT CGGGCTGCACCTGCTCCCTGGTTGATTATGAAACCAGAATA CGATGTGGAGGGCCACGATC	227
169	chr21	34971370	34971571	AAIAAATTGCCCCGGTTAAATTTTTTTAAATTCAATCAAAC AAAAAAACTACACAGGCTACTAAATTCTAAAGGTACCTACCTCC AGCATCTAGGACAATATCTACTCCAGGCCACCTGTTCTCC AACACAGCTTTCAGCAACATGAACCTTCCATTAGATAACATCAAT CACTCGGGCTGTAAAGGACAGGAA	202
170	chr21	43281516	43281717	TGTTICATTCTGGGGCCTCACATTCCTGGCTGTAAATGGAAA TAAAAAACACCCACCTCATGCAGTTGGAGGAGCATAAAGTGA GATACACACAGTATGCCCCAAATGATTATCCTCTGCC CCCCCAATCCCCAACTGTGCCACCGCATGTCTTCCCT CAACCTGCTTCTCTGATGCCCGTTGCTGAG	202
171	chr22	19951236	19951464	CGGGCATGACCCAGGGATGGTGGATTGGCTGGCGTGAAG GACAAGGGTGTGCATGCCTGACCCGGTTGTCAGACCTGGAAAAA GGCCGGCTGTGGCAGGGGCATGGCACTTGTCCCTC CCACCAAGGTTTACACACAGTTCACTGAAAACCACATCA CCAGGGCCCTCAGTGCCTCCAGCCTGGGCTGAGGAAGA CCCCCCAGCAGCTCAGTGA	229

FIG. 17AH

1700

172	chr22	29120871	29121053	CACTTAAACCATATTCTGTAAGCACAGGACAATTTCCCTCTA TGAGGAGGTGGAAAAAAATTCCAGTAACCATAAGATAATAA TATTACCTTATTCTGCTTAGTGCAATTCAAAGATTGT TATTCAAAGGACGGCGTTTCTCTTCCCTACAAGCTCTGTATT ACAAAG	183
173	chr22	31942804	31943008	TCCAACATCATATAATATGATTTTTGCCGTAGATT TTACTATGAGCAGCGATTACTACGGAGGTCTCGAAGAATGG AAAGAGGAGTGGTGGTTTCAGGCACGAGTGAAACTCTGT GTCAGGCTGACTGTCACATACAGGTCAGGTTCATGTTACACT CTTCAGTTTACTGACACATTACAGAT	205
174	chr22	37532326	37532535	GCTTGAGGGCTGACCGAACCTGAAGACTCATCTGGGT GTCTGGGGTGGCGTCTCAGGCAGATCCATTCTGCTTCCTG CTTGAGAGTCAGCAGGGGGCCTCTGGTGGAGACAGGA CTGTCAAGGGCCACCTCACCTCCCACGGCAGGGCAGGGG CTGAGCATCACAGAACCCCCCTGGCAGCTGCA CCC	210
175	chr22	38895342	38895566	CTACCCCTTGTCTCCCATCAACCCAAATTAAACAAATTAAAGGGTTC TAGATTGAAGAACACTTACGTGGGAAGTTAGCATGATGGAAAGG CAAACACGGGGTTAGGACAAACATCTCCCCCTCACTGTAT CTCCTCTTCCGGCTAGCTCATCAACCTCATACTATTGAAAAA AAATGAAGAAGTTAGTAACTAGTTTCTAAACATTCAAGAATT TATGTT	225
176	chr22	41572128	41572356	GIGGGATAATTGGCTTGAGGCCAGGGAGGGTTGAGTGA GCCAAGATCACGCCACTGCATTCCAGCCTGGATGACAGAGCG AGGCCCTGCTCAAAAAAGAGACTGTCTGTTTCAGGCT TCCTTGTGATCCGGCTCATGGCTGCCAACTCCCT GCCTCCCATTTGATCCTGATCCCTCATCCCCCTGCGATCTG ATGGATGGTCGGGATG	229

FIG. 17A

1700  
↗

177	chr22	43243504	43243704	GAGAAGATTCACCTACCGATAAAACTCAAGTGAACACCAAGT GACCGGGACCCGTGAGCAATCAATGCAAAGGAACACTCCA TAGGTTATGCTTGCCAGCTGGGATTGGCACCAATCAA AACACACCTGAAAAAAATGTTAAATTCAAGTTACTAATTATGTT AGCCAAATAATTATCCTGTAAAGATAACA	201
178	chr22	50506933	50507112	CTGTGAACAGCAGGGAGACGTGAGGCTGCCTTATGGCAATCA GGACCTCCACCTGGAAAAAAAGGGGGGTTGCAGGACCCAG GCCGGAGCCTGGGAGGCCAGGGACAGCGGGCTGCCAGGG CTGGGGGGGGGGCATGCCGGCACTATGGCATAGTCC CAAGCTGGTTCCCTGC	180
179	chr3	10301786	10302003	GAACAGCCCCTTCAATTCTGTTTCCAGGGTTGATTCC AAAGATAGTTCTCATAACTCCACAAACTCTGAATTTCAGCTGA AGCTGAGGGTCAGAATGATAATTGAGGAACCCAAACAAGGT CCAGAAAAGGAAGAGGGGATAGACTTCGAGAACCGGGCTCCA CAATGATCTACCTGAAGGGCTATCCAGGGCATCCTGGGAAATC GATG	218
180	chr3	25470207	25470429	CAGTGCAAGGGAGATCATGTTGACTGTATGGATGTTCTGTC GTGAGTCCTGGCAAATCCTGGATTCTACACTGGAGTC TCCTCCCTGCATGCTCCAGGAGAAAGCTCTCAAAGCATGCTCA GTGGATTGACCCAAACCGAATGGCAGCATGGCACACTGCTC AATGTAGGTTATTTTTCCCTCTACCAAGAAAAAAAT TTGTCTCT	223
181	chr3	31666228	31666447	AACACCAAGTCTAGCTCTTCACTTTTCTCATT GGTAGAGGGCTTTACTTAATAACTACCACTCCCTAAATCTGAAA GTATTCAAGTGGTTTAAGATTACCTCAAGATTAAAGTAGTTAGC ACATTGTGACATCTCATTAAAAGAGTATTGTGATTGTTTCAG CACCGGAATACTTAGATGATTAGAGAAGCTTACTTTTG	220

FIG. 17AJ

182	chr3	38750898	38751116	GCTGGGCAGGAAACAGGAATAAGGCTTACCACTACAGCA GAAACCAGAAACTCCTGAGCACCCTATCTTTTCTGTG ATTGAAGTTGTCAATTATGACCCCAACAAAGAGATTCACTG AAGAAGCCCTCCAAAATGATGAAGATGACAAGTACAAATACA TGACACGTTGTCCTCCCACCTGGGTGCAITGACCTGTGA CCAG	219
183	chr3	42678977	42679202	ATGTAGTAGTACAACCGAGTTGAGCAGAAATAATTCCCTGTAAATA CCACTGAGTGACAGTCCCCCCCCCTCAAGATGGAAGGCCGTGGA CAGAAACCTTGGAAAGGCCCTTATGAGCGAATTCAGGAAATGAA AAGCTAAACAAACCCATTGCTACCCATCCAAAGCACATTACAG TTAGCAAAATTAAAGAGACTGGTAGCTCATCCTACCCATA AAAGAGAAAA	226
184	chr3	47371337	47371515	AGAAATTGTTCAATTGTTGGCTTACCCAGAACCCAGCCTGA GAATGATCAAAGATAACATCCTTCTCTTAATTTTTTTTC TTTCCTTATGCCAAAGGCTTCACCCAAAAACTGCCTGGGTGT GCCAGTTGCTGAGACAATGATGTGTGCTGTACGA CGCTGC	179
185	chr3	48636498	48636720	GTTTGTGGGGGGGGGACCAACCCCCGGCCCTGGCCT CTTGCTTACATAGGCTTCCCGCTAGAAAGCCAGGAACAT GCCGCTACGGATCCGGTTGTAGTCCTGGAGCTGCTCAATGGG GCCTGTGGGATAGGGTGGAGCGTCAGGGTGGGTACCTGG CCTGCACAGGGACACCTGGCCCTGTGCCGACCCAGCTGG AGATGCTGCCAGAG	223
186	chr3	52029360	52029592	GTGCCCTCTCATCTGTCCCCAGTTGCCGTGAAACCAAGTCCTT GCCAGGGCTGGGATGAGGCGGGCCAAAAAAACTTACACT GGTTGTGTGGTTGGACTGGCCATGTCCTGCAC CTGGAAGACAAAGTGGAGATCAGGGTTAAGGGCTGGCAG GCTGGGCCACCTCTGCCCTGGCCATTCTGTCAGAAATGAGCC TGCAGGCAGAGAATGTGCAGCCTC	233

FIG. 17AK

1700

187	chr3	52272921	52273112	GGGTTGGGAGGGCCAGACCGGGCCCTGCCATCC CCGAGAGCCAGGCCAGGGCCCTGCCATCC GCCCTCACCGTGGATGCCCGTGGTGGCCATGGCTGGC GCTTCGGCTCCGTGGCCGGCTGGAGCCCTCCGCTTGA TGCCCCGGCAACGCTCGCTGGACCAAGAGAG	192
188	chr3	52954633	52954843	CTCACAAATACATTAGGTATAAGGCTTCTTAGAACCTGTTAGAT TTAAACAAACAACAAACAACAAACCTTTTACGACCA TTAAATAACCTCTTCTATAAACTATTCTAGATAATGTTCTACTC TTATAATAATGCAGACACAAAGTCTCTGACACAGAAAGTCATA TTTTCATGATGATAAACATTAAGAAAAAC	211
189	chr3	58486602	58486809	TTACTTTCTTAATTGTCAAAAGTTAGTTAAATTCAGATG ACTGATCAATTCTTACAAATTCAAGATGACTGACCCAGTCACA TTAAATGTAGGTGGACACTGTATAACAAGGCTCTACACATT GACCGCGTTACCCGGATTCCATGCTGGAGCTATGTTGGGG GGACTTCCCCACAGCTCGAGACCCTCAAGGCAA	208
190	chr3	113784266	113784435	CGGGGTTCCCTGCCATGGCTCAGCTTACGGCTGTCACTCCCTGTA AGTGTAGAACCATGATTCAAGTATCAACTGTGGCAGCAGCT TAGTTACCTTAGGTGCAAAAAGAGAAATTTTTTTGTTCAT ATATAATAATTGTGGATTTCATAATGAAATGAAAGAGA	170
191	chr3	121151106	121151326	TGCATCTGCACAGGCTTCCCTGGAGGACTTCATCAACAGCA CAGTTACACATCAAAGTCCCTAGCTCTCCAGCTGGCGCCT ATTTCACCTTCACCTTCAAATTCAAGACAGTTAACGCACT TTCCATTTCATTCTGACATCTGAGCTACACTAAGGAAAAAA AATGTAGTTAACAAACTGCCAGTCAAAGAATGTTTCCCAC CATCC	221

FIG. 17AL

FIG. 17A

1700 ↘

197	chr4	2941343	2941562	CATTTGGTCAGTGTGCTCTTGACACAGCTCCTGAAAAAACACG AAATCCCCGACTCTCGGATGACTCACCTGCTGGTGCCTGAGGA GCTCCACGCCACGGTTCTGGATGACTCACCTGCTGGTGCCTGAGGA CAGTGGCTGGAGCTCCCTGCAGCAGGATGGTGCCTGGCTG TGGCCTACCGGTTGGGGGGTAAGCCAAGGCTGTGG CTGGCTCAAGTCCTG	220
198	chr4	3432364	3432556	GAGAGCAGCGACAGGCCGGTCCACAGCCCCGGCTCAGGCC CAGCCCCCCTGGACCTCCCTGGACGGACCCCCGGCAGA AGTCTCCCAGCGGGCCCTCTGCACACTCCCAGTCCCCGCT CCCTCGCCAGGGCACCGGCCAGATCTGGAAAGGGCAG TCTCAGGAAGTGGAGGGGGCATCC	193
199	chr4	4411164	4411392	CCTAGCTGCCCTTACCGGTTCTCCAGAGATACTCTGCAC ATGCTGTGTGGGGGGGGCCTCGTGTGGGTGTGTGT GTCTGCTCCTGGCTGCCCTGGACAGTCCCACCT CTGCCCACAGGGTCTCCGTGGTCCCTCTCGCCCC CCTCACCTGCGTGTCTTCTGTGGTGTACAAGGTGTACAA GTATGACCGGCCCTGGCCC	229
200	chr4	47694916	47695114	AAATGCTGATGATTGTCAGGATTGAGAATCCCTGTCACCTCA GACAACCCCTAACAGCTTGTACATACTGCACATGAGCCTCGT CACTCCAGTCGTACAGTCGTGGCCATTACATTGCAGTT CCGGGGATGCAGATTCCACTGGCACACAGAAAAGTTCTCACC CCCTCCACAGAGCACTAAAAAAAGGG	199
201	chr4	56737032	56737222	TAATAATGTAAGAAACTCCTATGAGATAGAGTTCTTGGTAA GTATGATCATAAAATTACCAACAAAAAAACTAGTGTGATAAT AAAACCCATTGTAACATTTTTAATTTCCTTGTGAGAACAGG ATTAAAAAAATTCCAAGACCTAATGATAGGACTGAATAGCATC TTGCCTGTAATT	191

FIG. 17AN

1700

202	chr4	57337845	57338051	GTTAGATAATAATCAGATACTGTTTCTTCTTCAGTAAC TCTCTCTCTTGAAGGCATATTGCAGGTACAGGCTGAACA GAATTGAGAATGCCCTGAAGACAATAAGAAACTGCCAACCAGCA GACAGACAAACTGAAGGGAGCTTATGGACAAGTGGTAATTACT GCTTTAAATAACATGGTGGACAGTGATACTGAGT	207
203	chr4	68692857	68693056	TAGAATGGTTCTGTAGAAGATAAGAACACCTTCCAGA TGTTCCTAAATTAAAAAATTACCTTAATAGCAAATAAGACATT CTATTGTTAAATTTTTTCCGGAGACACTACCGCATATTCTTG AGCGCCCCATCCTGTTACATAAGCAGTAGAGCCAGGTGGAAATA TTCTGGGTAGCAGCTGGAGAC	200
204	chr4	70801580	70801743	ATCTAAATGATTGATAGATAACTCAACTATGAATGAATTAAACA TAACTTTTTTTGTAGCCCTATAACCTTAAATCAAGAGGGTA AGAATGACTCCACAGAAATTACCGGTGCAATTAAACAAGAGAAG TAAGTGTAAAGATCAGGAGACTCAGAACAG	164
205	chr4	77204471	77204689	GGAGGACTGTACTGTGCAAGTTTGTGCAATCTAAATGAAG GTCTCTCATTTTTTAGTTCCCTTGAGAAATATGTTATCGGGGA AGGAATGTAACCGTCATGTAATAAGACTCTATAAAACGAACT CAAGCATAGAAGAATCGTAGGAGATGATTAGGCAGATTATT ACTACGTACTTGGCTATTTCTCTGTCTCCTTTAAAGATTAA	219
206	chr4	77290666	77290851	TTCCCGCCCGCAGGGTCAATGGTGTGCTGTTGCCCTGTGTC TCGGTCCCACAGACCGCTTATTCTGCTCCTCTCCAGACTCAGC TCCTTCTCCCTTTGTGTAGATCAGCCTAAAGAGAGGAA TGCTTTTTCTCCCTGCTTAAATAATTAAAGATAAAACATATT GTCCTTATATCAT	186

FIG. 17AO

1700

207	chr4	78079669	78079892	TGGTGTCTTACTGCAGATGAAGGATTGGGGCAGAGCACT TGGCAGGGCATGAAGGGGCAACTCTCGGGTTGTTGAACG TCACCTGAAACAAGAAGAGATTCCAACCTCGAGAAAAAG GGCTGAGTTGATTGAGGCTACCCCGGAGGTAAGTGGCAGA AAAGATAACCTTGCTCAAGCAGCTGATGGGCTTGGAGGAGGT AACCCCCCCCCC	224
208	chr4	94749884	94750095	GCGCAGGCCCTTCAAGCAACCGGAGAAGCATAGTTGCACGGC ACCTGGGTGTGTGATCTCGGAGTGGGGGGAGGGTGCAGG AGGGAAAAAAATAAGACGTTGCAGAAGAGAACCCGGAAAGG GCCTTTTTGGTTGAGCTGGTCCCAGTGCCTGCCTCGGA TCCTGAGCCTCCGGAGCCTTGCAGTGCATGTCCCCGCTGCT GCAT	212
209	chr4	101336967	101337183	GTTTATGTTTICACATGAACAAAATCCAAGGTATAGTTAGAATT AAAAATAACTAATTTCATCTAAATCAACTTGTCTGGAAACCTACCA GACTCATGAGAAATTGCTTAACGGTAAGAAGCTTCACGGCTCT CTTTATCAGACTGAGGTCTATTGAAAAAAACATGAAATTAA ACACCACAGTAATCAGCATTTGATGTTTGTGAAAGTATTG ACAT	217
210	chr4	145032444	145032661	ATATAATTATTGGAGTGTCCAGICCACATGTTAATTTCACCTT GACTGAAGTCTTATGTCCTTGTGTTCTTCCACATTGGTT TGGTGAACAGATTCTCATGATCACCTGCTCTAGAAAGGATTAT AGTTAAAAAAATAACAGACCGAAATATGAGACATACTTGTGTTA AACACAAACAGGAAATTAAATGCACAGTATACTTGTGAT	218
211	chr4	183601587	183601797	GAAAGGGGGGGAGACTGGCATATTGCTTGGGGATG TGCTTTCTTTCGGAAAGCTGGGGAAAGTGTGGCTGTAATT CTCTCTGTCGTGCTTGCAGCCGCTGTCCAGTGTGTTAG TAGTGGCAACGGGCACTCCAAAGGGCAGTACTCCAGGACTACCC CAGGGCTGGAAAGGCACCGAGTGTGATGTGACTACCC	211

FIG. 17AP

1700

212	chr5	8899620	8899845	GGCTGGCGTGGGAATAAGGGCTCAGAAAAAGATGACACCACA CCCCCCCAGACACTAGCTCTCAGAAACGGGGGGTTACCT CTGAAGCTGGCGGAGGAAGTGTCCAGGAGTTGCCTGAATAGG TGTGCTCATTTGGCTGACAATTAAAGAAAAAAATTGA ATACAAGACCTTGATCCCTTCATTAGAGGCTCCAAGTTCAC AGTATCTGTCG	226
213	chr5	52161459	52161657	AAGCACTCATATGCCCTCTGAAAAAATTAGATTCAAATTGAC AATTGACTCCCTTTATTGAGAATGCAGCACTCAACTGGACATA GTCTAGTGGATGGTCCAACAGTATTACCCATGGACAA GTGTTACAGGCTTTTAATGACCTCTTGAAAGAATGGATATT GGTCTAACAGAACACAGGTATGT	199
214	chr5	54696172	54696340	CATCGAAATGTTCTCATCCACCTCACAAATGATCCAAATTGGA AACTGTGTATAAGGCTTTGTGAAAAAAAGCACAGGTATGGCAG AAATTGGTTTAGAATTCTATGTTCACTGAATAAAAT ATGTTGTTGTCATTTCCTTATCCCTGTTGGTTAAATT	169
215	chr5	60628763	60628977	GTCTCTGCAAGTCGGTGAGTCACGGGGCAGCCGGAGCCGTC TGTCGGTCGGTCAGTCCCTGGGGGGGGGGGGGGGG CTTCTCCTGGACAGCCCCTAGTTCCGGCGGCCGCA CCCTGGGGGTGTCCCTCACTGGCCCGGACGGTCCCTGAG CGGAGCGCCATTCCCTCCCTGGCCCCGGTGC CCGGGACGGG	215
216	chr5	65465905	65466072	CAAGACAGAAAAGACAGAGCTAGATCTAAGAGCCCACATAAAA ACGCTCTAAATCAAGGGAGAGACGGAAAGTCAGGAGTCGTC GCATTCAAGGGTGAAGTTAGAGAAATTAAACAAATAATTTTTTTC CTCAGAGGTCTGTAGTGCTAAGGGATAATAATTAAAT	168

FIG. 17AQ

1700 ✓

217	chr5	68669597	68669787	ATTGGGGAAATAGATGGGAGAAAATGTGATTATACTTACATGA TTTCCTTTTTTCAGGTAACAGACTGGGTTGACCCATCATTG ATGATTCTTAGAGTGTAGTGGCGTCTACTATTACTGCCACA TCATTAGGTGTGAATAACTCAAGTCATAGAAGAAAAATGGGC CTTCTACATTAGAAG	191
218	chr5	108516425	108516631	TCTCTGTCTTACTGTAACCTAGGGACCTTGCTGCAAGAAC GCCTGGTAGGTGAAAATAATGTTCTGAAATCAGTGACTTTGG AATGTCCTGTCAGAGGATGGTGGAGTGTTCATCTTCTGGC TTAAAGCAGATTCCATTAAATGGACAGCACCGGAAGCTCTTA ATTATGGTAAGAATAGACCCACATTTTTTTTTAAT	207
219	chr5	112884662	112884847	GATTGGGCTAAACACTTGGGCAGACTGAAATTGTGGATCTTC TAGAAATCTTACAGGTAAACCTTTGACTATTAAATTCTAC CGTCTCTTAAAGAGGCTTCACCATAAAATTAGGGCTATTC AAATGTAGACCATTATAGTCTTTATCCAGAGATTTTTGT GATTCTAA	186
220	chr5	127610184	127610387	GTTTATAATTAAATAATTTTTATGCACTACTAGAAGTTGT AAAATTCCATTGTAACCTTGTCTTACCGATAACAGCAGTGT GATGCTGTGTGAACCAAGGTGGACATTACAGGTAAACCCC CCAGGGGTGTTGACACAGAGGAACCTGGCAGTTATGCTGCTTGT TTTGACATTCAAGGTCTGAAATTAGAG	204
221	chr5	139862132	139862308	ATTGATGGATTATTTCATTCTTAACAGGAACATGAATCT GAAGGGTGGAGAACACCTTGTGAAAGCTGCAAGAGCTGGT CATTTGTGCACTGTGAGTTCTTATTAGCAAAGGTAAAGAAA GGCAAGTGTATTCACAGGCTACAGTTATGGAAAAAA AAAT	177

FIG. 17AR

✓ 1700

222	chr5	154173260	154173442	CTGCCAATCGGGAGAGATAAAAGGTATGCACTACCCACTATG GGGGCCCTGGACTTGGAGAACACCCCTCAGGCCTGACCTGGT ACCCCTCTCCCCATAGGGTCTGAGTCTGCCACCTACGTGCC GTGGCCCCCCCCACCCAGCCTGGCAACAGAGATCAAAC GGAGCCTGGCTGGCAC	183
223	chr5	170319512	170319708	ATTGGCTTCAGAGAAATTATTGCTGATGTGAAGAAGTTCTCC AGTAAGAAGTCATTGCTACATGGTTAGAACATCACTTACTTGC TTTTGGTTAGAATACTGCTCTCTTAAGCGATAAGTAGTTGT TTGTGACTACCAGTACCATGTTAAATTATGATTATTATTCCT CTGGATCCCTTATCATT	197
224	chr5	179050027	179050256	TCTTAACTCACCAGAAAAAACCTCTGCACITCATGGCCGAG CAAGACCAGGGCAAGCCCCGGGACCTTCACACGAAATCCCCT CCACCTTCCGTGCCAACATCATGGCTCTTACGGGGTCCGG CGTCGAAACAAACTGCAAAGGGGGAGGACCAAGAACTGAGA GCCCAATTAAAGCTGTCTTCGGCTTCGGAGGGGGGGGG CCGCACCGCCCCACCGA	230
225	chr6	13620279	13620498	AGGAGGGCTGAGGCAGGGCAAGTTGCTGAACCTGGGAGGCAG AAGTTGCAGTGAGCTGAGATCGCGCCTGGGTGACAGAG CGAGACTCTGTCTCAGGAAGAAAAAAAAGAAAGTAAGCAT GTGTAATTCTTAACCTTTTATATGATCAACAGCCAGAATA AAAGCTACTTGGCCGTAAGGCTAAAGACCAAGATCTGAGAG ATTCAAGGGC	220
226	chr6	15496537	15496767	AAACCCAAATCACCACAAGGCCAGTTCGGCTGTCAACCCACACA ATCTCAGGGAAAACTGAAAGTAGCAATGCAAAACCCGCAA CAGGTGCATCCCTCGGGGGCTCCAAAGTCCACTGGCC CGCCGTCAATGGCCTCAAGGTCAAGGTCAAGGTCAAGGT GTCACTGGCACTAAGGAGGTGGGGGGAG GCCCTGCAGCTGGGGAGGGCTG	231

FIG. 17AS

227	chr6	26476234	26476461	AATAACATATTATAGTTGAAAGATTTTTTAGTGTAGTGT TTCAAACTGTGCTCATTTATGCCAACCCATGAACGGCTCCCC CTCCATGTCCTTGTAGTTGAGCTTGAGGGCCAGTTGAAGCTTA TATATCATTACCAAACTTAGGGTGGTGAGTGGTGTGATGG CAGGGCTGGTTGAAATCTCCACTGGAGAGAACAGCAG GATTCAAGAT	228
228	chr6	30035118	30035342	CTCAGCGACAGAAAAAGGGAAAGGGTGTCTCATCCCCCTCCT CCTCTCCTCCCTGCTGCTGAGGCCATGCCATGGCGAGGCAGG GGCTGGGCTGAGCGAGACCGTCACTGAGACAACGGTTACCC TGACAACCGAGCCCGTGAAGAAGGGGGGGGGGGTGCCTG TTAGGGGTCTGGAGATACTGGAGGGAGGGGACAGGGAT TAGAAGAGTTGGAGGA	225
229	chr6	33263961	33264150	GCTGGCCCCAACACCCCTCTCAAGCCGGTCAAGCTGACCCCTGGCCCTC AAGTAAGAAGCTCTCAAGCCGGTCAAGCTGACCCCTGGCCCTC AGAGCCAAATCCCTCAGGGTGAGAGGGCAGCCAGGGTGA GTACAGAGATGGCTTACCCCTGGAGAAGGGATCAGCCAAGG GTGAGAGTAAAGCTGCAGGCCTGG	190
230	chr6	35705933	35706161	ACTTGCTGGAGGCTGAGGTAAAGGAAGCAGGGCCCC TAGCCCCGCTGGCCCTGGGGTTACGGCCGATCCCTGGCC CTGGAATGCCAGACTATCCTGGACCTCTCTGGCTGAT AGGGGTGGGGGGGTGCGGGAGATCTGGTGA GGGGGGCACTGGCCCCAGAACACCGCTCTCGCTCT CCATCTCTCCAAATACTCCAGG	229
231	chr6	38813192	38813411	AACACAATGATCTGAAGGCCCTTCTGGCTGTAAATGACTAA TGAACTGGGTGACCTACAGGTCCCCCTCAAGCAAAAAAAAT TCATGATTGATGACTTAATTGTAGTTGAAACATCATG ACATATACTTGGCTTCCCTCAAAGGATATTGCATACTGCCA TTAAGGAGAAGGATACTGAAGCCAAGCTCAGGTGATTGA	220

FIG. 17AT

1700

232	chr6	43969629	43969846	GGGCTGCCAGGAAAAATAACAAGACGTCCAATTAAATTGAA TATCAGATAAACAGAATAATTATTAGCATTAAGTCTGCCCCAT ACAAATATTGGGACAGACTTACACTATTAAAGTATCCAGTTAA CTGGAAGTCTTGTGTTTATTGCCAGATCTGGTAACCGTAG GTGCCTCCAGTTGAGGGAGGAGAGACACCCCTCCAC C	218
233	chr6	51921485	51921708	ACGGACAGCACCCTCGTTCAAATCCAAGCCGGAGAGGATGTT AGACCAAAAGGGTTCCAGTTGCATTTCAGCTTACAGTAACCTCC TCAATGGTTGTTGAATCTATTACAAAGGAAAAAAATGCCAGGA ATTATAATCATGAGCATAAAGACCACCCCCAGTTCTCCACCTCC CCTCCCTCAATTGGTGAAGAAGTCTCCACACAGATGGCTGTG GCATTTCG	224
234	chr6	52318712	52318926	ATAATTTCAGATGTTTAGTATGTTGATGAAATTTCCTTTTTTT TCAATTACCCATCTGAAAATTACTTCATAGTCCTTCAGCTAGT CTTCCCATCGTTGATACATAATTGCCAAAGTAGCCAAGTTGAAC TCCTACTTTAGGATTCTGAGTCACTACTTGGATCTTCAA AGTCCTTCGATTCTATGCAATCTGGGATGATAACAG	215
235	chr6	83810466	83810670	CTGACCATAGGCCAACGGCTAGCTCAATGTCACATCCAGCAT TACCAAGGGAGTTCATCGGAAGGGCCTGAAACATATGAAT TATCTTCAAAATAATTGGACCTAACGGACTTGCCAAAGATCTT TTTATATAGGTAAGAATAATTIACATATATAACAAATAATAAA AGAAAATGTTGCTAAACTATTAT	205
236	chr6	97246236	97246428	CACATACATACACAAACCCATAAAAAAAATGGAGGG CTATGAGAAAGGAAATTCAATGTCATAAAGAACAAAGCATCAC CCAAAATGTCAGCCAGTTCAATATTCAACACCACGGACTAGGAC TTCCACACACACATAGACAGGACTAGGACGTATCCGTGGCTT GTGTGACCAAGGGAGCTGC GG	193

FIG. 17AU

1700 ↘

237	chr6	111695291	111695498	TGCAATGCCAAAAATTAGGGTATAATGGCCAGCTTGTGAT CTAGGGAAATTGGATATTCACTTAACCTCGATTGATGACAA ATAGTTCTTGTATATTGCATTATCTCTATCTTGAGGAC ATACCAGAAGATATCTGTGATTCTGTGCTAACCTGAGATAAATG ATTGGAAAGGTCAAATATAATTGGAAATCTGCAATGAAGCTGA	208
238	chr6	112382373	112382555	AATCTGTGCCAAGCAACCAGGGAAATCTGCAATGAAGCTGA CCTCTGTGACCCACACAAGGGCTGTATTGACTACTCAGTA GACAGGCTTAGGTACGAGACTGGAGTGTGATGTAAGTGT CTTCTCTGGACCTGCTGAAAAAGATTGAGTCTAGACAGAAATT TTTTTTCCTGCA	183
239	chr6	117884360	117884582	CAGTGCCCAATTCAAGAATAGTCTGATTACAATCTTGTCTG GAGATATGGTCAATTAAATAGATTTTTATGATACAGAGTGTGC AGATCATCTAAATTGAAGCACCGTCATCTAGCGGAGTTCACT TGCACTGCCAGGTCTCCATTTCATGTTGTCAGTTACTGCC TTCTTAAATCCTGAAAGAAAAGAGAAAAAGAGACA CTTTC	223
240	chr6	118845012	118845201	CAACATTGATTCCCACCTTACTGAAATCCTGATAGCCAGGTGAAT AACTTGTGACCGAGATGATCCAGTCAAAACCTGCAGAAACATGTT TACACAATTAGTATAACTAAATTTAAAGAAATATTATATA TCATATTTCCTCAAGATTATTTAAAGAAACTCCCTCT TTGATTTTAA	190
241	chr6	119501438	119501622	GTACCTTGGCACATGGTTTAGAATAATTTCACCTACTTCAAT GTCTCTGCCAGGAAGAAACTCTGCTGCACATCATCATAACTCT CATGAAGAAGTAAACATCCCTAGGCCTGAATAGGGAGAA CACTCTGCAATGGTTTCCAAGGGCTAAATTAGGGAGAA AAAAAAAGTC	185

FIG. 17AV

1700

242	chr6	123673608	123673797	CTGAATCCAGTGCCTTCGGCAAGTTGTTGTCCTTCTAAAG AAGATAATGTCAACCTCCTTCATTTCACCTTGCTCCACT GTCTTGTTGTTCTCTCCCTTCCAGGTACAGGCTGCAA AACAAAGATAAGGTTAAAGAAGAGTTCCAGACAGAAATACTG CTCTGCTCAAAAAAG	190
243	chr6	130387399	130387616	ATATCATGCCCTGGTGTCCCATGTTTATTTGTTGTTATATA TATGAATATTCCTGAGTTAAAAAAATAATTCAACCTTTAACCT GTTTCAGGGTTGGGATAACCGATAAAAGCTTCACTTGATGG TATTCTGATTGCTATGACTCTGGGTGAATGCAGACGCTCTGG ATATCCACCCAGTGGGTGGGTGAGAAAACGGGACAA	218
244	chr6	142759315	142759520	CAGATTGATAAGGGTGTGTCATTTCAGATGGAGTAA GACAGCTACCAATATCATCAAGAAAAGTTCTGATAATCTAGGAA AACTTTGTTCAAGCTCCATTGGTCCAACTCAACCTATCTT ACATCCAAATCTAAATCCAGCTCTACCCACCTTCAAAAGGAA TAGCCACACAGGTGAGTCATAAGATGTCT	206
245	chr6	155728264	155728489	CTGGGGTGAATTGAGGCAATCTGGCTTAAGACTCATTTGTTCCAGT TGSGGCCTCCCATAAGGCTCTGCTTAAAGCTGTGAAGAGCCCAA GTCAAGTATTCTGGCATTAGATGACATTAGAAAAAAAAATAAG ACCTCTATAAAATAAGATAATCATGTTCAATTGTAGATT AAITAAA	226
246	chr6	161513872	161514051	TTACTGAACCCAAAGCAAAAAAGTGTCTTGTCTAAATAGTAA TTATAATTATGTTACTTTTCAGTATTGACAGTCTAAAAACTGCTA CAGCGTTTCTTGGTTGTTATGTGATTACTGAACCTTTTCGTT ACCAGGTGGCGACTCAAGGATTGATTCTACAAGCAATTG AACC	180

FIG. 17AW

1700 ↘

247	chr7	20180582	20180801	GAAGTGGAGAATGCAGTTACTCTCAAACCTCCAAAGAAATTAACTGTTAGTCACA CTAGTATTAAAACCTCCAGTTAATTCTCTCCAGTGTAGTCACA GGGTTTTTCATTCTGTCCAAGAGCTGACTGAAGGTCTTGTAACA CATCTCTGTAGTTCCATAGTGAAGCACTCCAGGTATAACAGA AAATCATAGGCAGGGTTCCACATCATCTATAAAGAAAAAAATA GA	220
248	chr7	27671945	27672112	AAATACATACATTTAGAATCCCATTTGCTCCGGAAATAAGCTTCTAT TGCATTGATACTGGTGGCAGGTGTAATAATTCTGTAGCTT TTCAAGCACATCTGCTGGGGAAAGATACTACCTTTAAAAAAAT TGCAAAAGGCACATTATTAATTAATGTAAGCA	168
249	chr7	44260229	44260444	AATCTGTGGAAGTCCATCCCTCAACCAGGGTGGCCAGTTTCTGG CAGGCTCAAACCGAGGTCASSCCTGGGTACACAGATTTCCTGG GGGCCAAATCCAAGTGAGGAGGGCCCCGGGGACCTCTGCA TCTGAGCAGGGCCCCAACGGGTOCCCACATCCCTGCA CAGGCCCCCCCCAGGGGACTCAGCGTAGGCCTCAA AGTCACC	216
250	chr7	55241592	55241799	TCTGTGTCTCTGCCCCCAGCTTGTGGAGGCCTTACACC CAGTGGAGAAGGCTCCAAACCAAGCTCTTGTAGGATCTTGAA GGAAACTGAATTCAAAGATCAAAGTCAAGTGTGGCTCCGGTGC GTTCGGCACGGGTATAAGGTAAGTCCCTGGCACAGGGCTC TGGCTGGCCAGGGCCTCTCATGGTCTGGGGAGC	208
251	chr7	73038331	73038544	GGGGCCAGGGGGCTGGGGGGCTAGGAGGACCTGGGGC TGGGAGAAAGGGGGTGTCCAGGGGGTGTGATGGGCTCCG GGAGGCCAGGACCCGGCTCGGGGGTGTGAGGGGGC CAGGGGGAGTCCCTGCCCGGACCCCCCCCCATCCCCACC TGAAGGCCAAGGGCCGTAGGGCCGGAACCGCCCTCACCTG TAGGCCAGGGCTC	214

FIG. 17AX

1700

252	chr7	77423269	77423474	GGAGATGATGTACCTTAAAGTCAGCTCTCACAAAAAGACAT TATATCCGAGTTCTGTTAAAGCCCCAGCCAGCAAGGCCATAA AGGAATGAAGACATAGGATAGATTCATGGTAGTAACACTGTTCCA GTTTAGCAGAGTGCAGGGCAGAAGGGCTAATTAAAGTAACAT TGCATATGAGGTACACGGCTTTTTTTAATT	206
253	chr7	87074245	87074420	GCAGGCCTGGCCAAAGCACTGAAAGCTCCAATTAGGATTGAAAAA AAAACCTGAGCAAATAACATGAGGAAAAGTTAACGTACATT CTGGCATTCAATTATCTAAAAATGTGATGTATAATGTAACATAGTA CATAAAGGATTAAGTTAGGTTTACCTTCTTTTTTTG	176
254	chr7	87179534	87179711	TGCTCTTCCCACAGCCACTGTTCCAACCAGGGCCACCGTCT GCCCACTCTGCACCTCAGGTTCAAGATCTACCA GGACGAGTGAGAAAAAAACTCAAGGCAATTCAAGGACACAG GATATAGGAACACTGACTGTTCACTAGGTTAAATAACATGCAC TTTTAT	178
255	chr7	102760012	102760205	AAGTCTACATAACCTCGGTICATAAGCTCCGATGGGAATAGCTG CAAGGTCAAAAGGTCCAATCTTTCCCTATCTCTCCAAAG GGGCAATAACCAGTATCTCTGCAGAAAAAAATCGATTCCAAAG CCCCCAAGACAGACCAGCTGCCCATAGCACCTTGTCTAT CCATTAGTCCTTACACCA	194
256	chr7	144098336	144098562	GGGGCAGACCCATGGCATTAGGCTTTCTGCTTCCCCGG CCTGGGAAATAGGACCTCTTCCATCTTCAACCTCTCCACTG GGAGGGAAACAATCTTCCCTGAGTCTGGCCTGGAGCGGGGG CCTAGAGTTCTGCTCTTGTGGGAACCCCTGGAGCGGGGG CGGGCACAGTCTCCAGCATAGCCCCCAGGATCTCCAG AGACTGCTGGCGGCTCTT	227

FIG. 17AY

1700

257	chr7	152549008	152549231	CCATCTGGCACATTGTTGCTCCATAAAAAATCTACAC TTTCCAAAGCATTAAAGTTGCCTTGTGGCATCTTCATAGG GCCGATGTGTCGTGCCAACAGGAATTGGACCTGTGGGAGTT CTGGGGTGGTGGTCCCGGAACTGTGGGGGGGGTTTA CCAGGGCTCCCTTGCTTGTGTTGAGAATGCGTACTC TCAGGAGGC	224
258	chr7	155465483	155465688	GTTAAATTCCAAGTGATCTACAACCTTGCAAGTCTTATGGAA TAATCATTGACATGTCCTTTTCTCCAAAGTGAACCTGAA GATGATTACTGGAGAATTGGCTATCTGGCAAAAGGTA GAAGTTTATGTCCCTCTGAGATGAATACTAGTCATTAAATTATA CTTAACATACAGCAGTAATCACCTTACAT	206
259	chr7	158444969	158445193	GGGGAGTTCTGAGTACCCCGAGGCCCTTCCGGCTGCTTCC TGAAGCTCCGTGCAAATACATTCCAACATTCTGAAATACATTTC TGGAAGTGTGCCAACAGCATTGCAACTCTCATCTGAATCTG TTTCTGAATCAAGGAACCTCCAGTTATCTTCAAGAATGTCA AGAAAGTAATGAAACATAAAAAATCCCTCACTAGAAAGAAAAT TGAATAT	225
260	chr8	18490123	18490323	TCTTCTTCCATCCATCTGCATGAATTTCCGGAGCCAAGAAT CCACTTTGTACACAGCAGCATTGGATCATGAGGAATGTCCA AAAATGGGTAGTAGTACTCCAATACGACTGATGGCTTGGAA TGTGTTCCGTTAGCTTCTCCTCAGTACTTCTGAGGGAGACT TTTTTTCTCTCATCTCTAAAG	201
261	chr8	31001076	31001261	AGAAGTTAAATTCTTAAACCATGTGATAAGATTCTGGGA GTAACATTCTAAAAAAAGGTACAGAGTCCCATTTCTATGTC TATACCTTGCTTTATGAGTACTTTTCTAAAGAGAAAGCTG TCAGATGTTGGCTTATTCATTGGCAAAAGGAAGTTAAATTAA AACATAA	186

FIG. 17AZ

1700

262	chr8	39068699	39068919	TITGAGGGCTTCAAATTCAATTCAAAATGGGGGTGCAAT GTCTTCAGAACATAAGCCACAAATGCAAAAAAATCTCCGAAACC AGTCTGTGGCAATGGCAGATTGGAGGGAAATGAAACCTGTGAT TGTTGTACTGAGGCTGTAAGTGAATGAAACTAGGAAATTGATA ATAAATTATTGTTTAATTATCCTTTCAAAATGTCAATATCT	221
263	chr8	42227229	42227425	ATGATAGTCTCTCTCAAATTAGACCACTTAGGGCTTATCTT TGGTATTGTTCCCTTATTGATTGATTGAGGCTTATATTGCTAA ACTTGAAATGCAAGGTAAAAAAATGTAATCTACTGTCCATTTTT TAGGTTGATACCCAAAGATCAGTATTACTGTTGTTCTATT TCACCTGGGAGTGATATT	197
264	chr8	52743959	52744172	CACCAAGTCATAATTCAATCAGAAAATGAAAGAGCTACTTA CGGAGTCCCCACAGAAATCTGGTTGCCATTATCATTCTTACTTG GTGCAACAAGTGGAGCAAAATGAAACAGCAAGGATAATTACT TTCCCAAGTGTCTGTCAGGTTCCAGTCGATAATCTGTGTTAACTATT TAGAGAAAAAAAGAAAGAAAGTGAACAACTTAATCCCAA	214
265	chr8	62370970	62371186	TTGGGCCGGACGTTACTCGGGCCACTACAGGGATGAAAAA TGACTATACTCACACATCCTATAATGCAATGCACGTGAAGCATA CGTCCTCGAAATCTGGAGAGAGAAATGCTCACCCAAAGCTGATGA AAAGGTAAGGCCCTGGTTATCAGAGCCCCCCCAGTCAGA GGTCAATGGAATTTTTATTGGACCTATTGTGAATTATAITA AA	217
266	chr8	86027371	86027588	TTGAATCTAGATTGCAAGAACATAATTGGTGAACCACTAAAT TTGACAGAAATAATTCACTACAGCTGCAGTGCCTAGAAGGTC TTTGGATAATTAGTTCTGATCTCCCTAAATAATAAGTG AAGATGAGGTTAGTAACTTTTCTGATCTCCCTAAATAAAAATA CCCTCAAAATTCAAAAGTGAATAAATTAAAGGATT	218

FIG. 17BA

1700

267	chr8	95686530	95686749	ACTCAGGGCCGCCATCAGGAGATGCCTTATCCAGATGAAG TCTGGGGACAGAGCATTATGGCTGCACAGAAGTGTCAAAA AAAACATGAAGGACAGATAATGTTGAAGTCTTCAGTGTTCAGC TGAGGAGATGAACCTTGTGTTAACGGGGCACTTTAAATCGA AATGGCTTATCCCCACCGCCATGTAAGTACCATGTAAGTTTT CTTGG	220
268	chr8	120580355	120580566	GTAAAATTAGAAATGTTTATTTCAATAACAGTGTAGGTTA AATATTGGACTTACGTTGAAGCAGGATACATTGGAAACCATA TGTTACAAGGAATGCATCATATTAGCCTCTGGTGAAGAGCT CAGATCTAAACATTAGAAAACAATAGTAACTACGTTGCTTACCA AAATCTTATCAGTTAAAGGGAGAGGGGG	212
269	chr8	124707889	124708056	AACATACTTTATCCAAAGCCAGAATGAAAGTCACCTACGGCTT TGGTAGGCCCTTTAATGGCGATGATTCCAGGTTAATCAATAAGAG TTACAAAAAAATCAATATTCCAAGGAGGGATTAACTACGACAC ATACAGAGAGAAGGGGGAGACCATGTCGGACAC	168
270	chr8	126033822	126034026	TGAGACATATTGTGACGTCCCTAAATGTTAGAATGTTTCAT ACAGAATTATGATGCTGGTAAAAAAAGTAGGATGTTTGTG AGCTTCGTAGATTATAATGTAATTGGAAACCTTGGAGTAAATTAA GTTTCTGAGTCTTACCAATATGTTTTTCTTACAGATTGT CTCCCTAACCCCTTAGTTTAATTGGAC	205
271	chr8	145095241	145095438	GCCAACCCATGAGCAACCTGGTCCAGGGCCCCAAG GCTGCGGGCTGGCTGAGCCACTCCGGGAGGGCCOACTGG CCCCAGTCCCAGCTTGGCTGGTACCCACTGCCACCCAAA GGTAACAGGTGTGGTGGTGGCAGAGTGGGGGG GGCAGAGCCCTCCACCTTAATCCCCCT	198

FIG. 17BB

272	chr8	145109784	145110001	GCAGGGGAGATCGGCCCAAGGTGCTGAATCTGAAAC AGAGATGGTGCAGACTGGCAGGGGAGCCACAGGGGCA GAGGGGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGG TGAAAGGGTGAAGGGGGGGGGGGGGGGGGGGGGGGGG CGTCCTGCATGCCACCCAGGTGCACAGGGATGTGGGG CATTGGACACCAGCCCC	218
273	chr8	145689504	145689731	CAACTGTCACTCCCACTGGAGGGCCTGTGTTATGCAGGGCC GGCAGTAAGGCCAGCTGCCTACAGCCAAGGCCAGCGACAG ACATCCTCTTCGGTAGCAGCAGCAGCAGCCTGTCCTCCAGC CGCTCAGGGGCTGCCACACCCAGCATGCCACCCCCCA CATCCCATAAGTGTGCCGTGAAGACAGCTACAATGAGGATG GTCACTGGACCAACAG	228
274	chr8	145735721	145735941	CATGGCCCCACCAAGGTGAGCAGGGTGGCCAGCAG CCTGTTCTCTCCTCCCTGCTGGACCAAGGGGTTCTGCTCC CGAGTTGGACTGTGGAAATGGTGGGGTGGCTGTGCTGCT CCATCGCTGGCTCCTCCCTGGGTGGGACCTTGCTGGCCAAAG CACTGGTGGCCCTCCCAAGTGTACCCCTGCCACCCACTTGT ACTTGAGCTGCC	221
275	chr9	3330393	3330600	CCACCATACTGTGGATGAGACCAAGGTAGTCACCTGGGG AACTCCCTTGAGTCAAAGTAATTCCCTCCAGTATTGGCTG TACATCTGTGTCTGTGAAGGATAACGTTGTTGGCTTTA GAGAAGAAAAGAAAATTACTAGCTTATTGTCACTTT ATTAAATTTTCTAAATGAAATGTAATTGAAA	208
276	chr9	36101901	36102125	AGACAAACAGAAAGTTAGGGGTGCTTCAAATTATTTCACCT GAGCTCCCTGTGACTCTTGTCTCTTTGAAACATGTTAAGTT ATGAAAGCTTACTTTCTGGAAAGCTTCAGAAACAAGTGATGTT GACTGGACAAACATGAGGAGGTCCCTCAAGCTAAACTTACGT GCATTTTTTCAAGATCAGATTGTGGAGATTCTAAAAAA TGTGGAG	225

FIG. 17BC

1700

277	chr9	36227107	36227287	TCCACAAGTCAAACTATCAGCAAATGATTACAAGCTCTTGTGT ATGTTTCAGTCTAACCTTCCAACATAACTATAAAATTTAAAAAAAT CAATCGCTCATGGATAAAAGTAGGAGTTAGGAGTTAGGAGTTATT TTACCTTCATGCTGAATTGCAACTCGGAGGTTCGTCCCCGCC AAGA	181
278	chr9	36881902	36882116	CAGAAGCGTAGAGGTACCCAGGCTGTTGGGGGGATGT CCCCCACCGAAACCCCGTCCGCTGCCTGCTGTGAGACGC CGACAGTGCAAAACTCACCAAGGACCATCCCTGTCAAGCTCG GTGCTGAGTAGCTGCCCTGTCCAGGGGGGACGTGTGGA GGTACCCGGGGAGGGTGTGCTGCCAAGTCACGGCCTGA GGAAATCAAAG	215
279	chr9	79917654	79917874	ACATCACTCAACTCAACTTTCAGAGATAAAAAAGTCTGATAG ATATTTCCCTTATACCCAAACAATTATTGAAITTCCTGCTATTGC ATATAAAAGGATAAAATTAAAGTTAACTTTCCTTAACCTGAAAT GTATATGAAACAAATTAAAAGAAATGTTATCTCTACTTAAAGATCTC GATTATTAAATGATGCATACCAAGGAAGTACTGGATCTAC	221
280	chr9	123249547	123249769	AATGTCACATTTCAGTCAAGCTTACCTTAGAGAAAGATGGTTTG CTGCCTGGTGGCTGTTGAGAAGAGCACTTTCTGTATTAAAGC CTTCACACTCTGAGCTCAAGATCCTCTGCAGCCATTAAATAGCA GTTCCTGGCTTACCTCCAAATCTTTCATTGAAATTCTAAATG AAATAAAACAAATTCTAAGAAAATTGTCAAGAAAATTA TAA	223
281	chr9	125618020	125618238	AGCATTGAAATTAGCAAATGACATCTTGGGCTCTGAATCAAATGG CATTATAACAGGAGCCTATGGTGCACAAAGACCCAGGGAGAATA ATGGGAAAGATGATCTCTCAATTTCGACACTCACACTCT CTGAGAACTGGTTAAAAAAATAATAAAAGCCAAATTAGTAAGA AGCTTAAATTACAAGAGCTTGTAACTAATAAAAGTAACTATTAT	219

FIG. 17BD

1700

282	chr9	140127351	140127552	CAACGGGGTCAGGATGGTGGCTAAGCGTGGGCCACACT CCCTCCCCGGTGGGGGGGGCAGGGGGGGGGCAGGC TGACTCAGCCCCCACCAGCAGTGGTGA CTGTGCCATCATGGTCAACGTAAGCAGTCAGGGACCG CAGCACCCCTGGTCAATGGCAGTCAGGGACCG	202
283	chrX	19433063	19433254	GAGCTGATCTACTTACTGATTCACTAAGATTGATGTTAT CCAACATTAAATTACTGTACAGTTTCTCAGTCAGAGTAT TCAATAGTTAAATAGCACCATTCCCTAGTTTTCTTAATCTTA AAAAAAATGGATTAGCATATCTTACTGGAAATCTGAGTCCAT TAGTGACTTCATT	192
284	chrX	37669980	37670201	CTACCTGCCTTTGAGACATCTCATCCC TTTTTTCTTCCAAAAGTACCAAGAATAGGAGTTTCTCTG TGACCTGAAGCCCTGGCTGAAACCCCTGAGTAAACAAAGCAT CTCCAACCTGAGCTGGCCCTGGGGAGTGCA AACAGGAAA ACTCTAACCTGCTCTCCATGAGGAATAAAT GTGGG	222
285	chrX	46884187	46884379	TCCAGCCCCACACAGAGCCCTGGCTACATCAACATCATG GAGTTGGCAGCATCTGTTGCCGCTATGACATGG ACATCTCTGGCTTCAGGAACCTCAATGAAGACCTTGAGAAAT GGTAAGTTCTCACAAAGTAAGTTGACTTAATT TTTTTTAAGTTACAGAA	193
286	chrX	47034374	47034557	CTGTTGACTAAACCCACTGCCTCTGATCTGAATGCTGTC GGTCAGAGCCCCGGCTTCGGCTCGAGTTAGTC CAGGACCTACACGATGGATGGAAGCCAA CCGCACCTGAACCCCCCCCCAAACAAAT AGCGCTCCCCATCG	184

FIG. 17BE

287	chrX	84563171	84563344	TGCTAAATGTGGACAGCACTAGTCCGAGTGACTTATCAGACAT ACCCCTCATCTGCAGGAAGAAAGGATAATGGTTAGTTTATT AACATTGAAATAAGCAATAAGTGAAGTTTTTTCCTA CTAGCAAACCTTAAATAATTACTTTGATAATGGCAAAATTA	174
288	chrX	100486597	100486813	GCACTGGCTATGCCATCCCCATGAGGCCTGGTTTAAG CAACCTATGGTCATGCAGGGATGCCCTAACACCCCTCCACGAT GTCTATGACTGGCAGGCAGTGAACCTCCATAGCAGCA GCCTCCGAAGCACTGCCAACCCCTCAGTAAGAGTCTGAT AGTGAAGAAAAGACCTGGAGAAAGTGGACACAGGATCTGT GT GTC	217
289	chr1	2422738	2422930	CTCCACACCTGCCCCATCTGGAAAGCTCGGACGCGAAGGTTAGA GGCCAAAAAGGTGACACCCCTGATGCCACAGGGCCCCCG ACAGCCAGCCCTGGCCAGGCCAGGGCCCCGAAACGTATGC TC TGTGCCGCACTCACACCTGTGCACACAAAATCTGCACACA AGCATACCTCTGTACACACCTCCGCAC	193
290	chr1	8073667	8073860	CGCGAGTTACTCGGTGACAAGGTCTCTGGGGTACTTTTG GGAGGCCCTGTCTTCACTCACTATAGTGCTCGAAGTAACCTCTG CTGACCACCTCTATAATCTGGCTTACTGGCTAGGAGGTATG GGAACCTCTGGGGAACCTCAGGTTGTCTTCATCGGAGTTA GGAGAAGCTCTGTGTTACAAACA	194
291	chr1	19651998	19652184	GCTGTGTATTATGTCTGGCAGACCTGGTATGCTGACGCTGT ACTTTTACATACAAGTTCAAGCAGCGCCCTCTGTGTGAGTA TGGGGACCCGCTCTGTCAAGTGCCTACACAGCAGGGGG GCAGTTGGGGGGGGGGCTGGAGTGTGGAGGAAGCG GGTTTCCCCTGGCTGGGTGA	187
292	chr1	46546310	46546483	TTACCTTGAATAATCCCCCAGTACCATCTGCATCCCTGAAGA GAAACAGAAACTGTCTTCATCCATTGGAACCTGCTGAAGTCA TTGGCTTAGGTGGCTTGGAGGCTAGAAGAGAAATGA ATATTACTCTGTAGATGTAATTATTACTGTTCTAATTTCCTC	174

FIG. 17BF

1700 ↘

293	chr1	156708309	156708501	CTCACCGTGTCCCCACAGAACAGATAACCACTGCTTCATCAGAGA TGGTTCCACAACCAACTGGGCCCTGATCCATGGGGG TTCTTCCTCTTGTCTGCAACACAGGCCTGGGGGCTCATCCCA TAGCCGTAATGGGGGAGACTGACCTTGGATGCCAAAGSC CAGCAGGGCAGAAAGACGGCATGCCTG	193
294	chr1	214566980	214567183	GGACATTAACTTAATCTGACAAAGCCATCAAGTCGGGGGAT TAAATGTTTACTTACTCAGTGCAGATTGTTTGTAAAAAC TTGTGTGCGTGCGAAACCAACCGAGAAATATACTTGGCATT CGATATCATCCTGGAAAAGAAGAGACAGAGAAAATAAAC ATTGGTGAATTAGTTATAACATTCAATTAA	204
295	chr1	227834200	227834409	CCATGGCACATTGTAAATAACGTGTGTATTTCCCCCCA GGGACTACTGGCATTCACTGATGTGGTCAATAGAATTCTCTCCA GAGGAGTGGGATGCGCTGGACCCGGCCAGGGAAATTGTAT AGGGATGTGATGTTGAGAACTACAGAAACCTGGTCTCCCTG GGTGAGGATAACTTCAATACACAATTATGTATTTCATA	210
296	chr1	228596592	228596794	GTTTACCTGATGCCCTCCACACAAATTGGTGTGGGGCTTGG TGCAATAAGGTACAGGGGCTGGCAGAGACACTGG AACTCAACAAAGATTCACTCCCATGAATTAGAACCCAGTAACCG CCGGCCCTACGAAGGGCACAGTGAATAGCCAGTGACCGTCACCA GGCAGTGCACCTGTAAGACCCCCCATTGGTGG	203
297	chr10	97515894	97516118	TCTCTGTTCTTGGAGGACCTCTCTCACGGAGACGGACCACA GCAAGCAGGGCTGGGGGGAAAGACGAGGAAGAGGA GAAAACAAAAGCTGCTACTTATGGAAGATAACAAGGGTAAGA CCAAAATTGATTGATCTGAATCCTTAAGAAAAAAATAGAA GGAAAATAAAAGCCCAGACCAAAAGCAAGTACTTGAAGGCT GGAGGCAAAAGC	225

FIG. 17BG

1700

298	chr10	124753930	124754127	GGTTTCTGAACCACCATGGAAAGGTGGACTTAAGTTGATTAGTG CTCTTCTGCTATAGCCCCAGATTGCTTGTGTTCTCTGTTAACCC CCCCACATTTTCTGCTTAAGGAAGACCTAGTACCTTAAAT TGGTACCAATTATGCTTGGCCTTCGATGATGGGACAAGTTA CTTCGATCACTGCAGCGGGAAAGGAA	198
299	chr11	18326808	18326983	CTTTTAGATAACAGAACACTTCATGTAACATAATTATTTCCTAACT TTCTATAAGTAATTAGTGAaaaaATAACCTAAAAACTACTTC TGAAACAAGAAACAGAGATAAAAATCTAAGTATTCTTTCCCCC CCAGAAACTAAGGAAGTAGAAACTTACCTAAGATGTAAGAGT	176
300	chr11	70281105	70281298	CATCTTTCTGTGTTCTCTCCAGGGGGGATGATGAGAT CTCATTTGACCCCTGATGACATCATCACCAACATCGAGATGATT GACGACCGSCTGGCGGGGGTGTGCAAGGGCGGGTACCG GGCTCTTCCCAGCCAACATGTGGAGCTGGCAGTAGGGC CCCCAGCCCCCCCCGGAGCTGGCCCC	194
301	chr12	58024954	58025154	GGAGTCATGGCTTCAAGCTGGTACCTCGACAGAAAGG CCTGGAAACTCCTGCTCTCTGGCAGAGGCAGGCCCTCAGCT CTGCAGGGGTCAAAGGCCCTGGTAGGGTCAAATAGCTCGGACTT GTTCTGGAAAGGGGAGGGGGAGCCCCCACTGGACTCA CAACTGCAGTTGTTCAAGCCAGCAGCCTGAAGG	201
302	chr13	111303318	111303538	GTCCTGTCAGAATGACTGTCTCCGAGCTGGGGGGAGTCC ACTCCACCGTGTGCTCGGCTCACCTGAGCGGTAGCTGCTCCG CAGGCAGAAGAACCCGGAAAGACATCGGGAAAAGGTCTTCA GAAAGTCCCTGGTAAGCGAGAGACAGGCAGTCAAGGAGGCTT GCCTTTACGCTTCGCACTGTTAGAATATGTTACTTTCTACA ATTACATAAGTT	221

FIG. 17BH

1700 ↘

303	chr14	102378807	102378991	CCAGGTACTAAAAAGAGATAACATGAAAACGCCAGGGTTA CTGAATGTTTATAAGATAAGGAAATATATGTCCTCACCATGGG GGGGTCTCGATTCACTAACCGTTGTTATGAAATGTCCTGCAAT AAAAGACTACTTTAACTTGTAAACCCCTCCTTTGTACAAGTAC TTATTTTGC	185
304	chr16	677573	677776	CCAGAGCCCCCCCAGAACTGCTCGGGAGTAAC TGCAAG ATCTCCTAGCGGGATGGGGGGGGCTGTGCAGATGCC AGGAGGGCTCGAGCTGGACACTCTGGCTGGACCCAGGCC AGTGGCCGCCCTACGTGGAGACTGTCCACACAGCTGCCTCAGAA GGGGGGGCTTCTCACACACTGAGCCGGTGGAGGAG	204
305	chr16	30129715	30129906	ATCTCGGTGGAGCACGTTGGGAGTGATGTACTTGAGGCC CCGCAGGAACTGGTAGAGGAAGTAGCAGATGGTCATTGCT CAGCTGCTGGCTTTCAAGCAACTTGACAGGTCAAGTCCATC AGTCCTGCACAAATGTAGCTGAGGATGGTTCGGAGCCCC CCAGGAGGGGAGTGGAGGG	192
306	chr17	30791457	30791651	AACAGTTCTGTGACTGTCAATCTTGATTTCCTGTTGGCTTT TTCCCCCCCCCTCAATCAATGCCCTTCTCTCTCTAGGCCAAG ACTGGTGTCTTGTAACCTTGATACCAAGAGGTACCAAGGAAGCA TTGCATTGGTAAGTAAAGCTGGTAATAAGTCATCTCAGCTAGC TAAATGAATGTATTCTAAAC	195
307	chr17	33318683	33318870	GAGACAGGGTGAACGTGTCAAGAGACAATCAGAGTCCTTGA GCAGAGCAAGTCTTCCCCCAGCTGCCAAGAGCCTCCTAC CATCCAGGAAGTGGATGAGTTCTCTGCGGCTGTCCAAGCT CACCAAGGAGGATGAGGAGAACAGGGCCTACAGGAGATTGC CTCCAGGGTGGGGAGGCTGCC	188

FIG. 17B |

308	chr17	42991418	42991604	TTCCTGGCTTCCAGGCTCAGGTTGGTTCATCTGGAGGCC GAGTGGGGGACACATTCTGGGTCCAGGCTCTTGAGGA CACTTGAATAACCTGCCTCAGTCTCCCTGAGGCAGCTGTCAC AGAGACCACCCCACCGAGGACAGTAGGAGCAGGAGGA TTAAGGGTTGGGGGGCTTAG	187
309	chr17	46688011	46688219	CGCGCAGTGCATGTTGAAGGAACACTGGCTCGAGCCCCATAGCC GGCCCGGTAGACGCCGGCGCGCTCTGGCCCGCCATGCC CCCCCGCCGGGGTACAAGGCCCTGCATCGAGGGGGCGAAGGA AGCGCCCGAACCCGCTCCATAAGCCCCGGCTGGGGCTGGGGTAG AAGCAAACGCACAAAGAAGTTTGTCTGGGAAGGCTCCGGTAG CGA	209
310	chr19	8136821	8137010	AGCCTGGCTCTGGGGGCTGACATGACTAAAGTCG GGTCCACAGCCTTCCACACTCACCCATTCCCTGGCCAC CCCTGCCCTAGTGCAGCACCCATGTCACTGGTGGTCCCTGTG GGCACTGCCTCGTGGCCGGTCCCAGGGAGGGCCATTGA TCTTGCATTCGTAGGGCTCA	190
311	chr19	10079936	10080166	GCTAGGCTGCACCAATGCTCACTAGGTGGTCACTAGGTGCC AGAGTGGCTCAGGCCAGAACCTCAAGGATGAACAAGTCT CTTACCAAGGGGCCAGGGGGCCATCTGTCAGGGGCTCC CAGGAGGGCTGGTCAACCTGAGAATGGAACAGAACATGGG GTGGGCTGCACCCCTGTCAAGAACACAGTGTGGACCCCC GACCATGCACCAAGGGCAGCCCC	231
312	chr19	49850421	49850641	CTTTGATTTGCTGTGAGAACCTCACCCAGAACATTGACCGGA AGAAGGCATGGGGGGCCACGATCATATAAGCTCTGGAGG CCACCCCTTTCTCAGGGATTGGTGTAGATCTGCCGGAGCT CCACACTCTCGAGGGGGGGCTGGGGCTGGGGCAGTGC CTGGCTGATGTGCACGAACAGGTGCCTCTGGTACTGAGGAG GGTGGCCGTGGG	221

FIG. 17BJ

1700 ↘

313	chr19	50862171	50862371	TTCTGTGCCCTCCCTAGTGACGCTCTCAGCTCTA CCCTCTCAAACCGTCATATGACGTCACTATGGCGTCA TGGCACCTACCTGGATGACGTAATCATGGGCCGT AACAGACCCCCAAGAAGGAAGGAGACTGGGGAGCTT TGGGATTCCGAGCACAGGATGATGTACTGGG	201
314	chr2	10140691	10140871	AGTTGGTTGGTCTGTCCCCCCCCATCCAGGATCCTGGTGAA CATGGACCGACAACATTGTGAAGCATTACTCCAATGAGGACACC TTCCAGCTGCAGATTGAAGGAAGCCCCGGGGTCTTACAAGGCTC ACCTGACGGAGATCTAAAGGCCTGGGCCACAGCTCCCC AGGAGTTCACTGC	181
315	chr2	61492544	61492746	GACAGCCTGTCTGCCTCAGGTATCATAAAGTTGCTGAATA GTAGAAGCTAAGTAACAAGTAGCTCCAAGGTTAGTAAGGCCAA CAAATCTACATTCAAGCAGACATCTCATGAGGOCAGTAATC CCATTATAAGGTGCATGGCTGCTGAAGAAAAGGGAGGAA AAGATCCCCAACAAACTATGAAACACAA	203
316	chr2	69409677	69409851	CAAAGAATGCAAGAGTCAGATGCGGGAGGAAATGATT CCCTGAGCCGGCGAAATCTAACAAACAATATGCGTGGCCTTC TTCCCCCGGAAGTGGTACTCTCCAATCAAGGTGTGTCTCTT ACTCAAAAATCTATCATCAGTCTGATAAAACACTTACATAG GAA	175
317	chr2	74763865	74764061	TCTTCTGGCCACTGGATGCCGTAGACAGGGCTGGCACA CAGCTCACCACACTGCAGGGGGGGGGGGGGGGGGGG GTCATTGGCACGGATAGAACCTCCAGGGAACAGAGGGAGGG GGGCCTCCGTGCCACGGCACGCCACCCATGCAGACCAAAAG GAGTGTGCTGCCGTGGGCTAGCAGGCCAGGG	197

FIG. 17BK

1700 ↘

318	chr2	166221500	166221712	CTTTACCAAAGGGATTGGCATTATGTTAAGTCTTAAATTACA GATCAAGAAAATGCATAACAGAACATGGGGGGCACACCT ATTAAATTTTATATTAGATTAAAGAAAATAATTAAATGTGTTTT TTTGGGGATTGATTTCAGAAAGCTAAATGCAACTAGTTCATCTG AAGGCAGCACGGTTGATATTGGAGCTCCCGCCGAG	213
319	chr2	242716342	242716529	TGTCCCCCTCGCTGGAGGCCAGAGGCCAAGATGATGTCATGC TGGCGGGCTTGGAGGGTAAGGGGGCAGTTGGACCCCCA GGTGGACAAAGGCCCTCATCTGGCTATTCTGAGAGACGGA CAGGAGTGGGTTTGTGTTGCTTTCGAAGGAAGTGACTTGA TTTGGAGGCGTGTCTACICA	188
320	chr20	33502059	33502245	TGGGGGGCTTGGCAAGTCCCTTGGCCAAACAAAGGCTAC CACTTTAGGGCTTCTCTCCAGGGGTTCCAGCTTAAGATG CTGCATTGGTCAAGCACCTGGGGAGACGCTGGCA TGSGTGACTCCACCAAGCCAGTCCCCCAATTAAAGGGTCAT GCCAGATGTGCAGGTGAG	187
321	chr20	46265018	46265189	GATGACGGGGGTCAATTCCCTTGACCAACTCCCCCTAGATT CAAGTTGAAAGAATCTCTGTTAGTGTCAACCAGCCCCCTGG AGTCTCCCTCATCTGGAGGAGTACATCTACATCCATA TGCATGGTCACTGTTACAAAGAGAACGGGATTTCGCACA	172
322	chr20	46291740	46291967	GAGGTGGGGGCCTACCTGGCAGCCTGGTACCTGGTACCCCA CACCCCCATCCCTAGCTGGGTGCTGTACCTGGTAGGG GTCGTGTTGAGATCAAAGTACTCTAGGAAGGCCAGTTGCAAAT TCACAGAAGAGGAATTGTGAGTCATTGATGGTCTCATGC ACCACTACGTGTTATTGTTGGCGCTGGTCAGGCACAGAAAG GCCCACTGCCAAGGAA	228

FIG. 17BL

1700 ↘

323	chr22	43572221	43572424	CGTGCCACGTGCCAAGCACTGTTCTGGGGCTGTGGGAACCA CGAAGCATGGGAGGGGGACTGGGGACAAGTCCCAGCCCA AGCAGTGGGGGGCTACGGCACCTGAAGATGTGGCGTGG GGTGCAACACGGGGTGTGATGTCAAGTGGCAGCTCTCTG TTTCCTCAGAATGGCTAAAGAAACACCCAGGAAG	204
324	chr3	9958898	9959130	CAGCCCCCTGCCACAGACACGGCTGA GCCCCCTGGGGGGGGCAGCACAGGGCCTCAGGCCTGG GTGCCACCTGGCACCTAGAAGATGCCCTGGCCCTGGTTCTG CTGTCCTTGGCACTGGCCCGAAGGCCAGTGGCCTTCTCTG GAGAGGCCTGGGGCCTCAGGACGCTACCCACTGCTCTCC GGTAGAGTGGAAACCCCTGGGAGACGA	233
325	chr3	12779951	12780167	CAGGTATCCAACAGGGCTGGCTGGCTCTCTGGGGAGGGTA TGCTTAGACCTGGCTCAATTACGACCCAACGGACCTCTCC TCTAACTGGACACTCACCTGATTCCCCAGAGGTACCCACT CCAGGAGCATCTGCACAGAAAGGAATCAGTGATCATATAACC ATGATCCTCTGGGGGAGAAGTCATGCTTCAGGGAAAGGT GGATCC	217
326	chr3	37778324	37778548	CAGAAGATGGAAAGAGAACAGTTAAGTTGGAATGCCTGCTGA TGCCAAATAACCACTCCCTTTCACTAGTAGTGAATTCAAGCGTGA TCCTTGATAACAAGGCCACCTGCTGGGGAAAGGAAGGTCTCA GCTTCATTGTTACTGCTCAGGGTAAGGGGGTTAAACA CTTTTTCAAAGATGAGAGATTATGGGACCAAGGAAGGTGT GTGCTGGTTAA	225
327	chr3	100712018	100712211	ATTTGAACCTTGAATCTCTGCAGAATAACTCTTACCCCGTA TGGTAAAGCAGATCATTTAAATCAACATTTAAACATCTAT ATCTTAATTTTAAACATCAAACATCAGGGCTACTTGGCTT TTGGCAATTCTGTGCATTCCAGGGCTAGTGTAATACTTCCA CAGAGTAGACAC	194

FIG. 17BM

1700

328	chr3	108402895	108403109	ATATGTTCAATTTCGACATACTTGTCTAAATGTCAGCTTTGTAT CTTGAGTTCAAATGAGTGTCCCCCCCAATTGCTGTATGA ATTAGATGATTGCCTAAATAAATATAAAATGTTGGGGGG CTGTTGTTATTATTGTTTATTTCTATTAGATGCCAGCTT CTAGGAAGACCTCTTGTGAAAGAACCTTCCTTAG	215
329	chr3	108822625	108822822	ATTATTATTGTAAGGTTTCAGACTGTTCCCTTAACAGCCA GTTAGGTATCTAGATAATTAAACTTACCAAGGGCTCATGCCACATC CATCATCCAGAAACACAACATGAATCCCCCTGCAAGTTTC ATTATCCACTGTAAGGAGAACAGTTATTACTAATAATCGGG CTTAATTGTTAGTCAGAAAG	198
330	chr4	114280067	114280259	GCCACAAAGCCAAGATACTTACATCCCATTGCCAGTTAAG AGCAGAACGCACTACATCTTCCCTGCAAGGGGGCACGAGCCC CACAAAAGAAAGTAAGGAGCATTCTTACAGAAATT CCATAGAAATTCTTGGAGGATTAGTGTATGAGGCTTCCAAATTAA GTGGATAGGCTGACACAGTC	193
331	chr4	154625988	154626169	ATTGGTTTCTTACAGTGAGGGGATGCCIACTGGGGAGAAA CCTTATGGTCCAGGAGCTGGAGAACTTCATTCCCTGGCAAGTGGATC GTGTGTCTTCATAAGGGGGACTTCATTCCCTGGCAAGTGGATC ATTGACAATATCATTGACTCCATTGAAAAGGCCACAAAACITG CTTGTGCTT	182
332	chr5	1501518	1501746	CACTGTTGGCCGGTGACCCCCCCAGGGAGGCCG GAACACCGGCAACTTACTTCCTCCACAGGGGGGGCTGC TCGGGGTCTTCTCCGAGGCCAGGGCATCATGGCAGGGCAACAGGA GAGGGGGCAGGCCAGCAGCATGTCAGTGTCACTGAGGGCCACCTGC GCCGGACGGGAAGAGGCTCAGTGTCACTGAGGGCCATTCCAGA AGACAGAGGGGGCATTCCAGA	229

FIG. 17BN

1700

333	chr5	135391356	135391545	CTCTTCTGGCAGAGGCTCTGGCATTTGAGAACAGCTGCATCGC GGCCCCACGACAAGAGGGGGAGGTACGGGACCCCTGTTCACCGA TGACCCGGGTGCTGACCCCCAATGGGGACTGTCATGGAT GTCTGAAGGGAGACAATCGCTTAGTAATTAGTTCCATCCC CGGGTGGAGCTCTGCCCAAGTGG	190
334	chr6	31691090	31691292	CCCTCCGGTCCCCCCCCAACCGGGGGCTGATTGCG CTCGCAGGCTCTCTCACCACATCTAACCATGGCTGTGTTTC TGCAAGCTGCTACCGGCTGCTGCTCGAGGGCCCAGGGAAC CCTGGGGTAAGCGATCCCTGGAGAGTTGTGATAAGACGCA GAGGGGCTGAAGCAAGATAAGGGGGCCCTAGTAGGGT	203
335	chr6	32948162	32948375	GGAGAAAAAGCGGGAAATTAGAAAAGCGGTACAAAGATGTCAG CGGACAGCTCAATTCTACTAAAAAGCCCCCAAGAAAGGTGA GTATATACTTICATGCCACTACAGATTGACTCCATCTGCCCTC TTGACTGTCTTTTATTGACAAATGAAGATTCAAGCTGAACGTC TTAACCTTCGAATTGTTCTGCAGGAATGAGAAAACAGAG	214
336	chr6	111213883	111214095	CAATTGAAGTTATTTGCAAAACCTTTGTTAAACTCGGCTTT TCCCCCCCCAGGAGGTGGAAATTCTGGTACCTGATAACCGTT CTGTCATTGATGCCACAATGTTCAATCCTTGGGTATTGATG AATGGAAATGAATCGGAATGTGAGIAGTTATATTGCTTCAAATA TTATTAAATGTGAATAGTCCTTCCATTCTTAACACTG	213
337	chr6	111693783	111693988	AAAAAATAAACTGATTACCTGGCTTTCTGGTACATCGAAAGGAT TACTGCAAAATGGTTCCCTGGTAATAATAGTCTCAGACAGGTCTG ATCCAACAAAGTTGCCATAATTCTCCCTACTTGGGGGGAC ATAAGTGGTTCAAGAATGTTAGCAGTTGGAGAGCTAGTAGATC CCATTTTAGATTGTGAAGGAGAGCTAGTAGATC	206

FIG. 17B0

1700

338	chr7	33014681	33014859	GTGACTGCTTGAACCCCTTTGGCCCCACATTCAAGCTGGGAGC ATCTAGGTGTTTATTATCTGGCCTGAAAGGCAGCAACTGTG AACTCAGTGTAAATCTTAATTCTAGCTGGTGTGATCCCCC CAATTCAAGTGCTTCATTGATGTACTCTGATGGATATTGGA ATTCTG	179
339	chr7	102113413	102113594	AACACAGTGGTGGCAATGCCTCCACCTCACCTCACGCC CTGACGGGAACCTCCACATCGTAGCCCATCTGGGGAGGATGTAG CCTCACACCATCGCAAAGGACACAGGGACACAGCTAACTAACT TATTCAAGCTTGGCCGATGGGGTGGGGGGGGGGGGGGGGGG GTGAATATTTTTATT	182
340	chr8	42761281	42761484	GAATCTCCCCCCCCCTTTTTCAGATCCAGAC ATTAAATGGGATGCACCATGCCAGACCGTGGAAAGATTTAACCG AATGGGGTTGTGATGGCCAGGTTCTCAAAAGATGTAAGAAT AAATTGCTTATCTTATGTGTTATCTTATGCATGGATATTGTTAG GTTGAAAACATGGATTAAAAAAAGTACTTA	204
341	chr8	85799811	85800039	GIAAGTCATTGCTTCTCCCCCCCCAGAAGCTCAGAAGAAAG CAATTGGAAAGAGCTAGTGCTGATCCAAGAGGAATGTG TCAGAGATTGAGATCACTTACAGAGGGCCTGCTGAAGGA GGGCCAGATGCCGATGGAGAAGAGATGACAGATGGTAGGAAAG GAGGAAGTTCGATGAAGATGGGGTCAATGAGCTGGTAGGAAAG AAACATTTGGTATTACAC	229
342	chr8	142516975	142517192	AAGCTTCCGGCCCTCTGACCTCCCTCCTCCAAACCTGATCCC CACTCCAACCCCTGCAGGGCTGCTCTGGGGCAGGACCCCTC CTTCCTCAGCACAGGGCCCTGATGGCCTGCAAGGGAGGC TCCCCTGACAGATGAGGAGGCTGAGTACCGGTGGGGGGGG TCAGGGGGGTCCCTGCTGGGGTGAAGCTGACCTGGAGTTTG GGACACGGCTGA	218

FIG. 17BP

1700

343	chrX	48819643	48819855	GTCTGGGGAGGTTCTCATTCATGCCATTCTGGCTACTTCCC CACTATGGGACCATTTCATGAAACTCCATTGCATAGTTCTCAG TGGGGGGGGAGAAGGGTGCCCAAGCCTCCCCCTCTCTG CCTCCCAAAATATGGCTTGAAAGGAGTTCCAAAATGAA GAGGACCCCTCTGAATCAGCAGGCCAACGGGATGA GA	213
344	chrX	53426626	53426800	CGAGTCTTCTGATTCTCAAACCTCCAAACTCTGTTATAAAGGTGG GGGAGAATCACTCATAATCTCATCAGCCAGAGATAACCAATG TACAAAATCCAACATGGGAAGAAGGGAAACTAGACCCAGG GCTCAGGGACCCAGAGTTACTGATTTCAGTCACAACTTGCTT CATC	175
345	chrX	106486216	106486436	CAGAAAAACTGGTATACTGTTGTGTTATTAGGCATATTGGTA TCCAATTACACTATCTAAATATTGGGGGGTTGTTTCT AACACTCTAGTGGCATATGTTGTTGCAATTGCTTATTGAAAG TGATACATCATTAGGAAGCTAAGGTTTATTGGTTGTTT TTAGGAAGCTGTTGATAACTCTTCCTGAGCTGGTGAATGTA	221
346	chrX	107683304	107683523	CCAGCCGGGAATTTCGTCGGGGTGAAGGGAGCTGCGGGGA GCCGGAGAAGAATGAAACTGCCTGGAGTCAGCCTGGCTGCC GGCTTGTCTTACTGGCCCTGAGTCCTGGGGCAGGCTGCA GAGGCTGCCGTAAGTCCTCCCTCCCCGCCCATCA CCGCTCTTCACCGCTGAAATTACCTTTTGGGGGGGT CCCTTIACTT	220
347	chrX	144904007	144904188	CGCAAAACTGCCAAAGACATTGCAAGATCCGCTGTCTGTGC GAGAAAAAGAAAACGTACTGAATATCAACTGTGAGAACAAAG GATTACAAACAGTTAGCCTGCTCCAGCCCCCAGTATCGAAT CTATCAGCTTTCTCAATGGAAACCTCTGACAAGACTGTATC CAAACGAATT	182

FIG. 17BQ

1700

348	chr1	8029357	8029556	AGTTGAAATGGTTTGTGTTCTTATGTTAACTGTAA CAGGGACCATATGATGTGGTGTCTACCGAGGTAATCTG GGGCACAGAATTATCTGAGGTAAAAATTCTACTCAATTATA CCTCAATAAGCTGGGGGGGGAAACTAAAGAATTTCAG CATCTGCTTATGTTCTGTTAATTGT	200
349	chr1	31191420	31191618	TCCCGCTCTCCGCCCTCTGCCCCCCCCCCCCCCCCCC CCTGGGCCCCGGCTGGCTCCGGCTTGCGGCTCAGGCC CCGCCTCCAGGCCCTGGCCCGCCGCCGCCGCCGCC CCGCCCTGCCCTGGCTCGCCCCGCCGCCGCCGCC GGAAAGGCCCTGGAAACTTCCGGACAGCTTC C	199
350	chr1	173797405	173797581	ATGCTCAACTTGGACTTAGAGATTATCTTAAATGTTTCTT TTCCCCCCCCCATTAATCAGGCAAAACACATTCTGGTCCTA AGAGATTGATGGCTTGTCAAGTTATCATCCCCAGGATG AGTAATAGAAAATTCTGTATTATCTAAAGCTTCTTTGATG C	177
351	chr10	75560494	75560698	TGCTGGGGCTGGCAGCAAAGCTGGGTAACACCTCCCC AGGACCATGGCCCCCCCCCACCTGCTCTCCCACCTTCTT ATCCCAGAACCTCCCTCCTAGCTCTGGCTCAGAGTTGAGG GGTGGGGTATGTGGGTGGGGGGGGGGGGGGTTACCT CAGSCTCTGGGTGGAGGGGGGGCTCTGCCAGGCCA C	205
352	chr10	134751935	134752145	CAAGACACAGAAACTTGTGACTGTGGAGAGAACGCACAT GGAGTGTGCCCAAGTGAATGAACTCCCCGTCCCCCCCC GAGAAGGGGGTGGCTCAGAGCCAGGGCACGGCAGCTCA GGGAGGGCACAGAGGGCACGGCAGCTCAACTCACAGC AGCTCAGCACGCCACTCCTGTCTCCCTCTCAGTCTGACTC AGCAC	211

FIG. 17BR

353	chr11	1248099	1248300	GGGCAGCGGTGAGGCCACCCCTGGGGAGGG CCGGGCCACACAGTGTGACCTCCCCACACGGCCATGTCTGA CCTGGGCAGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGG CAGGAGAGCGGGGCCAGGGAGAACCCCGCTGTCTGC AGGGAGGAGCTGCCTTACAGGCCACTGGGCACTGGCTGGGA G	202
354	chr11	17035402	17035596	GTGTGGCCGGCGAGGCCAACGAGCCCCGGCG GGGGAGAGGAGAAGGGGGGGGGGGGGGGGGGGGGGGGGGG GGAGGGACCCCCCCCCCCCCGGCACAGGGTSGCAGGGGGGG GCCCACTCACTCAGTAAGTAGCTGGGCCTCTCC GTGAAGGCCCTCCAGGCCAGGGGGGGGGGGGGGGGGGGGG TACCAAGCTGCTACAGTTAGAGAGTGGCTCTCAGGGGGGG GGTCGGGGGAGCTGGTCTGTCTCTGCAGCTTCTCGAG GTATAACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG ACAAGGCTGACCCTTTCTCTGGCCTCCAGGAGGGCTGA GGAGGGGGGCTGGAGTCCGGGCTCACTAGGGCTGCTAAGGG C	195
355	chr11	133788925	133789132	TACCAAGCTGCTACAGTTAGAGAGTGGCTCTCAGGGGGGG GGTCGGGGGAGCTGGTCTGTCTCTGCAGCTTCTCGAG GTATAACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG ACAAGGCTGACCCTTTCTCTGGCCTCCAGGAGGGCTGA GGAGGGGGGCTGGAGTCCGGGCTCACTAGGGCTGCTAAGGG C	208
356	chr12	54763917	54764101	GGGGGGGGGGGGAAAGGGAGAGCATTTAGGGAGTGGC GGGAGGGGGGGGGGGCTGGGAGCCCCGGAGCCCTGGTCCC GGCTGGGGTGGGAGTTGAATGGGAGGGGGTTCAAGGGTTGA GGTCAGTAGGGGAGGGATGGTCCGGCAGTTGG GATGGGTCCCCGGGGTGAA	185
357	chr14	23858119	23858325	GTCCCTCCACTGGGCCACCTCCGAGTTGGCCTTGGCAGGGAC GGGCTGGCAGCTGGCCTGGCCTGTCTCCTCGTACTG CTCCCGAGGGTGCAGTCATGCCGGGACTGCAGTG CATGGGCCAGGGCTTCTCGCCTGGGAGGGGGGGGG ACCAGGAGGTGGAGGGACTCCCTGTGCCCATCTCTAGAT T	207

FIG. 17BS

1700 ↘

358	chr15	31521620	31521620	CTGGCAGTTGAAATTGTTTACCTTGTCCTTAAGATGAGCACTA CTTAGGCATAGCTGATGCCCTCTGGGGGGGGGGGGGGGG GTGGGTGTCCCAAATCTGATGGTGGAGCAGGGGATCCCT GTGGGTAGAGAACAGGGAGGCCTGTGCCGTGGGCTGG CTGGTGTACCCAGGAGACTCAG	191
359	chr16	57071034	57071248	GGAGGGGGCTAGAGCTGGCTGGGGTACCTCAGGCCACA TGCCCAACGGGCTGCCCTCTCGCTCCACAGGCCCTCAGTAAC AACGGGGCTTTCTGTGGCCGGGTGATTGTTGCTGAGGGC CGTGAGTGCCTGCTGGACCCCTGGCAGAGCTGCACATCAGGT GGGAGGCTCCCTAGACCAAGGTACCCATCCCCCCCACATCA TGCTCTCT	215
360	chr18	67068444	67068638	GCGGAATCCGGGGCGCAGGGCCGATCGGCTGGCCATG GCCTCCAACCTTAACGACATAGTCAGCAGGGCTACGTGAAAA TCGGCAGGAAAGCTTGGGTGAGTGGCTCGGCTTC CTCTTCCCGGGCCTCGTTCGGGGGGCTGGCTGCCTGGG GGGGGGGGCAGGGAGGGTGACCCCGGGGGG	195
361	chr19	7707321	7707505	GACCACGGGGCTGAGCGAGGGGGAGAAGGGCGGTCTTG CTGGACAGGGAGCATGACCTGTGGGGGGAGCTTCCGCACAT GCATATCGCAGATGTGTCAAAGTGCCTGGCACACGGGGACCG GATCCCCCCCCCGCCACTGTGGGCTGGTAGGGGGCT TGGGATCCCTGGCTGCCAA	185
362	chr19	16186691	16186910	ATGGGATGGCGTGGTGCATGCTATCGCAACCCCTGTCTACAGA TAAGGAAACTGATGCCAGAGAGAGACGGAGGACTCTTGTGA ATATTGGGGGGACCCGCCCTGCAACCTCGGGGAGCTGACCC CCCCCCCAGGCTGACCCGTTCTCGCTGTGCTCCCCCAGG TGAGGGAGGAGCTGACGGCACCTCCAGAAGAAACTAAAAGGG ACAGAGGACGAG	220

FIG. 17BT

1700

363	chr19	18857752	18857977	AGGCCTCAGGCCAGCCAGCCAGGGCTTCC CCCTGGGCTGAGGTGTGTTGGAGGGGGTCTCAGCC AGTGGCACTGCCCTTAACTCACCTCACGCCGACAGCTGCC CCTATGGCACCATGTACCTCTCACCAACCGGGACACCCAGCT GGAGAAGTCAGTGGCTGACACCCCCCCCCCTTC TTGGAAACAAACAA	226
364	chr19	41062904	41063133	CCCCCCCCCACTTCTGTGATGGCAGGTGCTGCAGGAGAAATT CTCAGAGTTGCCAGCGAGACAGGTATGCCAGGGGGAAAC GGCTGGCAGCTGTGAACCAAGATGGTGGATGAGCTGATCGAGT GTGGCCATACAGCAGGGCACCATGGCCGAGTGGAGGAC GGAATGACCGAGGCCCTGGCTGAGCTGGAGCTCATGG CACACGGCCAGCTGCTGGCTGGCC	230
365	chr19	41255232	41255447	CTGGAGGCCACCCATGGAAAGAGATCTAGGTGGGGAC TAAGTAGTCAGCCTGTTGATCTTGAGGACACGGGAACCCCC CCCCCACCCCAAGAGATCTGTGAAGTGCTAGGGAGAGCT GGGTGGTGGTGGCTGGGCATGGAGACCCCTACCGGCTCTGGT AGTTTCAGAAGCCCTTTCACATCTCTGCCCTGGGACCC GGGCCGGC	216
366	chr2	46583262	46583470	AAGCTGTCCCACCCCCCTTCAGTTGCTCTGAAAAC GAGTCGGAAAGGCCAGCTGACCAGCAGATGGACAACCTTGTAC CTGAAAGCCCTGGAGGGTTTCATTGCCGTGGTACCCAAAGAT GGGGACATGATCTTCTGTCAAGAAACATCAGCAAGTTCATGG GACTTACACAGGTGACACCCCTCCCTATCTCTTTCAAAG GGGCCGGC	209
367	chr2	85788378	85788609	AGCACGCCCTTCCCACAGAGGCCCTCC CCGGAGACACTGGCGTCTCCGCCGCCCTCTGAG ACAGGCCTCAGGAACTCTCCGCCGGGGGCT AAGCCTACCTGAGCTGGCGAGGTGGGAGGGAG GACACCCGCCATTGCTGCCGGAGGGAGGTGGGT CTGGCGTGAACGGAGGCCAGG	232

FIG. 17BU

1700

368	chr2	119748069	119748292	TAATCAGACAAAAGGTAGAGGTTGCTAAGGATAAAATGTCATA ATACCTAAGTTTATTTATCTCCCCCCCTTTTGTTTG ATCTTAGGACTTCAAGGCACAGCAAGGAAGAAAAGGAGAATCA GGAGTTCCAGGTAAGGGCAGGCTGCATCTCATCCCTCGGA GTGATGACGGGGAAAGGCCCTGCCCTCAGGGCAGGTCTT GTTGACATG	224
369	chr2	220494592	220494817	GAGGAGGAAACAAGGGGGGGGAGAAAACAGAGAAGAG CAGAGTGAAGAGGTAGGAGTCCCAGGAGAGAGTGA AGTAGAGAGGGTGGAGGAAAGGGTGGAGGAAAGGGGAGC CACGGGATGGGGAGGGGGCCTGGGTCAAGCATCCCTTGCTGC CCAGGAGGGCCCTGGGTCTCATGCTCAGGGTCTACTCCAGG CCGTGCTCTGAGCTGAGGGACGG	226
370	chr2	231120095	231120316	CTAAGTTTCTCTGATTCTCTGTTGTTGTTGTTGTTGTTG CTCTGAGGGATTCTCTCTGAGGGGAGTGTGATG TGTTCGGAAATGTTGAGGAGGAGGAGGAGGAGG AGCTCCCTAGCAAGACGTCGGTCAAGGTAAGGACGGGGGG GGATTCTGGCCCTGGGCTGCAGAGTGTCAAGGGGGTTGAATT TGAGGCTGTTG	222
371	chr3	47604085	47604309	CTGCTTAGGTTAAATTCTGAAGGACAGTCACATCCAAGTC AGCCTGGTCACCTTGGCACCCCTCAAGTTGGGGAGCATGG GTCTGGATGTTAAATGACTCCTCTTCAAGGAGACTTG AGAACCTCTGGATTGTGGGGAGCACTAGGATCATCCTGG AAGAGGGAAAAAGTTGGGGGGAGACAATGCATAACAGC AACCAAATAATAA	225

FIG. 17BV

1700

372	chr3	51975395	51975616	CCCTCCCACTATGGGAAAGAATAAGCTCACCTCGGCCCTTGGA AGTCGCTGGAGATCTCCTCATTCACCTGGCCACCCCTTGGAA TTTGCCTCTGTCCCCAGAGTCGGCTGGATTATAACGG TCTTTTCTTGGTCAAGGGTCCAGACCCCTGGATTGACC AAGGGCCCCCTCGGTTCCCTCAGAACCCAGGAATG GGAAGGGGGT	222
373	chr3	184100821	184101020	GGTGGCTGCCAACCTGACAGTCCAGGAGATGGACTGGCTGG TGCTGGGGAGCTGCAGATGGCCCTGGGACTGGCAGGCGAG GCCAGGGCTGCCATCAGTGGACACATTGCTGCCAGGAAGA GCTGGCGACGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGG GCATGCAACAAGTGGAGACACAGACAAGTGCAGGGTGG	200
374	chr3	195295592	195295797	GCAACAGCGAGGTCAAGCAACAAAGTTTATTTGCAAGCTAGCAAGGGT AACAGGGTAGGGCATGGTTACATGTTCAAGGTCAACCTCCCTTG TCGTGGTTGATTGTTGTTGCTTTATGGGGGGGGTAGGGGA AAGCGGAAGCAGAAAGTAACATGGAGTGGGGTGCAGGCCCTCCCTGT AGAACCTGGTTAACAGAGGCTGGGGCAGTTACCTG	206
375	chr4	843503	843720	GGCGTCATTCAAGCTCCATGAAGATCATCTGGCGTGTGCTC GTACGGCTGGCCCGCAGGCTATGGGTGACAGGGCGGTAAAG CGCCTCTTATAAGCATGCGGGGGGGGGGGCAGGACCCC CCCCCCCCCCCCCTGAAGGGCTCAAGCGGGCCCTGG CAGCTCTGCTCACCTGGGGTGCAGGCCAGGCC CGGCCGATAGTGTCT	218
376	chr4	56322103	56322271	CATGATAGTAATCATAGGCCTGATGTTCCCAGAACTTCAAATGGC AAATAACCCATTATGGGTGGTGCCTAAATTACACGAAACAA TCATTATTATAAGTTTCCATAATAAGAAATATTTCAACTAGAAA TAAACTTGGGGGGGGCTTTTCAAAAAAA	169

FIG. 17BW

1700

377	chr5	10263323	10263500	CAGTGGCATGAAATCCATCCAGACTATGACCGAAGTCCGAGC CAGACAGGTGAAGGAGATGAACCCCTGCTCTGGCATCGACTG TTTGCACAAAGGGGACAATGGTGAGGGAGCTGTCAACGCCCTG CGTGGAGGGGGATSTCTGATTAACTGAATAATCATCC ACTGTATGT	178
378	chr5	131326471	131326684	TTGCATACCCACCCACTTGCACCCACTCCACCCCCCTGCCACAG CCCCTCCACACCCCCCCCCACAACCCCCCTTCACCTGAGATGC TAAGCCCCACGGGGAAACACCTGGTACATGTCGGGGCATCAT CATAGTAGTGGTAAGTAGCTGAGGCCAGACCCAAATCACAG ATCGCCGTGCCCGGCCACTGTCTACAAGCCAAGGCCACGGC CAGAG	214
379	chr5	140803186	140803361	TCAGCTATCATTTAGCAAATGTGACCCGACAAGTAATCAGG TGAGATTATTCTCTGCCTCTTAATTGTTGGTGTCTGGCACA AGTCCTTTAAGGAGATGTTTTGAGGCCATATATTAGTTCATATA TGGGGGGGTGGGGCATATATTAGTTCATATA G	176
380	chr5	176831004	176831203	GAGACAGACTCTGGGAGGCCGTGCCGGCAGCTCAGTGG GCCGTTCTGGTCAAGGAAGGGGGCTGCTCCCGCTCGCCG GCAAGGGCTGGAGGGAGCAGGGGGCTGAGGAGGCC GGGGCGGGCTGGGGGAATCTAGCTGCCGGGCCCTC GGCTCCTCTCTCCCCCCCCACTCTAACCTCCCCGG G	200
381	chr7	99711301	99711513	ACCTCCCTCTCATAGAAGTAAAGCTCCCCCCCCACCA GAGGCCAGGGAAAGGAATGAACCTCCTGGGGCTGGAGGCC ATAGAGTGGCTGGCAGGAGAGGGCATGGGTAAGGAAG GGAGCCGGAGGAACCTGGCTGAAGGATGGCGTCTCACCC CCCCCGGCTGTCTCCCCAAATCACCTCGACATTCTTAGCTT CAAGG	213

FIG. 17BX

1700

382	chr9	85677168	85677387	TCTGACTCTTTCTTAACCTAGCGGAATGGAAAAGTCAA AGCGAGTAGGTGAAGGAAGACAAGCAACTCACAGAGAAAGACT AGGGGGGGGGCCATTGGGTCTGCACACCTGCAAGTAAGTG GGAAGAACGCAAGCAGGAAGGCCAGAATAACCTCCCCCCCC CAGCTCACACTCCGACCTGGGAATGAACTGGCTTCGTTCCCC AAGTCGGCCC	220
383	chr9	95874190	95874385	GGCAGGCCAGAGGGCGAGCTGAGTAACAGGGTAACACCTC CCTGCTGCCCTCCCTCTCTCCCTCTCTCTCTCTCTCTCT TACCCCTCCCCAGGGGCTCCATCTCCGCTCAGGGCTTCTCCA CCCCAAGTCTGGCTCCATTCTGGCCTCTGTGGTACAGA CCCCCCCCCTAAGGTGCTCGTTGGGG	196
384	chrX	153713886	153714114	AGAAGGCCCTTCTCCATTGTCAGTCACCTCAGGGTGCAGGGA AGCGGCTGGCCAACCGTTCGATGTCGTCAGGGTCAAGCGGA CTCAGGGACCTCTCGTAATCCTGGGGAGAGAAGGGCAGGAA GATTCAAGGGGCTGCCAGTGGGGGGCATACACTGCAAC AGAGGTAGGGGCTGGGACCAACTGGGCTGAAGAGGATGCC CAGGGTTACCCCTCTCACCCCTCT	229

FIG. 17BV

## CALIBRATION PANELS AND METHODS FOR DESIGNING THE SAME

### RELATED APPLICATIONS

[0001] This application is a continuation of U.S. application Ser. No. 17/811,192 filed Jul. 7, 2022, which is a continuation of U.S. application Ser. No. 14/975,001 filed Dec. 18, 2015, which claims priority to U.S. application No. 62/093,754 filed Dec. 18, 2014, which disclosures are herein incorporated by reference in their entirety.

### FIELD

[0002] This application generally relates to calibration panels and methods for designing the same. More specifically, the application relates to panels of amplicons for homopolymer calibration or recalibration for use with nucleic acid sequencing data and methods for preparing the same.

### BACKGROUND

[0003] Nucleic acid sequencing data may be obtained in various ways, including using next-generation sequencing systems, for example, the Ion PGM™ and Ion Proton™ systems implementing Ion Torrent™ sequencing technology (see, e.g., U.S. Pat. No. 7,948,015 and U.S. Pat. Appl. Publ. Nos. 2010/0137143, 2009/0026082, and 2010/0282617, each of which is incorporated by reference herein in its entirety). In some cases, such nucleic acid sequencing data may be processed and/or analyzed to obtain base calls using one or more calibration or recalibration processes. Such calibration or recalibration processes may be based on measurement values obtained for randomly selected subsets of nucleic acid templates undergoing sequencing. In some cases, a random selection of nucleic acid templates may result in subsets of nucleic acid templates that lack sufficient representation of long homopolymers. Thus, a desire exists for new and improved methods for designing or selecting sets of amplicons that improve calibration or recalibration processes.

### BRIEF DESCRIPTION OF THE DRAWINGS

[0004] The accompanying drawings, which are incorporated into and form a part of the specification, illustrate one or more exemplary embodiments and serve to explain the principles of various exemplary embodiments. The drawings are exemplary and explanatory only and are not to be construed as limiting or restrictive in any way.

[0005] FIG. 1 illustrates an exemplary system for nucleic acid sequencing and/or analysis.

[0006] FIG. 2 illustrates exemplary components of an apparatus for nucleic acid sequencing.

[0007] FIG. 3 illustrates an exemplary flow cell for nucleic acid sequencing.

[0008] FIG. 4 illustrates an exemplary computer system for performing nucleic acid sequencing.

[0009] FIG. 5 illustrates an exemplary process for label-free, pH-based sequencing.

[0010] FIG. 6 illustrates an exemplary method for performing nucleic acid sequencing using a homopolymer calibration panel.

[0011] FIG. 7 illustrates an exemplary method for preparing a homopolymer calibration panel.

[0012] FIG. 8 shows homopolymer representation statistics for an exemplary recalibration panel.

[0013] FIGS. 9A and 9B illustrate exemplary distributions of amplicon length for a two-exome sequencing run.

[0014] FIGS. 10A and 10B show exemplary distributions of GC content for a two-exome sequencing run.

[0015] FIGS. 11A and 11B show exemplary distributions of coverage, expressed as percentage of median coverage, for a two-exome sequencing run.

[0016] FIGS. 12A and 12B illustrate exemplary distributions of percentage forward end-to-end for a two-exome sequencing run.

[0017] FIGS. 13A and 13B illustrate exemplary distributions of percentage reverse end-to-end for a two-exome sequencing run.

[0018] FIGS. 14A and 14B illustrate exemplary distributions of strand bias.

[0019] FIG. 15 illustrates an exemplary plot showing the correlation between forward and reverse end-to-end coverage.

[0020] FIG. 16 illustrates an exemplary plot showing the estimated per-homopolymer coverage relative to the median panel coverage.

[0021] FIGS. 17A-17Z, 17AA-17AZ, and 17BA-17BY include an exemplary table, which lists a chromosome, insert start, insert end, and insert for each of the amplicons of the exemplary recalibration panel of FIG. 8, according to one embodiment.

### SUMMARY

[0022] According to an exemplary embodiment, there is provided a method for nucleic acid sequencing, comprising: (a) disposing a plurality of template polynucleotide strands in a plurality of defined spaces of a sensor array, the template polynucleotide strands comprising a set of homopolymer recalibration template polynucleotide strands; (b) exposing a plurality of the template polynucleotide strands, including the set of homopolymer recalibration template polynucleotide strands in the defined spaces, to a series of flows of nucleotide species flowed according to a predetermined ordering; and (c) determining sequence information for the plurality of the template polynucleotide strands, including the set of homopolymer recalibration template polynucleotide strands in the defined spaces, based on the flows of nucleotide species, to generate a plurality of sequencing reads corresponding to the template polynucleotide strands. The homopolymer recalibration template polynucleotide strands may comprise amplicon sequences that together comprise at least a minimal threshold number of homopolymers of each homopolymer length between a predetermined minimal homopolymer length and a predetermined maximal homopolymer length for one or more of homopolymer types A, T, C, and G. In one embodiment of the method, the predetermined minimal homopolymer length and the predetermined maximal homopolymer length is for each of homopolymer types A, T, C, and G.

[0023] In such a method, the minimal threshold number may be 10, for example. Alternately or in addition, the minimal threshold number may be 25, for example. Alternately or in addition, the minimal threshold number may be 50, for example. The homopolymer recalibration template polynucleotide strands may comprise amplicon sequences that are comprised in high-confidence regions of a reference genome (e.g., NIST NA12878) with no variants, for

example. The homopolymer recalibration template polynucleotide strands may comprise amplicon sequences that include, at most, one homopolymer of length 6, 7, 8, 9, or 10 per amplicon sequence, for example. The homopolymer recalibration template polynucleotide strands may comprise amplicon sequences having a minimal distance of 7 bases between any homopolymers of length 4, 5, 6, 7, 8, 9, or 10, for example. The homopolymer recalibration template polynucleotide strands may comprise amplicon sequences that do not overlap. The predetermined minimal homopolymer length may be 5, for example. The predetermined maximal homopolymer length may be 10, for example.

[0024] According to an exemplary embodiment, there is provided a system, including: a plurality of template polynucleotide strands disposed in a plurality of defined spaces of a sensor array, the template polynucleotide strands comprising a set of homopolymer recalibration template polynucleotide strands, wherein the homopolymer recalibration template polynucleotide strands comprise amplicon sequences that together comprise at least a minimal threshold number of homopolymers of each homopolymer length between a predetermined minimal homopolymer length and a predetermined maximal homopolymer length for one or more of homopolymer types A, T, C, and G; a machine-readable memory; and a processor configured to execute machine-readable instructions, which, when executed by the processor, cause the system to perform a method for nucleic acid sequencing, comprising: (a) exposing the plurality of the template polynucleotide strands, including the set of homopolymer recalibration template polynucleotide strands in the defined spaces, to a series of flows of nucleotide species flowed according to a predetermined ordering; and (b) determining sequence information for the plurality of the template polynucleotide strands, including the set of homopolymer recalibration template polynucleotide strands in the defined spaces, based on the flows of nucleotide species, e.g., to generate a plurality of sequencing reads corresponding to the template polynucleotide strands.

[0025] In such a system, the homopolymer recalibration template polynucleotide strands may include amplicon sequences that are comprised in high-confidence regions of a reference genome (e.g., NIST NA12878) with no variants. In one embodiment of the system, the predetermined minimal homopolymer length and the predetermined maximal homopolymer length is for each of homopolymer types A, T, C, and G.

[0026] According to an exemplary embodiment, there is provided a method for preparing a homopolymer recalibration panel, comprising: extracting, from a set of amplicons used in sequencing-by-synthesis, a set of candidate amplicons satisfying a first set of criteria, wherein the first set of criteria includes amplicons known to belong to high-confidence regions of a reference genome with no variants; and selecting, from the set of candidate amplicons, a reduced set of amplicons satisfying a second set of criteria, wherein the second set of criteria includes amplicons that together comprise at least a minimal threshold number of homopolymers of each homopolymer length between a predetermined minimal homopolymer length and a predetermined maximal homopolymer length for one or more of homopolymer types A, T, C, and G.

[0027] In such a method, the minimal threshold number may be 10, for example. Alternately or in addition, the minimal threshold number may be 25, for example. Alter-

nately or in addition, the minimal threshold number may be 50, for example. The reference genome with no variants may be NIST NA12878, for example. In one embodiment, the predetermined minimal homopolymer length and the predetermined maximal homopolymer length is for each of homopolymer types A, T, C, and G.

[0028] The reduced set of amplicons may comprise at most one homopolymer of length 6, 7, 8, 9, or 10 per amplicon, for example. The reduced set of amplicons may comprise amplicons having a minimal distance of 7 bases between any homopolymers of length 4, 5, 6, 7, 8, 9, or 10, for example. The reduced set of amplicons may comprise amplicons that do not overlap. The predetermined minimal homopolymer length may be 5, for example. The predetermined maximal homopolymer length may be 10, for example.

[0029] The method may further comprise determining underrepresented homopolymers of the set of candidate amplicons; and augmenting the set of candidate amplicons with a predetermined number of the underrepresented homopolymers. The method may further comprise disposing the reduced set of amplicons in a plurality of defined spaces of a sensor array. The method may also further comprise exposing the reduced set of amplicons to a series of flows of nucleotide species flowed according to a predetermined ordering; and determining sequence information for the reduced set of amplicons based on the flows of nucleotide species, to generate a plurality of sequencing reads corresponding to the reduced set of amplicons.

[0030] According to an exemplary embodiment, there is provided a homopolymer recalibration panel, comprising: a set of candidate amplicons extracted from a set of amplicons used in sequencing-by-synthesis, wherein the amplicons in the set of candidate amplicons: (a) are known to belong to high-confidence regions of a reference genome with no variants; and (b) together comprise at least a minimal threshold number of homopolymers of each homopolymer length between a predetermined minimal homopolymer length and a predetermined maximal homopolymer length for one or more of homopolymer types A, T, C, and G.

[0031] In various embodiments, a panel comprising amplicons with predetermined base sequences as described herein may be synthesized using any suitable nucleic acid synthesis methods known in the art.

[0032] In some embodiments, recalibration may include a single-pass calibration process in which a recalibration engine changes or modifies a set of default/initial parameters (e.g., homopolymers of various lengths being treated/weighed the same or according to some factory pre-determined set of initial homopolymer-specific parameters or weights). In other embodiments, recalibration may include a multi-pass or iterative process in which previously calibrated or recalibrated parameters may be further changed or modified by the calibration process.

[0033] The reference genome with no variants may be NIST NA12878, for example. The minimal threshold number may be 10, for example. Alternately or in addition, the minimal threshold number may be 25, for example. Alternately or in addition, the minimal threshold number may be 50, for example. In one embodiment of the method, the predetermined minimal homopolymer length and the predetermined maximal homopolymer length is for each of homopolymer types A, T, C, and G.

**[0034]** The set of candidate amplicons may include amplicon sequences that include, at most, one homopolymer of length 6, 7, 8, 9, or 10 per amplicon sequence, for example. The set of candidate amplicons may comprise amplicon sequences having a minimal distance of 7 bases between any homopolymers of length 4, 5, 6, 7, 8, 9, or 10, for example. The set of candidate amplicons may comprise amplicon sequences that do not overlap. The predetermined minimal homopolymer length may be 5, for example. The predetermined maximal homopolymer length may be 10, for example.

#### EXEMPLARY EMBODIMENTS

**[0035]** The following description and the various embodiments described herein are exemplary and explanatory only and are not to be construed as limiting or restrictive in any way. Other embodiments, features, objects, and advantages of the present teachings will be apparent from the description and accompanying drawings, and from the claims.

**[0036]** According to various exemplary embodiments, panels of amplicons for homopolymer calibration or recalibration for use with nucleic acid sequencing data and methods for designing the same, are disclosed herein. Such panels of amplicons may improve downstream processing (including variant calling), since such panels may improve calibration and recalibration of nucleic acid sequencing data and/or reduce certain systematic errors and improve overall sequencing accuracy (especially in the case of long homopolymers).

#### Design of Calibration Panel

**[0037]** In various embodiments, a homopolymer calibration panel may be designed to have a substantially uniform representation of homopolymers of various lengths. Homopolymers of relatively short lengths (e.g., 2, 3, and 4 bases) may be well represented in a sufficiently large set of sufficiently long sequences selected using any suitable arbitrary or random approach. However, homopolymers of relatively long lengths (e.g., 5, 6, 7, 8, 9, and 10, or more) are naturally rarer than homopolymers of relatively short lengths. Thus, homopolymers of relatively long lengths may be insufficiently represented (or at least under-represented compared with shorter homopolymers) among sequences selected using any suitable arbitrary or random approach.

**[0038]** In an embodiment, a set of sequences containing a desired uniform representation across homopolymers may be defined by computationally and combinatorially populating a set of sequences with desired types and quantities of homopolymers. For example, a set of sequences may be populated to include exactly or at least  $n(\text{MinL}, \text{MaxL}, \text{NumT})$  homopolymers, each homopolymer having a length between lengths  $\text{MinL}$  and  $\text{MaxL}$  (e.g., each length between  $\text{MinL}=1$  and  $\text{MaxL}=10$ , or each length between  $\text{MinL}=5$  and  $\text{MaxL}=10$ , etc.). The set of sequences may further be populated to include at least one type of nucleotide among  $\text{NumT}$  types of nucleotides (e.g., each type among  $\text{NumT}=4$  types A, C, G, and T), where  $n$  is an integer (e.g., 10, 25, 50, 100, etc.) that may be a function of parameters  $\text{MinL}$ ,  $\text{MaxL}$ , and  $\text{NumT}$ .

**[0039]** In some cases, one or more sets of sequences containing such homopolymers populated computationally and combinatorially may not have been empirically tested and may not be ideally suited for a given underlying

sequencing technology. One method to address this may include designing a set of sequences containing a substantially uniform representation across homopolymers using steps including: (1) identifying an initial set of candidate amplicons or oligonucleotides known to function properly on a particular sequencing platform or technology (e.g., a set of amplicons from the Ion AmpliSeq™ Exome Panel or any suitable panel used with some given underlying sequencing technology); (2) selecting, from the initial set of candidate amplicons or oligonucleotides, a subset of amplicons or oligonucleotides meeting one or more selection criteria (e.g., one or more minimal numbers of occurrences of homopolymers of certain lengths and types, for example, those having at least 50 (or some suitable integer, such as 10, 25, 50, 75, etc., for example) homopolymers of each length 5, 6, 7, 8, 9, and 10, and/or of each type A, C, G, and T); and (3) augment the subset of amplicons or oligonucleotides with additional amplicons or oligonucleotides comprising a desired number of under-represented homopolymers (e.g., by adding a substantial number of homopolymers of length 9 and/or 10 or other comparatively rare and under-represented length(s)). Augmenting the subset of amplicons or oligonucleotides with additional amplicons or oligonucleotides may help achieve a desired level of representation uniformity across homopolymers.

**[0040]** In an embodiment, a set of sequences containing a substantially uniform representation across homopolymers may be designed using steps including: (1) identifying an initial set of candidate amplicons or oligonucleotides known to have been used with a particular sequencing platform or technology; (2) selecting from the initial set of candidate amplicons or oligonucleotides a subset of amplicons or oligonucleotides that together include all n-mers up to a predetermined maximal homopolymer length (e.g., up to  $n=5, 6, 7, 8, 9$ , or 10) for bases A, C, G, and T with a predetermined minimum of n-mers of each length and/or type of nucleotide (e.g., at least 10, 25, 50, 75, or more, of each length and/or type); and (3) performing an empirically-based pruning or refinement selection to reduce the impact on throughput. Such pruning or refinement selection may include minimizing the number of amplicons or oligonucleotides by selecting amplicons that have several n-mers but maintaining the quality of the selected amplicons (e.g., in an example further discussed below, a final panel may have 384 amplicons, a 23% reduction from a starting point of 500 candidate amplicons, where 500 may be the product of 50 n-mers times 4 bases times 5 (for  $n=6, 7, 8, 9$ , and 10, with shorter n-mers being automatically included), divided by 2 (for strands that produce complementary homopolymers)).

**[0041]** In various embodiments, a homopolymer calibration panel may be designed using steps including: (1) identifying an initial set of candidate amplicons or oligonucleotides that are known to have been used with a particular sequencing platform or technology (e.g., a substantial number, such as 300, 400, 500, or more, amplicons or oligonucleotides from the Ion AmpliSeq™ Exome Panel or any suitable panel used with some given underlying sequencing technology) and that are inside high-confidence regions of a reference genome (e.g., NIST NA12878) with no variants so that the real homopolymer length(s) may be known; (2) selecting, from the initial set of candidate amplicons or oligonucleotides, a subset of amplicons or oligonucleotides that together include a predetermined minimum of n-mers of each homopolymer length from a prede-

terminated minimal length to a predetermined maximal length for all bases A, C, G, and T; (3) filtering out amplicons or oligonucleotides that violate one or more of the following constraints: (a) having more than one same-base n-mer with n=6 or more, (b) having more than a minimal separation of 7 bases between n-mers with n=4 or more bases to obviate or reduce additional errors and de-phasing that may be introduced with neighboring homopolymers (except that several G or C n-mers may be on the same strand if they are separated by at least 7 bases and at least 3 bases from A/T n-mers, given that long C and G homopolymers may be rare and GC-rich regions may be particularly difficult to sequence), (c) having an overlap with another amplicon or oligonucleotide in the set, and (d) having a homopolymer of length longer than 10; and (4) if desired to achieve a determined or selected level of representation uniformity across homopolymers, augment the subset of amplicons or oligonucleotides with additional amplicons or oligonucleotides comprising a desired number of under-represented homopolymers (e.g., by adding a substantial number of G and C homopolymers of length 9 and/or 10 to represent problematic homopolymers in the set).

[0042] FIG. 1 illustrates an exemplary system for nucleic acid sequencing and/or analysis. The system includes an apparatus or sub-system for nucleic acid sequencing and/or analysis 11, a computing server/node/device 12 including a base calling engine 13, a recalibration engine 14, a post-processing engine 15, and a display 16, which may be internal and/or external. The apparatus or sub-system for nucleic acid sequencing and/or analysis 11 may be any type of instrument that can generate nucleic acid sequence data from nucleic acid samples, for example, a nucleic acid sequencing instrument, a real-time/digital/quantitative PCR instrument, a microarray scanner, etc. The nucleic acid samples may include amplicons for calibration or recalibration as further described herein. The computing server/node/device 12 may include a workstation, mainframe computer, distributed computing node (part of a “cloud computing” or distributed networking system), personal computer, mobile device, etc. The base calling engine 13 may include any suitable base caller and may be configured to include various signal/data processing modules that may be configured to receive signal/data from the apparatus or sub-system for nucleic acid sequencing and/or analysis 11 and perform various processing steps, for example, conversion from flow space to base space, determination of base calls for some or the entirety of a sequencing data set, and/or determination of base call quality values. In an embodiment, the base calling engine 13 may implement one or more features described in Davey et al., U.S. Pat. Appl. Publ. No. 2012/0109598, published on May 3, 2012, and/or Sikora et al., U.S. Pat. Appl. Publ. No. 2013/0060482, published on Mar. 7, 2013, each of which is incorporated by reference herein in its entirety. The base calling engine 13 may also include a mapping or alignment module for mapping or aligning reads to a reference sequence or genome, which may include a whole/partial genome, a whole/partial exome, etc. In an embodiment, the mapping or alignment module may include any suitable aligner, including the Torrent Mapping Alignment Program (TMAP), for example. The recalibration engine 14 may be configured to recalibrate base calls or related intensity values or parameters based on an analysis of base calling and alignment performed by the base calling engine 13. Recalibrated base calls or related intensity values,

thresholds, or parameters may be fed back into the base calling engine 13 for improving the accuracy of base calls. In an embodiment, the recalibration engine 14 may implement one or more features described in Jiang et al., U.S. Pat. Appl. Publ. No. 2014/0316716, published on Oct. 23, 2014, which is incorporated by reference herein in its entirety. The exemplary system may also include a client device terminal 17, which may include a data analysis API or module and may be communicatively connected to the computing server/node/device 12 via a network connection 18 that may be a “hardwired” physical network connection (e.g., Internet, LAN, WAN, VPN, etc.) or a wireless network connection (e.g., Wi-Fi, WLAN, etc.). The post-processing engine 15 may be configured to include various signal/data processing modules that may be configured to make variant calls and apply post-processing to variant calls, which may include annotating various variant calls and/or features, converting data from flow space to base space, filtering of variants, and/or formatting the variant data for display or use by client device terminal 17. Variant calls may be made using any suitable variant caller, including the Germ-Line Variant Caller and the Torrent Variant Caller (TVC) Plug-ins for Ion Torrent™ sequencing technology. In an embodiment, the variant caller may implement one or more features described in Hubbell et al., U.S. Pat. Appl. No. 2014/0296080, published on Oct. 2, 2014, which is incorporated by reference herein in its entirety. In an embodiment, the apparatus or sub-system for nucleic acid sequencing and/or analysis 11 and the computing server/node/device 12 may be integrated into a single instrument or system comprising components present in a single enclosure 19. The client device terminal 17 may be configured to communicate information to and/or control the operation of the computing server/node/device 12 and its modules and/or operating parameters.

[0043] FIG. 2 illustrates exemplary components of an apparatus for nucleic acid sequencing. Such an apparatus could be used as apparatus or sub-system for nucleic acid sequencing and/or analysis 11 of FIG. 1. The components include a flow cell and sensor array 100, a reference electrode 108, a plurality of reagents 114, a valve block 116, a wash solution 110, a valve 112, a fluidics controller 118, lines 120/122/126, passages 104/109/111, a waste container 106, an array controller 124, and a user interface 128. The flow cell and sensor array 100 includes an inlet 102, an outlet 103, a microwell array 107, and a flow chamber 105 defining a flow path of reagents over the microwell array 107. The reference electrode 108 may be of any suitable type or shape, including a concentric cylinder with a fluid passage or a wire inserted into a lumen of passage 111. The reagents 114 may be driven through the fluid pathways, valves, and flow cell by pumps, gas pressure, or other suitable methods, and may be discarded into the waste container 106 after exiting the flow cell and sensor array 100. The reagents 114 may, for example, contain dNTPs to be flowed through passages 130 and through the valve block 116, which may control the flow of the reagents 114 to flow chamber 105 (also referred to herein as a reaction chamber) via passage 109. The system may include a reservoir for containing a wash solution 110 that may be used to wash away dNTPs, for example, that may have previously been flowed. The microwell array 107 may include an array of defined spaces (e.g., microwells), for example, wherein the array may be operationally associated with a sensor array so that each

microwell may be associated with a sensor suitable for detecting an analyte or reaction property of interest. The defined spaces may include nucleic acid samples, which may include amplicons for calibration or recalibration as further described herein. The microwell array 107 may be integrated with the sensor array as a single device or chip. The array controller 124 may provide bias voltages, timing, and/or control signals to the sensor, and collect and/or process output signals. The user interface 128 may display information from the flow cell and sensor array 100 as well as instrument settings and controls, and allow a user to enter or set instrument settings and controls. The valve 112 may be shut to prevent any wash solution 110 from flowing into passage 109 as the reagents are flowing. Although the flow of wash solution may be stopped, there may still be uninterrupted fluid and electrical communication between the reference electrode 108, passage 109, and the microwell array 107. The fluidics controller 118 may be programmed to control driving forces for flowing reagents 114 and to control the operation of valve 112 and valve block 116 to deliver reagents to the flow cell and sensor array 100 according to a predetermined reagent flow ordering.

[0044] In this application, “defined space” may refer to any space (which may be in one, two, or three dimensions) in which at least some of a molecule, fluid, and/or solid can be confined, retained, and/or localized. A space may be a predetermined area (which may be a flat area) or volume, and may be defined, for example, by a depression or a micro-machined well in or associated with a microwell plate, microtiter plate, microplate, or a chip, or by isolated hydrophobic areas on a generally hydrophobic surface. Defined spaces may be arranged as an array, which may be a substantially planar one-dimensional or two-dimensional arrangement of elements, including sensors or wells. Defined spaces, whether arranged as an array or in some other configuration, may be in electrical communication with at least one sensor to allow detection or measurement of one or more detectable or measurable parameters or characteristics. The sensors may convert changes in the presence, concentration, or amounts of reaction by-products (or changes in ionic character of reactants) into an output signal, which may be registered electronically, for example, as a change in a voltage level or a current level. In one embodiment, the output signal and/or change in voltage or current level, in turn, may be processed to extract information or signal about a chemical reaction or desired association event, for example, a nucleotide incorporation event and/or a related ion concentration (e.g., a pH measurement). The sensors may include at least one ion sensitive field effect transistor (“ISFET”) and/or chemically sensitive field effect transistor (“chemFET”).

[0045] FIG. 3 illustrates an exemplary flow cell for nucleic acid sequencing. The flow cell 200 includes a microwell array 202, a sensor array 205, and a flow chamber 206 in which a reagent flow 208 may move across a surface of the microwell array 202 and/or over open ends of microwells in the microwell array 202. The flow of reagents (e.g., nucleotide species) can be provided in any suitable manner, including delivery by pipettes, or through tubes or passages connected to a flow chamber. A microwell 201 in the microwell array 202 may have any suitable volume, shape, and aspect ratio. A sensor 214 in the sensor array 205 may be an ISFET or a chemFET sensor with a floating gate 218 having a sensor plate 220 separated from the microwell

interior by a passivation layer 216, and may be predominantly responsive to (and generate an output signal related to) an amount of charge 224 present on the passivation layer 216 opposite of the sensor plate 220. Changes in the amount of charge 224 may cause changes in a current between a source 221 and a drain 222 of the sensor 214, which may be used directly to provide a current-based output signal, or indirectly with additional circuitry to provide a voltage output signal. Reactants, wash solutions, and other reagents may move into microwells primarily by diffusion 240. One or more analytical reactions to identify or determine characteristics or properties of an analyte of interest may be carried out in one or more microwells of the microwell array 202. Such reactions may directly or indirectly generate by-products that affect the amount of charge 224 adjacent to the sensor plate 220. In an embodiment, a reference electrode 204 may be fluidly connected to the flow chamber 206 via a flow passage 203. In an embodiment, the microwell array 202 and the sensor array 205 may together form an integrated unit forming a bottom wall or floor of the flow cell 200. In an embodiment, one or more copies of an analyte may be attached to a solid phase support 212, which may include microparticles, nanoparticles, beads, or gels, and may be solid and porous, for example. The analyte may include one or more copies of a nucleic acid analyte, which may include nucleic acid samples. Exemplary samples may include amplicons for calibration or recalibration, which may be obtained using any suitable technique, as further described herein.

[0046] FIG. 4 illustrates an exemplary computer system 401 for performing nucleic acid sequencing. Such a computer system 401 could be used as computing server/node/device 12 of FIG. 1. The computer system 401 may include a bus 402 or other communication mechanism for communicating information, a processor 403 coupled to the bus 402 for processing information, and/or a memory 405 coupled to the bus 402 for dynamically and/or statically storing information. The computer system 401 can also include one or more co-processors 404 coupled to the bus 402 (e.g., GPUs and/or FPGAs) for performing specialized processing tasks; a display 406 coupled to the bus 402 (e.g., a cathode ray tube (CRT) or liquid crystal display (LCD)) for displaying information to a computer user; an input device 407 coupled to the bus 402 (e.g., a keyboard including alphanumeric and other keys) for communicating information and command selections to the processor 403; a cursor control device 408 coupled to the bus 402 (e.g., a mouse, a trackball, or cursor direction keys) for communicating direction information and command selections to the processor 403 and for controlling cursor movement on display 406; and one or more storage devices 409 coupled to the bus 402 (e.g., a magnetic disk or an optical disk) for storing information and instructions. The memory 405 may include a random access memory (RAM) or other dynamic storage device and/or a read only memory (ROM) or other static storage device. Such an exemplary computer system with suitable software may be used to perform the embodiments described herein. More generally, in various embodiments, one or more features of the teachings and/or embodiments described herein may be performed or implemented using appropriately configured and/or programmed hardware and/or software elements.

[0047] Examples of hardware elements may include processors, microprocessors, input(s) and/or output(s) (I/O)

device(s) (or peripherals) that are communicatively coupled via a local interface circuit, circuit elements (e.g., transistors, resistors, capacitors, inductors, and so forth), integrated circuits, application specific integrated circuits (ASIC), programmable logic devices (PLD), digital signal processors (DSP), field programmable gate array (FPGA), logic gates, registers, semiconductor devices, chips, microchips, chip sets, and so forth. The local interface may include, for example, one or more buses or other wired or wireless connections, controllers, buffers (caches), drivers, repeaters and receivers, etc., to allow appropriate communications between hardware components. A processor may include a hardware device for executing software, particularly software stored in memory. The processor may include any custom made or commercially available processor, a central processing unit (CPU), an auxiliary processor among several processors associated with the computer, a semiconductor based microprocessor (e.g., in the form of a microchip or chip set), a macroprocessor, or generally any device for executing software instructions. A processor can also represent a distributed processing architecture. The I/O devices can include input devices, for example, a keyboard, a mouse, a scanner, a microphone, a touch screen, an interface for various medical devices and/or laboratory instruments, a bar code reader, a stylus, a laser reader, a radio-frequency device reader, etc. Furthermore, the I/O devices also can include output devices, for example, a printer, a bar code printer, a display, etc. Finally, the I/O devices further can include devices that communicate as both inputs and outputs, for example, a modulator/demodulator (modem; for accessing another device, system, or network), a radio frequency (RF) or other transceiver, a telephonic interface, a bridge, a router, etc.

**[0048]** Examples of software may include software components, programs, applications, computer programs, application programs, system programs, machine programs, operating system software, middleware, firmware, software modules, routines, subroutines, functions, methods, procedures, software interfaces, application program interfaces (API), instruction sets, computing code, computer code, code segments, computer code segments, words, values, symbols, or any combination thereof. A software in memory may include one or more separate programs, which may include ordered listings of executable instructions for implementing logical functions. The software in memory may include a system for identifying data streams in accordance with the present teachings and any suitable custom made or commercially available operating system (O/S), which may control the execution of other computer programs such as the system, and provide scheduling, input-output control, file and data management, memory management, communication control, etc.

**[0049]** According to various embodiments, one or more features of teachings and/or embodiments described herein may be performed or implemented using an appropriately configured and/or programmed non-transitory machine-readable medium or article that may store an instruction or a set of instructions that, if executed by a machine, may cause the machine to perform a method and/or operations in accordance with the embodiments. Such a machine may include, for example, any suitable processing platform, computing platform, computing device, processing device, computing system, processing system, computer, processor, scientific or laboratory instrument, etc., and may be imple-

mented using any suitable combination of hardware and/or software. The machine-readable medium or article may include, for example, any suitable type of memory unit, memory device, memory article, memory medium, storage device, storage article, storage medium and/or storage unit, for example, memory, removable or non-removable media, erasable or non-erasable media, writeable or re-writeable media, digital or analog media, hard disk, floppy disk, read-only memory compact disc (CD-ROM), recordable compact disc (CD-R), rewriteable compact disc (CD-RW), optical disk, magnetic media, magneto-optical media, removable memory cards or disks, various types of Digital Versatile Disc (DVD), a tape, a cassette, etc., including any medium suitable for use in a computer. Memory can include any one or a combination of volatile memory elements (e.g., random access memory (RAM, such as DRAM, SRAM, SDRAM, etc.)) and nonvolatile memory elements (e.g., ROM, EPROM, EEPROM, flash memory, hard drive, tape, CDROM, etc.). Moreover, memory can incorporate electronic, magnetic, optical, and/or other types of storage media. Memory can have a distributed, clustered, remote, or cloud architecture where various components may be situated remote from one another, and accessed by the processor. The instructions may include any suitable type of code, including source code, compiled code, interpreted code, executable code, static code, dynamic code, encrypted code, etc., implemented using any suitable high-level, low-level, object-oriented, visual, compiled and/or interpreted programming language.

**[0050]** FIG. 5 illustrates an exemplary process 500 for label-free, pH-based sequencing. In one embodiment, a template 582 with sequence 585 and a primer binding site 581 may be attached to a solid phase support 580. The template 582 may include nucleic acid samples comprising amplicons for calibration or recalibration as further described herein. The template 582 may be attached as a clonal population to a solid support (e.g., a microparticle or bead), for example, and may be prepared as disclosed in Leamon et al., U.S. Pat. No. 7,323,305. In an embodiment, the template may be associated with a substrate surface or present in a liquid phase with or without being coupled to a support. A primer 584 and DNA polymerase 586 may be annealed to the template 582 so that the primer's 3' end may be extended by a polymerase, and a polymerase may be bound to such a primer-template duplex (or in close proximity thereof) so that binding and/or extension may take place when dNTPs are added. In one embodiment, step 588 may include adding dNTP (shown as dATP), and the DNA polymerase 586 may incorporate a nucleotide "A" (since "T" may be the next nucleotide in the template 582, and "T" is complementary to the flowed dATP nucleotide). In one embodiment, step 590 may include performing a wash. In one embodiment, step 592 may include adding the next dNTP (shown as dCTP), and the DNA polymerase 586 may incorporate a nucleotide "C" (since "G" may be the next nucleotide in the template 582). More details about pH-based nucleic acid sequencing may be found in U.S. Pat. No. 7,948,015 and U.S. Pat. Appl. Publ. Nos. 2010/0137143, 2009/0026082, and 2010/0282617, each of which is incorporated by reference herein in its entirety.

**[0051]** In an embodiment, the primer-template-polymerase complex may be subjected to a series of exposures of different nucleotides in a pre-determined sequence or ordering. If one or more nucleotides are incorporated, the

signal resulting from the incorporation reaction may be detected. In one embodiment, the nucleotide sequence of the template strand may be determined after repeated cycles of nucleotide addition, primer extension, and/or signal acquisition. The output signals measured throughout this process may depend on the number of nucleotide incorporations. Specifically, in each addition step, the polymerase may extend the primer by incorporating added dNTP only if the next base in the template is complementary to the added dNTP. With each incorporation, a hydrogen ion may be released, and collectively, a population of released hydrogen ions may change a local pH of the respective reaction chamber. The production of hydrogen ions may be monotonically related to the number of contiguous complementary bases (e.g., homopolymers) in the template. Deliveries of nucleotides to a reaction vessel or chamber may be referred to as “flows” of nucleotide triphosphates (or dNTPs). For convenience, a flow of dATP will sometimes be referred to as “a flow of A” or “an A flow,” and a sequence of flows may be represented as a sequence of letters, such as “ATGT” indicating “a flow of dATP, followed by a flow of dTTP, followed by a flow of dGTP, followed by a flow of dTTP.” The predetermined ordering may be based on a cyclical, repeating pattern including consecutive repeats of a short pre-determined reagent flow ordering (e.g., consecutive repeats of pre-determined sequence of four nucleotide reagents, for example, “ACTG ACTG . . . ”). The predetermined ordering may be based in whole or in part on some other pattern of reagent flows (e.g., any of the various reagent flow orderings discussed in Hubbell et al., U.S. Pat. Appl. Publ. No. 2012/0264621, published Oct. 18, 2012, which is incorporated by reference herein in its entirety), and may also be based on some combination thereof.

**[0052]** In various embodiments, output signals due to nucleotide incorporation may be processed, given knowledge of what nucleotide species were flowed and in what order to obtain such signals. The output signals may be processed to make base calls for the flows and/or to compile consecutive base calls associated with a sample nucleic acid template into a read. A base call may refer to a particular nucleotide identification (e.g., dATP (“A”), dCTP (“C”), dGTP (“G”), or dTTP (“T”)). Base calling may include performing one or more signal normalizations, signal phase and signal decay (e.g., enzyme efficiency loss) estimations, signal corrections, and model-based signal predictions. Base calling may also identify or estimate base calls for each flow for each defined space. Any suitable base calling method may be used, including as described in Davey et al., U.S. Pat. Appl. Publ. No. 2012/0109598, published on May 3, 2012, and/or Sikora et al., U.S. Pat. Appl. Publ. No. 2013/0060482, published on Mar. 7, 2013, each of which is incorporated by reference herein in its entirety, taking into account that more accurate base callers may yield better results.

**[0053]** FIG. 6 illustrates an exemplary method 600 for nucleic acid sequencing using a homopolymer calibration panel. In step 601, a plurality of template polynucleotide strands may be disposed in a plurality of defined spaces disposed on a sensor array, the template polynucleotide strands comprising a set of homopolymer recalibration template polynucleotide strands, wherein the homopolymer recalibration template polynucleotide strands may comprise amplicon sequences that together comprise at least a minimal threshold number of homopolymers of each homopoly-

lymer length between a predetermined minimal homopolymer length and a predetermined maximal homopolymer length for one or more of homopolymer types A, T, C, and G (e.g., each of homopolymer types A, T, C, and G). In step 602, the plurality of the template polynucleotide strands, including the set of homopolymer recalibration template polynucleotide strands in the defined spaces, may be exposed to a series of flows of nucleotide species flowed according to a predetermined ordering. In step 603, sequence information for a plurality of the template polynucleotide strands, including the set of homopolymer recalibration template polynucleotide strands in the defined spaces, may be determined based on the flows of nucleotide species, e.g., to generate a plurality of sequencing reads corresponding to the template polynucleotide strands.

**[0054]** FIG. 7 illustrates an exemplary method 700 for preparing a homopolymer recalibration panel. In one embodiment, the recalibration panel may include amplicons that are 200 base pairs in length. In another embodiment, the recalibration panel may include amplicons longer than 200 base pairs in length (e.g., amplicons that are 600 base pairs in length). In step 701, a set of candidate amplicons satisfying a first set of criteria may be extracted from a set of amplicons used in sequencing-by-synthesis, wherein the first set of criteria may include selecting amplicons known to belong to high-confidence regions of a reference genome with no variants. In step 702, a reduced set of amplicons satisfying a second set of criteria may be selected from the set of candidate amplicons, wherein the second set of criteria may include selecting amplicons that together comprise at least a minimal threshold number of homopolymers of each homopolymer length between a predetermined minimal homopolymer length and a predetermined maximal homopolymer length for one or more of homopolymer types A, T, C, and G (e.g., each of homopolymer types A, T, C, and G).

**[0055]** FIG. 8 shows homopolymer representation statistics for an exemplary recalibration panel 800. The upper table includes exemplary homopolymer compositions of amplicons in the exemplary recalibration panel 800. The amplicons shown in the upper table may represent a subset of amplicons on the Ion AmpliSeq™ Exome Panel. The “base” group of amplicons may represent selected amplicons that are inside NIST NA12878 high-confidence regions with no variants and reasonably high sequencing quality that further satisfy one or more of the following exemplary constraints: (1) only one n-mer (where n is 6 or more) may be allowed on a given template; (2) a minimal separation of 7 bases between n-mers, where n is 4 or more; and (3) no overlapping between the amplicons and/or non-mers longer than 10-mers, with a few exceptions (e.g., several G or C n-mers can be on the same strand if they are separated by at least 7 bases and at least 3 bases from A/T n-mers). An exemplary reference genome in addition to NIST NA12878 may include NIST RM 8398. The “backup” group of amplicons may represent additional amplicons with higher sequencing quality (including mostly G/C n-mers). The “extra 9GC\_10GC” group of amplicons may represent additional amplicons with 9G/9C/10G/10C n-mers. The lower table includes an estimation of sequencing quality, e.g., based on a good two-exome run on an Ion Proton™ system implementing Ion Torrent™ sequencing technology. FIGS. 17A-17Z, 17AA-17AZ, and 17BA-17BY include table 1700, which lists the chromosome, insert start, insert end,

and insert for each of the **384** amplicons discussed in FIG. **8**. Among these, inserts **1** through **288** may correspond to the “base” group; inserts **289** through **347** may correspond to the “backup” group, and inserts **348** through **384** may correspond to the “extra\_9GC\_10GC” group.

**[0056]** FIGS. **9A** and **9B** illustrate exemplary distributions of amplicon length for a two-exome sequencing run. FIG. **9A** includes an exemplary distribution **900** of amplicon length for a full AmpliSeq™ Exome Panel run (e.g.,  $\frac{1}{2}$  of an Ion Proton™ system run implementing Ion Torrent™ sequencing technology). FIG. **9B** includes an exemplary distribution **920** of amplicon length for the subset of **384** amplicons discussed in FIG. **8** and listed in FIGS. **17A-17Z**, **17AA-17AZ**, and **17BA-17BY**.

**[0057]** FIGS. **10A** and **10B** show exemplary distributions of GC content for a two-exome sequencing run. FIG. **10A** includes an exemplary distribution **1000** of GC content for the full AmpliSeq™ Exome Panel run (e.g., half of an Ion Proton™ system run implementing Ion Torrent™ sequencing technology). FIG. **10B** shows a distribution **1020** of GC content for the subset of **384** amplicons discussed in FIG. **8** and listed in FIGS. **17A-17Z**, **17AA-17AZ**, and **17BA-17BY**.

**[0058]** FIGS. **11A** and **11B** show exemplary distributions of coverage, expressed as percentage of median coverage, for a two-exome sequencing run. FIG. **11A** includes an exemplary distribution **1100** of coverage, expressed as percentage of median coverage, for the full AmpliSeq™ Exome Panel run (e.g.,  $\frac{1}{2}$  of an Ion Proton™ system run implementing Ion Torrent™ sequencing technology). FIG. **11B** includes an exemplary distribution **1120** of coverage, expressed as percentage of median coverage, for the subset of **384** amplicons discussed in FIG. **8** and listed in FIGS. **17A-17Z**, **17AA-17AZ**, and **17BA-17BY**. The leftmost peak in FIG. **11B** may represent additional 10G/10C n-mers.

**[0059]** FIGS. **12A** and **12B** illustrate exemplary distributions of percentage forward end-to-end for a two-exome sequencing run. FIG. **12A** includes an exemplary distribution **1200** of percentage forward end-to-end for the full AmpliSeq™ Exome Panel run (e.g., half of an Ion Proton™ system run implementing Ion Torrent™ sequencing technology). FIG. **12B** includes an exemplary distribution **1220** of percentage forward end-to-end for the subset of **384** amplicons discussed in FIG. **8** and listed in FIGS. **17A-17Z**, **17AA-17AZ**, and **17BA-17BY**. Similarity in the distributions in FIGS. **12A** and **12B** may indicate that little or no bias is introduced in the recalibration subset by selecting amplicons that over- or under-perform relative to the full set of AmpliSeq™ Exome amplicons.

**[0060]** FIGS. **13A** and **13B** illustrate exemplary distributions of percentage reverse end-to-end for a two-exome sequencing run. FIG. **13A** includes an exemplary distribution **1300** of percentage reverse end-to-end for the full AmpliSeq™ Exome Panel run (e.g.,  $\frac{1}{2}$  of an Ion Proton™ system run implementing Ion Torrent™ sequencing technology). FIG. **13B** includes an exemplary distribution **1320** of percentage reverse end-to-end for the subset of **384** amplicons discussed in FIG. **8** and listed in FIGS. **17A-17Z**, **17AA-17AZ**, and **17BA-17BY**. As with FIGS. **12A** and **12B**, similarities between the distributions in FIGS. **13A** and **13B** may indicate that little or no bias is introduced in the recalibration subset by selecting amplicons that over- or under-perform relative to full set of AmpliSeq™ Exome amplicons.

**[0061]** FIGS. **14A** and **14B** illustrate exemplary distributions of strand bias. FIG. **14A** includes an exemplary distribution **1400** of percentage forward of all assigned reads (strand bias) for the full exome run (e.g.,  $\frac{1}{2}$  of an Ion Proton™ system run implementing Ion Torrent™ sequencing technology). FIG. **14B** includes an exemplary distribution **1420** of percentage forward of all assigned reads (strand bias) for a selected subset of recalibration amplicons. Similarity in the distributions in FIGS. **14A** and **14B** may indicate that little or no bias is introduced in the recalibration subset and that forward and reverse strands may be sequenced equally well (with the exception of a small number of amplicons with long G/C homopolymers).

**[0062]** FIG. **15** illustrates an exemplary plot **1500** showing the correlation between forward and reverse end-to-end coverage. The data may pertain to a selected subset of recalibration amplicons. The majority of amplicons may have good end-to-end coverage on both strands. Still, the inclusion of at least some recalibration amplicons of a poorer quality may be desirable to maintain a fair representation (e.g., of typical AmpliSeq™ panels) which may include difficult to sequence amplicons (e.g., amplicons with long G/C homopolymers). The area in the bottom left corner of plot **1500** may include 24 poor end-to-end amplicons having a sum of percentage forward end-to-end and percentage reverse end-to-end that is less than 20%.

**[0063]** FIG. **16** illustrates an exemplary plot **1600** showing the estimated per-homopolymer coverage relative to the median panel coverage. Here, data may pertain to homopolymers of length 5 through 10 for nucleotide types A, C, G, and T.

**[0064]** In an embodiment, series of measured intensities obtained for panels of amplicons for homopolymer calibration or recalibration as described herein may be used as training subset(s) within the recalibration engine described in Jiang et al., U.S. Pat. Appl. Publ. No. 2014/0316716, published on Oct. 23, 2014, which is incorporated by reference herein in its entirety, instead of (or in addition to) the series of measured intensities obtained for a randomly selected training subset as described in Jiang et al., U.S. Pat. Appl. Publ. No. 2014/0316716, published on Oct. 23, 2014.

**[0065]** Unless otherwise specifically designated herein, terms, techniques, and symbols of biochemistry, cell biology, genetics, molecular biology, nucleic acid chemistry, nucleic acid sequencing, and organic chemistry used herein follow those of standard treatises and texts in the relevant field.

**[0066]** Although the present description described in detail certain embodiments, other embodiments are also possible and within the scope of the present invention. For instance, while described embodiments may include recalibration panels with amplicons that are 200 base pairs in length, the embodiments may also be tailored to recalibration panels with amplicons that are longer in length (e.g., 600 base pairs in length). For example, those skilled in the art may appreciate from the present description that the present teachings may be implemented in a variety of forms, and that the various embodiments may be implemented alone or in combination. Variations and modifications will be apparent to those skilled in the art from consideration of the specification and figures and practice of the teachings described in the specification and figures, and the claims.

---

SEQUENCE LISTING

Sequence total quantity: 384

SEQ ID NO: 1 moltype = DNA length = 224  
FEATURE Location/Qualifiers  
source 1..224  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 1  
ccgtctggaa ggtgctgcac acagtggc ttcaaggacgc tgggggggg ctgggtgaga 60  
ccctgcctgg cttccaggc ccaagggagt ggggggggggg cggggccggg cgccccactc 120  
ttgagggtgc tgtagacgcg caatccgc tcctgctct tgcgtgtgtc atgcagggtc 180  
acgcggcagg tgaagcgcag ctggtgcata gccagaccac gtc 224

SEQ ID NO: 2 moltype = DNA length = 218  
FEATURE Location/Qualifiers  
source 1..218  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 2  
ctctgggggg gggcctgggt gcttatgccccc tcccggtctg gtgcaggggg actccgggtgg 60  
cccaactgaaac tgcccaactgg agaacgggtt ccgggaggtt tttggcatcg tcagctttgg 120  
ctcccgccgg ggctgcaaca cccgcaagaa gcccgtatgc tacacccggg tgcgtgcata 180  
catcgactgg atcaacgagg tgggtgcata ctccacag 218

SEQ ID NO: 3 moltype = DNA length = 191  
FEATURE Location/Qualifiers  
source 1..191  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 3  
ttgacagcca aatgcaagac gagttctca ggtgcctgaa aacacatcca cacctctgag 60  
attccccccc cagtgcacct gcctccaacc cagttttcc ctcttgaact cttgctaacc 120  
ccttatecccc ccataggaag ctggcagaac cctggagaga tgagactgca agggcagagc 180  
ctcatgtctg g 191

SEQ ID NO: 4 moltype = DNA length = 228  
FEATURE Location/Qualifiers  
source 1..228  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 4  
accctttgag ccaactagtg tggacatgtat ggggggggat gcccggcgcc tgcttgcctt 60  
ggcccaagggtg gacgagaacc actcagagtt tactctgtac gaatcacggc tggtggacat 120  
ctcggtatca ccgttcatgtatca actcattggat ttcacaatgcg atttgtatgt tactgttttt 180  
gattggccag tcatgacagc cgtgggacac ctccccccccc cgtgtgtg 228

SEQ ID NO: 5 moltype = DNA length = 218  
FEATURE Location/Qualifiers  
source 1..218  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 5  
tgctaacaat ccattctgaa agcaagggtg gcttcctcgta taagtataaa ctataacatt 60  
ccaaatcagg cccaaatgtac cttttctggat gattcacaca atatatttc tggtttcaga 120  
tcacgatgatca caatgcctgaa catgacaaag agcaaaaaaa tggttaacat atgcagataa 180  
ttatcaaaca ttccacacac aattaatttt ttccaaat 218

SEQ ID NO: 6 moltype = DNA length = 227  
FEATURE Location/Qualifiers  
source 1..227  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 6  
ctgatccaaa ggcagaaagg gagggggcgaa atggaggagg gtaccgtgaa cagagtaagg 60

---

-continued

---

```

ggctggag aggaaattag agggaggcg ctggccagca ggcaaggcct gagcccttc 120
ccccggcggt gttggggggg gggccgtgtc catataaac gcgtggccgg cccctgcgggt 180
ccggagcccg cgccggcagg ggctgcgcga gccgatggcg gggcgca 227

SEQ ID NO: 7      moltype = DNA length = 182
FEATURE
source          Location/Qualifiers
1..182
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 7
gtggatgcta atgtgttaca ttttccag ttagacagt aattaaatcg gaagaaaaca 60
tggtttctag cagcagactg caatccacag atgaagctca tttttact tgactactg 120
aaacaaaaaa aaaactataa atatctgtgg aaaatattta cttaccaa ac tgggactata 180
ac 182

SEQ ID NO: 8      moltype = DNA length = 200
FEATURE
source          Location/Qualifiers
1..200
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 8
gattcaggct atcgattgtc gcaatcttg ataattcttc ttttacatat gtggggggg 60
atatatctgg aaagccttgg ccaagatca caacagaagg gtctgcggcc aatttggtaa 120
attcaatctt aagttggaaa aaaatactaa tattaataag ttctacaatt ataaagtctg 180
ctttttata tcttatagtc 200

SEQ ID NO: 9      moltype = DNA length = 217
FEATURE
source          Location/Qualifiers
1..217
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 9
cataggacag cttactacac ttttagago ttagatgttc taaatattaa tagttttaa 60
tgacattgtc ataaattgtc agtgtctgt aagaagctt attttaggt aagtgtttc 120
atattttac tggttctaaag gtggggaaa aaaaccagggt agtttactaa gtgataattt 180
gtttaaacaac ttaggtgtc gttcaaggaa tgataac 217

SEQ ID NO: 10     moltype = DNA length = 217
FEATURE
source          Location/Qualifiers
1..217
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 10
ccaacacaaa gtgttcctc ggtaactgtc attaaaaaaaaa aaagatactg ctattcaatt 60
cagcaaatgc atcatccatc caacagtcaa acaaacattt gtaccacaaq tacatttcaa 120
gaagccccaa atatctaaa tcttttgc cacttatggc ccagttatcc attaattcat 180
ccaaggatt ttctaaatca ttgttattac ctaatat 217

SEQ ID NO: 11     moltype = DNA length = 220
FEATURE
source          Location/Qualifiers
1..220
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 11
cataagatta aacagtggga tgcagacaaa ttgaacaca tacagactct ggaggtaacc 60
acatgccttc tggcttgc tggactttt gaaatacttt taagaatatt tacattaccc 120
actaagttagt ggtcaatgaa tgaggttac ttatttttt ttttctctc cagggtcac 180
accagggaaat atgggtttt gctgtaaagcc ccagtggaga 220

SEQ ID NO: 12     moltype = DNA length = 192
FEATURE
source          Location/Qualifiers
1..192
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

```

-continued

---

-continued

---

```

note = Description of Artificial Sequence: Synthetic
      polynucleotide

SEQUENCE: 18
caccataaag aagatgagat caataaaaaa attgaagata taactgagct taaggcgtac 60
tattgttaa tgccccccca caaactgttt tctcttgacc caagcttaca cttttatata 120
aactagtata ctttcatgg caacagcgt tcaggagttg taaaggaaat gtgaatggat 180
tatttcttg tcttcgttgcacg aacatcttgcacg a 221

SEQ ID NO: 19          moltype = DNA length = 161
FEATURE
source
1..161
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
      polynucleotide

SEQUENCE: 19
caaattcagc attgcaatca cacaatataa gactccataa atgattttat ttacttctga 60
tactctacct gtttactcg atccatcttc agctttttt ctgcagaata aacatcttt 120
accatccat tatctgcttt caacatctt aaaaaaaaaaag t 161

SEQ ID NO: 20          moltype = DNA length = 219
FEATURE
source
1..219
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
      polynucleotide

SEQUENCE: 20
aggctctctg agggagaagc caaagaggac tccttggatg aagagtttt tcatcacaaag 60
gcaatgcagg gcacagaggt aggacagaca gaccaaactg acagcacagg aggaccagct 120
ttcccttctca aagtagaaaga ggatgattt ccctctgaag aactactaga ggatgaaaac 180
gtctataaatg caaaacggtc taaaatggg aaccctggg 219

SEQ ID NO: 21          moltype = DNA length = 167
FEATURE
source
1..167
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
      polynucleotide

SEQUENCE: 21
ccattgtatt aattnatgtt attnatgttca attaattaaa ttgttattt atttttttt 60
taatctcaga acccttttggg agatgttca gaaatttgcatttgcatttcc tcggaaaatg 120
cttcaatccc ttacaatcc atatgtatctt caggtatgtat cggtca 167

SEQ ID NO: 22          moltype = DNA length = 220
FEATURE
source
1..220
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
      polynucleotide

SEQUENCE: 22
cccccaacca tgagaactac tgatctaatac taatcataaa ctcacccaaatc atcaggagaa 60
aagtctgtgc caaggcccaat ttcaatcaag gctttgttca agggagaatt gggcccaagaa 120
gtcaagatgtg aagacacaag acttaatgtg aaggcttcaat atgtgtcggt gatgtgtcat 180
tggaaaaaaa aggaacgggg aggttaagaaa aggttaaagg 220

SEQ ID NO: 23          moltype = DNA length = 204
FEATURE
source
1..204
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
      polynucleotide

SEQUENCE: 23
tttaggtcttg tgaatttgcattt ataatgtttt agaattaataa ccattaagaa agacaatca 60
gactatgttca aataatatttttttcca caattttatgg ttgttgcatttcc atacctggcc 120
agtactgttca agtctgttcaat atgtttgcatttca aggtgtcgat tacaagcaag 180
aggaacgttcaat atat 204

SEQ ID NO: 24          moltype = DNA length = 214
FEATURE
source
1..214
mol_type = other DNA
organism = synthetic construct

```

---

-continued

---

note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 24**

```
acatacacaac tcctatagca acaggattca aataaaact tagaccaga tttttttt 60
aactgtttaa gagaaaagga cttaagcca taaatgtaa agaccacata gtcgaacaaa 120
tggaaatgg ttggacctga cattctata gaaacgtaca tcattaacag tcaacatacc 180
aaagtgtgct gaaccactgc tactttact ttgc 214
```

**SEQ ID NO: 25** moltype = DNA length = 220  
**FEATURE** Location/Qualifiers  
**source** 1..220  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 25**

```
cccttacaca gcggcacggt gatcttgg catgccagct cttggccga ggccggccga 60
gcgcgcgtag agcgcgtcag cagcgccaaag gggccgcga gcgagggtcac ttccaaacagg 120
taaaaaactt ccatgtgtg cgcctcgcc cggtgcgcctc gcccctccagg cggcgccg 180
aggggtgccg gggggggggc ccacgagaga gccgcagacg 220
```

**SEQ ID NO: 26** moltype = DNA length = 199  
**FEATURE** Location/Qualifiers  
**source** 1..199  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 26**

```
tcttctaatt tctgttagctc cgatgtatgc ggcatttagtc gagcattgca caccatgaga 60
ttcttaactg catgtaaaat ttgatcttc tgagtgtca gagaaaggag ttccatatt 120
ccttctcgca ttcgaattt taagtctatt tttcttgaat tggtgcgtc ctaaaaaaaaa 180
aatgcatctt gaaaatataa 199
```

**SEQ ID NO: 27** moltype = DNA length = 168  
**FEATURE** Location/Qualifiers  
**source** 1..168  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 27**

```
tttcttcaa ccagggttt gaaagtctgc caagaaaaaa aaaatttca atgataattt 60
tctgggtatc atcctgaaat tatgaaaggaa cacaccaaa tacaatgtat caatgtcatt 120
atcaactacag agtcttaggg tacctaataat taattttcat attataaa 168
```

**SEQ ID NO: 28** moltype = DNA length = 191  
**FEATURE** Location/Qualifiers  
**source** 1..191  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 28**

```
aaccataatt ttttttagt tgaggtgaaa aaactgcttc ataagttcgat atgtttaaa 60
agccactgt tgagggaa tacagcgtat gtaattaaga gataaaccac gatagagtcc 120
tttcgaattt ccatgggttc catagacata cttcatagta tcccgcatgg tactgaaaga 180
caatgattaa a 191
```

**SEQ ID NO: 29** moltype = DNA length = 218  
**FEATURE** Location/Qualifiers  
**source** 1..218  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 29**

```
acccagcatg ctcagggacc acatcctacc ccccccaaat taggagcaat cacccacccg 60
tgggagttatc tggcccaaggc tttgtacactc tctgacgggc agttcggtct gctggcttc 120
agccgcctgt ctccgcacac ttgtatgtt tttctatggat tttttttttt ttaatttttttt 180
agcatgtgtt ctactccccc tttccagctt cccctgccc 218
```

**SEQ ID NO: 30** moltype = DNA length = 183  
**FEATURE** Location/Qualifiers  
**source** 1..183  
mol\_type = other DNA

---

-continued

---

```

organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 30
attacagatt gaagattctg aaatgtctcg tgtaattcga gtcagtgagt aagttgaata 60
ttctttaaat ttaatttattt gtaattgtt tgaagaactt ttaaaagtact tttttttct 120
taattcagat tatctttatg tgatcttgct ggtcagaac gaactatgaa gacacagaat 180
gaa 183

SEQ ID NO: 31      moltype = DNA length = 183
FEATURE
source
1..183
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 31
attcagaaaca tcaccccttga aatgacaaca aacaccaattt aggtatgtt tagcattcaa 60
aggttaaggc aacaaaactt gattaagaaa atatcatacc ttgtccagaa agaatatgaa 120
aagcagtctg gacacagtgt aagacttctt gatagcttag gtcctttaaa aaaaaaagg 180
ggg 183

SEQ ID NO: 32      moltype = DNA length = 191
FEATURE
source
1..191
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 32
cctggggcg aaacccttag agcccccaat actcaactgcg gcatctcaga tcagctggcg 60
cccaggcagg cacaggagaa actcacctga agggcatcca tgggggggggt gccccctcc 120
agccagatgg tgctacctgc aggccacagg taataacata acgtatgtt accatgaga 180
tgtaagcaat g 191

SEQ ID NO: 33      moltype = DNA length = 180
FEATURE
source
1..180
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 33
aataacaagaa ataacagatg tttggtagct tgatttagcat taaacaattt tacaatttt 60
accttgaagg aagtatagct ggttaacaacg ctggctccag ggaaagttaga gcagacacccg 120
gttctgact ttggctttct aaatattctt ggagaaaaaa acacaacatg atcatttacc 180

SEQ ID NO: 34      moltype = DNA length = 221
FEATURE
source
1..221
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 34
ccagtgtcaa ttcccttcca aagaatgccca agttctccat caagaaaagaa gctcccttg 60
tcttggaggc atccgaaatc ggttcagatc tgttgcagg agtggaaaggcc aaagggttaa 120
gcatttaggtt agtacatttcc ctccaagttt atatgtttaa atgcaaaagaa aaaaaagact 180
agctgacata ctttttaatc atcagatgtt tgtaggcata a 221

SEQ ID NO: 35      moltype = DNA length = 172
FEATURE
source
1..172
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 35
ttgtttcatt ttttttttaa taggtccctga ttgtttcttg aatgttccct ctactgagtc 60
ttactggatt ctgccaaacg ttaaacccctt cagtccttct gttaggtcggg cttcacataa 120
agcagtttca cacgggaaat ttatgtgggt gattgggttatacttttta ac 172

SEQ ID NO: 36      moltype = DNA length = 225
FEATURE
source
1..225
mol_type = other DNA

```

---

-continued

---

```

organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 36
cccccccca ctgagggcct tgcttgccct atggagctgt ggctctccac ttattgagca 60
cagcccttctt agtccacatg tgctgcctt ccaggtgcaa gtgctggca gtcaggacgg 120
tgcctaagtg gacccatgc atggctcagc cataggacat gccacacaag cagccgtgga 180
cacaacgccc actaccacct cccacatgaa aatgtatccct caaac 225

SEQ ID NO: 37      moltype = DNA length = 219
FEATURE
source
1..219
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 37
gggtcctgctc gctgatagcc atgagcatgc cagtggggcgc ctgggcatct ccctgacccg 60
cgtgtcagac ggcgagaatg tcattatatac ccacttcaac tccaaaggacg agctcatcca 120
gggtggggcct ggtggagcca tgctgggtgg cggtgggggg ggcagtggga acctcaaggc 180
ctctgctcat tctctccac tctgtccctg ccctgaaagg 219

SEQ ID NO: 38      moltype = DNA length = 192
FEATURE
source
1..192
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 38
ggcgccact gtcgtcatgg tgactgggt ctgggctgc tgcgccacct tcaaggagcg 60
tcggaaacctg ctgcgcctgg tcaggagggc gcaggggcac ggggtggggg tggtgcaagat 120
ggggccaaagg aggttgtcgc tgcaggctt gAACCTGTGC CTTGCTGTGC TCACCTGGGG 180
gggggggtcac gg 192

SEQ ID NO: 39      moltype = DNA length = 199
FEATURE
source
1..199
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 39
ggacttgttag ccatgtgctt gctttgcago gtgtgtacca actcctccca agtcgtctgc 60
aggccggagg aaggttaagct gcccctctgt gccagccctg cgggtggccgg gcccatctgg 120
gggaagcctg tggggcctt gatcggtggg ggggtgtggg ctccctctgg gctctgcccc 180
tttggtcccc ccccaagtc 199

SEQ ID NO: 40      moltype = DNA length = 217
FEATURE
source
1..217
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 40
tgtgcacccga gtctgtgctc cgtccctact tctggctcta gttaggccctc actatccct 60
acagggtgc ttagaggctc caagggttag aacatacctg tttaacactt cctgaatatac 120
ctggaaaggag aaaaaaaaaat gcagcatctg agtaagtgtt attcctcaga gggAACAGAC 180
atcgaactgg caggaggtt ctcttagggga agggaga 217

SEQ ID NO: 41      moltype = DNA length = 225
FEATURE
source
1..225
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 41
gagtcgtgc agccaaaggct ggagccatga gacggcccccc tggcaatggaa gaggccggca 60
gcgaagggtcc aggtggctgg ggtctatggg gagtccagga gtccaggagg ctgtgtgtg 120
cctgtatgtat ctggaaagggt ggtgtgtgt gtgtgtgtgt gtgtgtgtgt 180
gtgtgtgtgt gtgtattggg aagtggggggg ggggcattta tgcca 225

SEQ ID NO: 42      moltype = DNA length = 204
FEATURE
source

```

---

-continued

---

```

source          1..204
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide

SEQUENCE: 42
gaaataagat cagtaatgg aataaatgac aagactaattt gggttaccc ttcttgacag 60
ccttgaagaga cacaaggagga gatggagaa gaaatggtaa tggaaatgccg aacgggagaa 120
aagtttcatt ctatctgtc tcaaacatca ctggagcatg agaaacatgt tctaattaaat 180
ggaatcgagaa aaaaaaagaaa caaa                                         204

SEQ ID NO: 43      moltype = DNA  length = 211
FEATURE
source          1..211
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide

SEQUENCE: 43
cttcactac caggtgctat catcaagagg atgagaaaaaa aaaatttagag aagaatggag 60
aaatcatctt acctcataag tgctggtag accagatctg tagaaacacag ggagaaagag 120
ctctgatgtt aagaggatga caaataggtaa agcaatggaa aagacttagga aggatgcccc 180
aaagcggttag acttcagaag gggtccccag g                                         211

SEQ ID NO: 44      moltype = DNA  length = 175
FEATURE
source          1..175
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide

SEQUENCE: 44
gtttatgaaa taaacctcaa ctaatagttt aaagtctatg ttaaaaaaaaa agaattggca 60
aatattttt ttttgtaccg ttgtaaaatg tctttgttat cattttggtag tgattctgt 120
tcggtaatt tacacttcga gctttgtatg tcttcacata tagtaatgt agagt           175

SEQ ID NO: 45      moltype = DNA  length = 218
FEATURE
source          1..218
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide

SEQUENCE: 45
tctggcttac tgctccctgg agggtagag ggggtgcggg attgtggtag acctactaat 60
tgtgtatccc gtgtcccaagt atttggctaa aataactctaa tggagatgt tactgaagcc 120
ggcccccggga gtgggtgaagg tgggtcttcg ctgtgagctg tgcatgatct ttttttttc 180
tc tagcatct tctgtgcctg cccgaggggcc atccttct                           218

SEQ ID NO: 46      moltype = DNA  length = 190
FEATURE
source          1..190
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide

SEQUENCE: 46
gtcctgcccc aggctgcctt ggcagcgggc gcggtggctc tgggcacagg actagtgggt 60
gttaggagcca gcccggcaag tctgaatgca gctctatacc ctccctggcg aggggtctcg 120
ggcccccgtgc tgggtggctgg cacggctgtt gggggggggc tctgtgtgtt cggccctggg 180
ctagccctgg                                         190

SEQ ID NO: 47      moltype = DNA  length = 208
FEATURE
source          1..208
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide

SEQUENCE: 47
agggacatct tgcgtttttt ctgtccaca gggttcccgctt gagtccatcg ccgtttcac 60
ttgttaggtt ggctgggggg gcagggttag tgagccctttt gctggacatg cgcatgactc 120
gtggcccccc actcacacag acacaccatt cacgcttact gagaacctac ataatgcgag 180
gcactgtcctt aggcacgaat gttccag                                         208

SEQ ID NO: 48      moltype = DNA  length = 222

```

---

-continued

---

```

FEATURE          Location/Qualifiers
source           1..222
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 48
tatcttact tatttcttt atgttaaggta ttttacactc atattcaatt ttttgcttc 60
agctaataatgat tcttaaagaa cagaataatc agaaaaaaaac tctaacaatgg 120
acataaaaccacca tgaacaagaa aatacacccaa gcacgagtgg aaaggcaagt ctgttaggct 180
tgtcttttattt ctgtttgatt tcatttgc acagttatgtt tc 222

SEQ ID NO: 49      moltype = DNA length = 170
FEATURE          Location/Qualifiers
source           1..170
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 49
cccagtaaga ctagaaatat tcaacaacac aaaattcaact cagaaaaaaa aaactcttaa 60
gtataatcta ataaaagaac catataagaa taagaaagcc taatatctaa acctaccatt 120
ttcaaggAAC tcaggtcctt tcagcacact ggatttagag tgttgaactg 170

SEQ ID NO: 50      moltype = DNA length = 167
FEATURE          Location/Qualifiers
source           1..167
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 50
actggattt ctcccaatg catttcatgg ccatctgaat taactgattt ttttcaattt 60
ggtccttttc tgctgcatga agaacatcg gattgttaggc aacatcaatg actgtgttaag 120
catctgtgtg tgtaattgaa agggtatgaa atgaaagaaaa attgttg 167

SEQ ID NO: 51      moltype = DNA length = 202
FEATURE          Location/Qualifiers
source           1..202
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 51
ccatTTAAAT gaaataggta acaaattttg tgtaatgaa aacaacatTT TGATGTTGTT 60
tatatctact tagttataa tacaactttt ttcccttataa gtgagagagc agcagtagat 120
gcatcttaca ttgtatggat agatgaaactc ttcaaaagaa ccaatgttat tgaaaacttt 180
ctaatacaaa aaagagagtt cc 202

SEQ ID NO: 52      moltype = DNA length = 208
FEATURE          Location/Qualifiers
source           1..208
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 52
caggagaact acaattttt aggttaatgc tgataaaaagt ttccctctgag catcacctaa 60
ctcaggacac agccaactt caatgttataa tagtgacatc aatgtgttta cgatccttca 120
tcatttcaac tccccactttg gaccttcaaa aaaaaatag aacaataaat tagcctgctc 180
tctggcccatg taacacaggtt gaatcatc 208

SEQ ID NO: 53      moltype = DNA length = 217
FEATURE          Location/Qualifiers
source           1..217
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 53
aacactgcag ggccttcct tcaggcaacc gtacgggcaa cttctgaagg gataccagga 60
gtgacagaat agtctctatgc ctgggccttc gagaccgcattt acaaagaggg caaaggaaatt 120
ttactttttt agcttgcacatc ctggatcctt ttttttggaa acttgattt ggaagaggaa 180
cacagctgtt atgaaaccatc tctttgcaga gctcaca 217

SEQ ID NO: 54      moltype = DNA length = 219

```

---

-continued

---

```

FEATURE          Location/Qualifiers
source           1..219
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 54
tggagaccc tcgatgtttt gctcaaggac accctctaa gccacccctcg catgcagccc 60
tccaccagcc ccgtgtctac ccccaactcac agcaccagct cctgcagcag ccacctggtc 120
gcgcgtcaaga aggccaaaat tggctggaa ggagagatg cacacccgtc gcaccttaggt 180
ccaagggtga ggggtcccccc ccccaaccac cagcaccac                           219

SEQ ID NO: 55      moltype = DNA length = 171
FEATURE          Location/Qualifiers
source           1..171
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 55
ctggataaaa tattgtaat attacaaga accatgcat tttttttta actgcggtaa 60
aatttacccg tttcatctgt agcgggatca ttccaccatg ctgcaaagac acataacaag 120
gaatgaaaaa atgcagttagt gaaaatgtac taaggaaagc aggataaggg g               171

SEQ ID NO: 56      moltype = DNA length = 219
FEATURE          Location/Qualifiers
source           1..219
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 56
gaagaatcat attcattgct tgcccccggca tcagaatccct ctaaacactg ctttattttc 60
tttgcaggaa cttttttttt ctgataago tcttggttgc ttcccgtaag ttcttgaatt 120
tgaatttcta ctgcgtatgtt atagaatcca gttccctcagt tagagctatg 180
ggaggcagcgc cgatatacat gattagacag agttaattta                           219

SEQ ID NO: 57      moltype = DNA length = 162
FEATURE          Location/Qualifiers
source           1..162
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 57
ttagctcaca caattaaaaa aatatatgtt tttcttacct ctgaacttct tgggggggtt 60
ggcaagggtcc cccgcctctg ggtgcaccac acatctgtca tccactaacc tagaaggaaa 120
agcatgttta cattgttaac atccttgtca tattcataaa cc                           162

SEQ ID NO: 58      moltype = DNA length = 224
FEATURE          Location/Qualifiers
source           1..224
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 58
acagaaaaccc tcccaactgg atatccacag tccctcaaga tggctttgtc tctggagaac 60
tattttttt ctgcgttccc ttgtatgttgc ctgtactgttgc ggataaaaaag tcgagacatc 120
gaaaagaagca agatagggg caagatgttgc ggggttagaaa gaggaattgtt gggaaactcac 180
ggctttggcc agcgatccac ttgttctgtc tgctgtacg gggg                           224

SEQ ID NO: 59      moltype = DNA length = 221
FEATURE          Location/Qualifiers
source           1..221
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 59
gaggctaaacc tcataccac actcatcttgc tatccccctt gaggagaact gggacatca 60
cctggattcc tctagccaaac aaaagttctg gaacataaca cttaccatgg ggtccagggg 120
ggggaaaccc tggctggact gccccaggct ggggggctcc agctgggtgc tgctgttg 180
ggccctgcagt tgaccctgcc tgcccaatct gttcagaagg g                           221

SEQ ID NO: 60      moltype = DNA length = 221

```

-continued

---

```

FEATURE          Location/Qualifiers
source           1..221
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 60
cttccacaga tgcctcaatt tcactcgctg ggatccacgt gtcgttaactg ccatggaaa 60
tttagcagcac agtagaaatca atggataga atttgaaaaa agtaaatgtc agatcttta 120
ccctttccat ttacagaaaaag tattgggtt ttggcttatt tttttgtat tctattcatt 180
ataaaagtatg aaaaggcacat tacgaaacat ttcaaaaata g                     221

SEQ ID NO: 61      moltype = DNA length = 218
FEATURE          Location/Qualifiers
source           1..218
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 61
ggtagggaggt cagcctccc tataggcaag gggactgggt tacctaccag ggcgcgttca 60
ggaatggagc ggctcaaagt ccattcctgg ctattgtatc cagcctctgt aacggggaaa 120
ggcgcgttcaag agggggggct cctggcttag aaggacccc aacactgagc caccctccct 180
ccctctgcgttcccttcaacc ccccccac cctgcctg                           218

SEQ ID NO: 62      moltype = DNA length = 203
FEATURE          Location/Qualifiers
source           1..203
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 62
cagagacatt ctgtttccc aagcttgcgt tccacttacc ctatgttgg tgactttatt 60
ccccagggaa cagctccatc ctttccgtat cggaaagaact ttacattttt ctcctcttag 120
acatggctgc atatggcacc accatttctg ttccacttatt gaagctataa gagagaaaaa 180
aaatttgcgttcccttcaacc ccccccac cctgcctg                           203

SEQ ID NO: 63      moltype = DNA length = 173
FEATURE          Location/Qualifiers
source           1..173
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 63
tttctaagga gaaaaaacat gtatttata agaagataga tttttttaa agatttatga 60
actgcgaggc tggaaaagca cattaattc tagatgtga aattgttca tttttttcc 120
caggatctt gtcagtttgt taatgtctgat gtccttatt ggctagtggg cat         173

SEQ ID NO: 64      moltype = DNA length = 172
FEATURE          Location/Qualifiers
source           1..172
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 64
ctacaaggca acacaggcag gattatttctc accttacct gctctatag tcttggctt 60
attcaactgc atggggctca taggtattgg gtataaaggc ttgagagaat taaaaaaaaa 120
attaacattt ggcacctatg ataataacaa ggtattctgt cactttagt tt             172

SEQ ID NO: 65      moltype = DNA length = 221
FEATURE          Location/Qualifiers
source           1..221
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 65
gacatcagca tccctcccccc cccatctcgg agaaccgggg cttcttaccc agtgaaggct 60
ggggggctgg ctgtcagatg tggcaaaaaa ggtaactgaa gccaaggctg tgcctgcgt 120
caccaccacg acccagacgg gaacgaggcc gcctatctca tagacaccat ctgcaaagg 180
agagaaaaggc agctgatacc ttcccagcac ctcaacaccc a                     221

SEQ ID NO: 66      moltype = DNA length = 229

```

-continued

---

```

FEATURE          Location/Qualifiers
source           1..229
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 66
gacaaacgag tctcattggg attagctct actgagttt gggtgaatg tttagaatga  60
tagcatcaag tataatcttat aacctccatt tttttctaa tgccaccata gcagagacat 120
gtgtttatca gaccttgaaa ataaccaccc aaaagtcgat attaagaggg aaaaaaatca 180
gaagtcactg ttaaggacc agaaatttga agccatgtt gttcagcaa                229

SEQ ID NO: 67      moltype = DNA  length = 221
FEATURE          Location/Qualifiers
source           1..221
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 67
gagctggggg gggcttaggg ctgcgggctc cccagcaccc gcctcctgcg ctaccttagc  60
cccgagactc acaatgcggg acaggaggcc gatgaggattt ccgcgttaca cgctcctctc 120
gtggggctcg tcgatgtatca ccaccccttgc ccgcagcgcg aggaagtctt ggggggaggt 180
ccgggggttag gccccaccctt ccagcctaag ccccaaggcc c                    221

SEQ ID NO: 68      moltype = DNA  length = 183
FEATURE          Location/Qualifiers
source           1..183
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 68
aataaaatcat gataggcaact tgaagatgac ttgtaaaattt gatggctaac ttatcctg  60
gttttctaaa ctctgtttt tttcttctta ccacagcctc agttatgtt aatttcctt 120
atcaagcaac agtaaaaaaaaa gttgtctata gcttcctcg ggttggagta gggaccagct 180
atg                                183

SEQ ID NO: 69      moltype = DNA  length = 186
FEATURE          Location/Qualifiers
source           1..186
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 69
gtatggagg gtacttctcc aagggtctagg gagggaaaggg gagggggggt actggcttc  60
cctcccccctc ccctttaaag aaaaaatgtt acacaatgtt agaggactac tcttcaagta 120
tagaggcaag aaacaagtgc atgtcagatt gtcctgacct tcatggagca agatggctac 180
gtcttc                                186

SEQ ID NO: 70      moltype = DNA  length = 228
FEATURE          Location/Qualifiers
source           1..228
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 70
ggggctcccc tcccaggatgcc atcatctacc ttgcctataa agagggttga cttggcaag  60
gtatgcagggg tggctgtggc aggccgttcc actgctggac taggcacccccc cccagggtt 120
cttcgaaactt ggccttctt ctccggggcc ctggattttt ggaaccagac ctgtaccact 180
cgcttttgcg gccccaccctc ctccggagatg cagtcgagca tcttgcgt                228

SEQ ID NO: 71      moltype = DNA  length = 207
FEATURE          Location/Qualifiers
source           1..207
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 71
ctgtgcgcctt gcccccccccc actccctcagg ccaagctttt gggctggaga agggaccctg  60
gtccccctcca ggaggcactg ggctccggca ggcagggggtt cgccagggtt ggcgggaggaa 120
ggcaggactg ctggggatcc caggggatcc ctctgaggggg gtcagtcgt tctcatatac 180
cccagggttgc ttctcatcca gggctc                207

```

---

-continued

---

```

SEQ ID NO: 72      moltype = DNA  length = 170
FEATURE          Location/Qualifiers
source           1..170
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 72
aaatcagggg ttttccca acccttaat acattaacat ccctgggacc agactgttag 60
tgaacgttt acacacattt gcattgtaaa tgataattaa aaaataagt ccaggtattt 120
tttataaagg cccccctcc catttctgtat cgtttgtca gtctcttaggg 170

SEQ ID NO: 73      moltype = DNA  length = 201
FEATURE          Location/Qualifiers
source           1..201
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 73
gtgatctgc aggttactac tttgggttta aagccactgc agcgggaggtaaacaactg 60
agtcaaccag ctcccttgaa aaaaaagtga agaagaaatt tgattggaca tttgaacaga 120
cagtggaaatg aagtcaacca aaggagctg actttttta tgctatatacg aacagtgggg 180
atctttgtat aatttttata c 201

SEQ ID NO: 74      moltype = DNA  length = 220
FEATURE          Location/Qualifiers
source           1..220
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 74
ggggcatggg tagcaagaaa cttaaacggg tgggttttac acaagagctg tttggccgtc 60
tgatgtatcata tcagatccctt acctgtcagg taaattttat ttaacattt tattgataag 120
ttttatgcac aagttaactt tataccttaa tacctcttaa acttgaataa tgaaaatgt 180
ggcttacaaa aaaaagataa atgatggtgc tttccatag 220

SEQ ID NO: 75      moltype = DNA  length = 216
FEATURE          Location/Qualifiers
source           1..216
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 75
aaactacata attttagaaat gtttactgtt gaaaaatgtt tagttttat attctaaaat 60
ggacgtgttc cattatcaga tggaaaattt aactgtacaa tttagaatttc tacaatgtt 120
ctcaaaaatca ctccataattt tatcacttctt attatttttt ttagcctgcc ctcaagtatg 180
tgttgtatc attgggtggc tttgtgggtt ttgtaa 216

SEQ ID NO: 76      moltype = DNA  length = 209
FEATURE          Location/Qualifiers
source           1..209
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 76
acgtatgggtt tagagcacat cttatcaaag gacattttt ctattataaa caatgtcttc 60
caagccccctt gggggggatc ccacacgttc cagaaggacg aaaaagcaat cgagtgcac 120
ttatgtcgtt ctagatccct ctgttatcgtt cttgtttgtt aactcctggaa gagactatgt 180
ctttaagaag aaagccggctt ggtggtaag 209

SEQ ID NO: 77      moltype = DNA  length = 228
FEATURE          Location/Qualifiers
source           1..228
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 77
actggcttc atcagtcctt gaaacaactt caaaatatca aatacttca tcacttagata 60
tgtatcctat cagattttt taaagtaaga tacatgttcaaaaacacttc ggatttgtaa 120
acctcacctt cacatctgtt aacactgtca ttaatccatg attaatatcc agaagaactg 180

```

---

-continued

---

```

agagaaaaact ctgagagttc ttgttcttag agtctaattt tttttcc          228
SEQ ID NO: 78      moltype = DNA  length = 175
FEATURE
source
1..175
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide
SEQUENCE: 78
ataaccatga attatttttt ttagaggta gagagaaaata tattgaaagt tataaaatga  60
tttgttattt ctttttaggtt aaatggccaa atacatggaa acagttaaat tactagacta 120
cactgaaaaa cctctttagt aaaatttacg tgacattctt ttgcaaggac taaaaa 175

SEQ ID NO: 79      moltype = DNA  length = 221
FEATURE
source
1..221
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide
SEQUENCE: 79
gggggtccgg gctgggcctg caagactgca ggccgtgcctt ggcaggaaagg ctggagccgc  60
agtgtgggca cttccgcctt cggaccgcattt gcctgggcctt gctggggggg tagctggatt 120
tgacggcgtt tttggggagggg cggggctccc tggaaagggtt gtcacggctg gaccaagggg 180
ggagcctgccc tctgggcctt ctgggcaggc aggggctgccc a 221

SEQ ID NO: 80      moltype = DNA  length = 220
FEATURE
source
1..220
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide
SEQUENCE: 80
cgccccatcat tatatcctttt aacctctatg atcaaacaatc tccttgcattt cagctgcctt  60
cattcgggaa aacagaaaaca atttgcgtca atttctccat aaccttgattt ggctcatcg 120
taaaatggac agactagaat ttcttcagg tgagaatttga caggaaacta taactagaaa 180
agatgaaaggg tttttttttt ttctctgtttt aaaaagaaaa 220

SEQ ID NO: 81      moltype = DNA  length = 225
FEATURE
source
1..225
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide
SEQUENCE: 81
ttaacctgaa gtcgtttact tctcaacaaa attatgttgc ggggtgggggg tagggatgtt  60
tcttttaattt tcactttttt gggctgcattt gttttttttt taacctctgtt atttctgttt 120
tcttcaaatg acagactacg ttcttcgtat ggtggacgtt atggagatgtt taagtccgtt 180
cttcctctg gctatgcccc ttgtatgaaa gccttacttc caaaa 225

SEQ ID NO: 82      moltype = DNA  length = 217
FEATURE
source
1..217
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide
SEQUENCE: 82
gaggcatcac tggggccaca cagggttatgg gctgtatgggg ccagggttcc aggggtgggt  60
ggcgcccccccc gggcaggaga cacttcctt gactgcctgt ttctgtctt cctcggtccc 120
cagcccaactt ttccccccgaa gcctgtccca ctggggatgtt agtccgggtt gtcggccattt 180
ttaaaatgtt ttcccaagggtt tagaggaggc aaagggtt 217

SEQ ID NO: 83      moltype = DNA  length = 173
FEATURE
source
1..173
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide
SEQUENCE: 83
gagactgtgc agcgtacaca attatgtcag cagcagtttgc ttgtataata tatgacaacc  60
aagtttactgtt gtaagcttgc aaataattttt tatgtaaaaa ttagagatttt tatgaatttcc 120

```

-continued

SEQ ID NO: 84 moltype = DNA length = 212  
 FEATURE Location/Qualifiers  
 source 1..212  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide  
 SEQUENCE: 84  
 ctaatttgc tggttccaaat ttttttttc ccacagact ctatgtccaa atgatgaatc 60  
 ttttcttata actgaccagg ctctcaaaaa agaattgcta gcctgagaaa tgtggaatgc 120  
 ctggcttctc tgacttagtg tgacacagg tgcctccagg tgaacatacc tgtacaagtg 180  
 aagccatcac ctgtgtatcc ttcttgac ag 212  
 SEQ ID NO: 85 moltype = DNA length = 171  
 FEATURE Location/Qualifiers  
 source 1..171  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide  
 SEQUENCE: 85  
 cctataccaa ccggtaatgg ataaagggtt gatccactt gacctaggaa atcatcttt 60  
 gtctataaga gaaggaaagaa aaaaataat tagtaatacg ctcaatccct tttagaagagt 120  
 ttgtatgatg agtgcacca ttacacagaaa atgttagatg ttttttttat t 171  
 SEQ ID NO: 86 moltype = DNA length = 221  
 FEATURE Location/Qualifiers  
 source 1..221  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide  
 SEQUENCE: 86  
 ccactataacc cattattcat ccattcaagc ttgggtttgt cagggtgtgt ggccatgggt 60  
 cttygcacatcg gtttcacagg actactgtg acaatgtgg cgctatgtata tgccatactg 120  
 gagcttattg gagtgaatgc aaacggattc ctgcattcaa cagggtctgg ggggacacta 180  
 ctgtgcagt tagaatgtgg agtaccaatg ggtgtatgtc g 221  
 SEQ ID NO: 87 moltype = DNA length = 224  
 FEATURE Location/Qualifiers  
 source 1..224  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide  
 SEQUENCE: 87  
 ttggcagagc cccattccac gggaaagtg gtgggtccg ggcgaagcgc cgccattcc 60  
 ctcacactcc cccgggtcag caggctctgt gaccccccgg gggcccccaa ccattggcca 120  
 gtccacacccc ccccccacagc gctgagggtc caggagtggc actattctgt tacctctctg 180  
 cgccggaggcggc agtggaggagg tgatgaggcc gacggccggc gcgc 224  
 SEQ ID NO: 88 moltype = DNA length = 226  
 FEATURE Location/Qualifiers  
 source 1..226  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide  
 SEQUENCE: 88  
 tgagttact tgaaaactgc atgtctggtc agcatgacaa aagtcttaaa ccagggacaa 60  
 cacagtttttt ttcacagatg aattgtctta acaggcaggc tacttttgc gctctgttt 120  
 gaagagaata ccgtactaaa aaagaacaga agccacatcc aaccatgtt cagatttcag 180  
 atgtctcaca ttgcctccag aagacactcc tccccgagta gacaga 226  
 SEQ ID NO: 89 moltype = DNA length = 191  
 FEATURE Location/Qualifiers  
 source 1..191  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide  
 SEQUENCE: 89  
 ccccgacacc cggccgcgtca ctggccggg ctttgtgtgt tgcaggacca cggagggttg 60  
 atcqtqaaca tcactqccac cttqqqqaaac cqqqqcqacq cqctccaaqt qcatqcaqcc 120

---

-continued

---

```

tccgccaagg ccgctgtggg tatgaccacc cccccccgcc caggttgcc cacgtgggtc 180
cccaatgggc c 191

SEQ ID NO: 90 moltype = DNA length = 198
FEATURE Location/Qualifiers
source 1..198
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 90
gacaacatcg gctacgctg gattgccatc ttccaggtgg gggcaagat ggtgggacgg 60
ggaccctggg gcacggcagg ggagcgggtg caggactcg cccccccagc ccagaccca 120
ggggcacggg gaggaggagg aggggtcgta cgcccaag tcaagccact gccccccag 180
gtgatcacgc tggaagcc 198

SEQ ID NO: 91 moltype = DNA length = 225
FEATURE Location/Qualifiers
source 1..225
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 91
gcaggcgatg ctctggggca ccgggtgtcc atcctcccaa ggtgctgtct gcaggggagg 60
gtgccagcca aaagccaccg ggtagaaaaga aaacaaggac ctgtctaaga tggggggctca 120
tcggcaga tggggccctt gcctacctt gggggggcact tggagtcata gcagtgtctca 180
gtccccagat ggtccacagg tgcacccctt ttgtccctca tggca 225

SEQ ID NO: 92 moltype = DNA length = 206
FEATURE Location/Qualifiers
source 1..206
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 92
cgctgggtgc cccccccccc ccaggatgaa gggcggcgag ggggacgcgg gcgagcaggc 60
cccgctgaaac cctgaggcgc agagccctgc aggtctcgcc acgtacccggg agttctgtca 120
ccggcggtac ctggaccccg tggggccag tcagcactcg ctggggcgc tcagctggcg 180
ccgcctctac ctcagccggc ccaagc 206

SEQ ID NO: 93 moltype = DNA length = 178
FEATURE Location/Qualifiers
source 1..178
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 93
ccatttttgt aagtgaatta aagtctttt ggcctaaaat aatgttcttt aaaaaatgtt 60
tattattct gtgtatattt tttttttttt agtccaaagc acttagaaatc tccacaccac 120
taactgttgt taggtacatt aaggagaata gcccttggt gactccaggc tctacagc 178

SEQ ID NO: 94 moltype = DNA length = 217
FEATURE Location/Qualifiers
source 1..217
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 94
gcccagtacg agagcaagaa caagtcagcc cacaagaga tctacaccac cgtcacctgc 60
gccacggaca ccaacaat ccagtgttc ttgtatgtc tgacggacgt catcatcgcc 120
aaaaaacctgc ggggtgtgg actctactga gcccagccgc cctgccccgc acccttgcac 180
tgcttggccc ctcggccccc accaggc 217

SEQ ID NO: 95 moltype = DNA length = 209
FEATURE Location/Qualifiers
source 1..209
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 95
aagtgagagc actcttagtcc caggtcttgg ggggggggag cactcacctg tccccatgt 60

```

---

-continued

---

```

gacagcgagg gctatgacca tggcagctgt caccactcca ccccaggtct tcacttcagg 120
ctcgctgcgt ctgttgaaga cggcacact gctgacgtgg cactgtccag gtgaaggcga 180
cgcgtcatgtt gggaaatgc cctcaaccc 209

SEQ ID NO: 96      moltype = DNA length = 176
FEATURE
source
1..176
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 96
caagtctcca gggcccttct gtacgtact tcttggctgt gctctgcacc ccaccagcac 60
caggactcc aggagagcct gtcttttac tcaacagact gttgaagaag ctggccaaaca 120
cccccttcaact tgctgcattt tctaaggaaa attaaagaga aaaaaaagca tcatat 176

SEQ ID NO: 97      moltype = DNA length = 232
FEATURE
source
1..232
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 97
gagggttaatt tctttggcgc gctcgagctg accaaggggcc tcttgcocct gctgcccgc 60
tcaagggggcc gcatacgtagc tggggggaggc ccagcggttg agtgcggccccc cccactggag 120
aaaaaaaggag cccccctgggg tggggggaggc cttaggggagc cccttgccaa agctgagctg 180
cccccaactccc aatccatccg caggggacat gccatatccg tgcttggggcc 232

SEQ ID NO: 98      moltype = DNA length = 205
FEATURE
source
1..205
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 98
gcctccgggg ggctctcttg gcagctcggtt cattgtgacg aatgaagtcc acccccaggc 60
ctcgctcttg caggacacag caagcggtgg tggggcaga cagcattatg ttgtgtccct 120
aaacacccccc cccacaccc gagccctgcg caaccacat ctagccccca tacatcctt 180
caccttagct tcgggacctg gtggg 205

SEQ ID NO: 99      moltype = DNA length = 218
FEATURE
source
1..218
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 99
cagcctcccg gcccacccctg gggcaacccc aacctccaggc cacgcgcctc cccacctcg 60
gtgttggtag ccatgtcaggc cactttctgg aaggcatcca ggaaggccac agcagccagc 120
acggtgttcc tggggagagg gacaaacggca tggatggtcg gtgggtggc gggggggggg 180
cctctgcctt cccctccctt gctggggatc cctggggcc 218

SEQ ID NO: 100     moltype = DNA length = 229
FEATURE
source
1..229
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 100
cttcacacagg tcctcaactga gcagcgagtc gtcccccggt tcctctccgg ccaccaacca 60
cagtcctccccc gccagcacac ccaagcgctg gcccattggc cctatcatcg tccccctgg 120
ggggccacacgc gtgcccacgc ccccccctgt ggtgaccatc gctccaaacca aaaccgtgaa 180
tggtgtctgg aggagtggaa gccggcagggt gagtggggcc gggcaggaga 229

SEQ ID NO: 101     moltype = DNA length = 215
FEATURE
source
1..215
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 101

```

---

-continued

---

```

gctacactgg ggcctgggt gccttcagg aacggttcca cgggggggg gccccagtgt 60
ggcccttca gcccacgtgg gcccacgtga ggaaggctgg gatgtccctg acaggaacat 120
tccttgag cgccttcaca cgcagggttcc ggtcatccgt caacagcacc acctcccgca 180
gtagccggat tggctctct gtggccatga gagag 215

SEQ ID NO: 102      moltype = DNA length = 187
FEATURE
source
1..187
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 102
ctccggccgc gctggagcgt cccacccagt gcctacgcgt ggctccggct ccagcgcgaa 60
gccccggccc gagggggggt cttccgggtt ggcggccccc cataaggccc ccccccaca 120
ataatcgccc gggcctcgga gtatcttc ccacccgaaa ataagcgaac agttctgcc 180
atcgaag 187

SEQ ID NO: 103      moltype = DNA length = 226
FEATURE
source
1..226
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 103
agtaccggga tcacatccac agctgtggg ctgtcagcgt cctctggagc agccccatca 60
tctcagaatgt gtgggggggc aggaaggaggag gagggggccc ctttgggggtt ggttagctca 120
tagtccatgaa ggagtaggggg ggcttctggg acaccagagg aaagctggcc tggggccatg 180
tcctccatag tcatacgtcaga ggctggcaga ggtgcagggg ggggac 226

SEQ ID NO: 104      moltype = DNA length = 169
FEATURE
source
1..169
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 104
atataacttt aaaaaattta gtgtgccagg ttatttttt ttcataatgc ttagatgtga 60
aatatgttgtt gatgtttgtt agaaaacgtt cgtgactatg agcaccaggcc ccacccatgg 120
ctggagcagg agcgaacaca gtgggcacag tatagagaat ccgcagaga 169

SEQ ID NO: 105      moltype = DNA length = 212
FEATURE
source
1..212
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 105
aatatattaa cttttttttt atcttggaga aatttcaagg gcaaataatat ttgttgaaca 60
aatggataag taacaaacca acctctgcata attgaaggat ttctgtttcc agactctca 120
atttctttt actctctttt gattgagca aaatctcatc tctggatgca cgagcttctt 180
ctaattcacg ttggtaatcc ttcatcttag cc 212

SEQ ID NO: 106      moltype = DNA length = 216
FEATURE
source
1..216
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 106
gaggccaccc cggcctctca aagaggacgc acacctggga ctgcaggggcc agcacttgct 60
tctccagggtt cttctggcc tcctctctt ctcctctgtt ctccctaaaga ctgttcttctt 120
cctctccatg ctgcggatc cgactgtca ggtttagttt ctggcgtgtc tcctctgaa 180
gaagctctgt tgttttttt ttaaaaggca acactt 216

SEQ ID NO: 107      moltype = DNA length = 204
FEATURE
source
1..204
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

```

---

-continued

---

```

SEQUENCE: 107
ccccaggaa accacgtcg aagecctcg tgacatagat catccaagat gtgttaacggt 60
ccttcagggtt gtacacgcacg gctggctcat tcaggtgcgt cagcatggcc atgtttcga 120
tccctgtcgaa cttggggggg ttcatggcgt acacatcctc tggttgacc accagggtct 180
aaaaagaaag aggacacagt cttc 204

SEQ ID NO: 108      moltype = DNA  length = 202
FEATURE
source
1..202
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 108
ggaggtggag aaggcccctg cttggatctc acatgggctt taaaaaaaaa agtccctcaaa 60
gcaaagctgc cagcaggctg agtgcgtctc ctcctccca ctgtcttc cagaggaaag 120
cgccatgagt agcgaccgca tggactgtgg ccgcaaagtc cgggtggaga gcggctactt 180
ctctctggag aagaccaaac ag 202

SEQ ID NO: 109      moltype = DNA  length = 189
FEATURE
source
1..189
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 109
ggcccttcta ggggatctct ctggggctcg aagatcttc taaggcccct ctgcattcac 60
tcccctctag gaagtcctac ctgcgtcccc tttagcgcacg gtcccacccag gatgtggc 120
ctcatggaaag gcaccacgtc ataggggggc acatgctctg tctggggggg gaagcaggaa 180
gggaaacc 189

SEQ ID NO: 110      moltype = DNA  length = 225
FEATURE
source
1..225
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 110
ctactgcctt ggggggtgtca gtgccagccc cccacaatc ttttctgccc cccccaggag 60
gtgcaccatgt gtgtggctcg tgccactat aaggaccctc ctttctgcgt ggcccgctgc 120
cccaagcggtg taaaacctga cctctccatc atgcccattctt ggaagttcc agatgaggag 180
ggcgcattgc acccttgc accaactgtc gagtc 225

SEQ ID NO: 111      moltype = DNA  length = 196
FEATURE
source
1..196
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 111
tgtcattgac ttccacaacc ctgcggatgt ggtggaaatc aggacactgt cccggccatc 60
tgagacctgc aatgagacca ccaagcttgg ggactttgtt cgataccatt acaacttttc 120
tttgctggac ggcaccacgc tggtcaccc tggggcccg gggggggccg ggactggca 180
gggggggggg cacagg 196

SEQ ID NO: 112      moltype = DNA  length = 189
FEATURE
source
1..189
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 112
aggcgaccgt aaatctccgg caatggcgag tttatagcgg ggacaattgg cttccccagc 60
tcaacccccc cccaaaccat gcctccgaag tccctttgtt gtaaagctcc aggggtggaa 120
ggggggaaagg gttccccacgc actcacccgt gctcacgtga actttgcgc tccaggggtt 180
gacgacggg 189

SEQ ID NO: 113      moltype = DNA  length = 203
FEATURE
source
1..203
mol_type = other DNA
organism = synthetic construct

```

-continued

---

note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 113**

```
caaccactgt gatgcctcct aacaggctgt ctgtctctgc tgagctgttg cttgtcatac 60
ttccacataa agatctttg tctacctgac cagattctga aagatcttct cctgcagggt 120
agtctgttcc tcgtgcacct gaaaataaa tcttcttgcata acagcaaagt aaacttctat 180
aaaataacatt ttttttacta aaa 203
```

**SEQ ID NO: 114** moltype = DNA length = 188  
**FEATURE** Location/Qualifiers  
**source** 1..188  
**mol\_type** = other DNA  
**organism** = synthetic construct  
**note** = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 114**

```
agccaaggca ctcggggcag cgccaaacgac tcccagcacg accccgggtta agtttccagc 60
cgctgccccac cgccggccct tgggtctgcct ctccctgcag cggcggggac ggccgtgcgc 120
ggagccgggc atctcccgcg cccccccgcg tctcccgcgcc gccccccccc tggaaagtta 180
acacccctt 188
```

**SEQ ID NO: 115** moltype = DNA length = 208  
**FEATURE** Location/Qualifiers  
**source** 1..208  
**mol\_type** = other DNA  
**organism** = synthetic construct  
**note** = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 115**

```
agcctggggcc acccttctgc ctggggcgcc cagcctggcc ttgggggggg gggggcagaa 60
agagcaggcc tggggggctc ctacatccat catgggtgcg caccagtcgc ggggtgcgc 120
gtggccgcgc cccgggggtt gtaaacaca gcacttaggtt gaagtactcg gtgaacttct 180
tcaggccaga cacgaagccc acgtggat 208
```

**SEQ ID NO: 116** moltype = DNA length = 207  
**FEATURE** Location/Qualifiers  
**source** 1..207  
**mol\_type** = other DNA  
**organism** = synthetic construct  
**note** = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 116**

```
gttcaggccc ccccccaccc ccccaagatc ggtcacccctc cagcctccctc ctccgtcacc 60
tccttcggcc tggggcaagcc cgtgtgtccc ccgttattgtt tccgggtgtga gtgggtggca 120
ctggagacat tttccggcag cttccaaag ccgtgtctcactgcctc acagggtggct 180
gccttgatga tggtgtgac aggttaag 207
```

**SEQ ID NO: 117** moltype = DNA length = 221  
**FEATURE** Location/Qualifiers  
**source** 1..221  
**mol\_type** = other DNA  
**organism** = synthetic construct  
**note** = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 117**

```
ggccagcccc gggagctggg ggggtccctg agatccgcac aaggggaggga gcagggccct 60
ggggagcccc tgaccactcc accaggtccc accttttcc caaggacgtc tttctgtatct 120
gcttctctct ggtgagcccc gcttccttcg agaatgttcg tgccaaaggta gggcaaggct 180
ggggggccctt gtggggagag ggcttgcccc agtacctgccc c 221
```

**SEQ ID NO: 118** moltype = DNA length = 205  
**FEATURE** Location/Qualifiers  
**source** 1..205  
**mol\_type** = other DNA  
**organism** = synthetic construct  
**note** = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 118**

```
cagaagtcaa ttcatggcc agaccctcaa actgttggtg agtatacaca ccaaagatct 60
gcggggcagtgc tggtgtgggg agtgtttaat aaccagctca ccaggataaa aaaaaaaagcc 120
caatttggtt tttccaaagt gtaaatactc ctgttatgtc accaaatgcag 180
tcactgaaca tggagttggg aactg 205
```

**SEQ ID NO: 119** moltype = DNA length = 206  
**FEATURE** Location/Qualifiers  
**source** 1..206

-continued

-continued

---

```

source          1..229
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide

SEQUENCE: 125
gatcatgacc ttgtcaatgt ctacaaggag acgtgggtcg ggtcagtcg ggccttctgt 60
acacacaggc tgacctact gcacccctcg gtcttcgac ctgcacttag tgcccaccca 120
gtaccaggac acaaggggac ccacacggg aacgagaagg gggcgatggc gttgccggac 180
ggagagaaac tgggagccca ctgcctgccc cccccgggg gccgctgtg 229

SEQ ID NO: 126      moltype = DNA length = 213
FEATURE           Location/Qualifiers
source            1..213
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide

SEQUENCE: 126
ggcgcgctgt gagtccctcg cctggcgggg ccctcgaccc tgcctctgaaa aacctcacag 60
cctgggtctca cagcctggat gctcaggaca gtctcaagggt gggactggg gggcgagggt 120
ggcgctctgt cacagcaagg tccaaaccca ttgctctgac cctatgaccc tgaccccac 180
ccagatctgg ttcaacaaca aaggctggca ctc 213

SEQ ID NO: 127      moltype = DNA length = 214
FEATURE           Location/Qualifiers
source            1..214
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide

SEQUENCE: 127
ggaaagccgac cggggcgtgc ggagggatgt ggcccggaat attggtaggg ggggccccgg 60
tggcgtcatg caggatcccc gcgctggggo ctcgcctcg cgctcaactt ccagcgcagc 120
ggccgcggca goggacactc cagaacggcc tggccctgca gcggtgcctg cacgcgacgg 180
cgacccgggc tctcccgctc attcccatcg tggt 214

SEQ ID NO: 128      moltype = DNA length = 228
FEATURE           Location/Qualifiers
source            1..228
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide

SEQUENCE: 128
tctttctcac gaatggcaact cacctcaa at gtctctcatg gcttttgtc tgataatcaa 60
taccacagga ctcctggttt tcatgtaaaat tggagaaccc ggaatcatcc cttttgatc 120
ttactatttt ctgttcatgt gatattttt ctccacaaaat atcttgtgaa gcaactgtt 180
cattagttt tcaaaacctg aaacgaaaaa aagtgaaa 228

SEQ ID NO: 129      moltype = DNA length = 219
FEATURE           Location/Qualifiers
source            1..219
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide

SEQUENCE: 129
gaggaccctg gggagaaagt ggggtttcag tgaagctaa aggggtcatg gtgtggggc 60
gttagtggaaa ttgacttggg ggggttcagt catggaaaggg aatagggtg tggacaatcg 120
gggtcagagt gtcagaaaggg tcagaggta aagggcacag ggcacatacc acattgcctt 180
tggcaccagg agcaccatca attccagtca caccctagg 219

SEQ ID NO: 130      moltype = DNA length = 185
FEATURE           Location/Qualifiers
source            1..185
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide

SEQUENCE: 130
ggcagggcg gcgccgggtg cccccagaag tggccagca gtaccaggac atcatgcaac 60
gcagccagt gcaaggggca cagctggago aggggggtgt gggcatccga cggggtaggg 120
gtttggagat gggcatctgg tgggggagga gctccaggat aggcatgggg ggggaggtgc 180
tcagg 185

```

-continued

---

```

SEQ ID NO: 131      moltype = DNA length = 217
FEATURE          Location/Qualifiers
source           1..217
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 131
caagggagat aggccaggAAC aagggggggg ggtccttccc ccttaggGA aactgagaca 60
caaggAGTC ctggAGACCC acatGGAGT tcttgaggTC tcgctttccc agcggtaaa 120
ctggggccCA gagcggtggg agggtccG cctcAGcccc acttccACT acttcccAG 180
ggctccTTCA ccgtcatGGG agtcATCAGC ggcccc 217

SEQ ID NO: 132      moltype = DNA length = 220
FEATURE          Location/Qualifiers
source           1..220
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 132
tttagttact gtggatcccT tgcccaacAT cagagaATTc atactggAGA gagaccCTAT 60
gaatgtAAgg aatgcAAAAA aaccttCAGG cagcatGCAC accttGCTCA tcaccAGAGA 120
attcacatTG gggagtCACT gtccacACCC aaccCAGTCa atccaAGT cctatAGATC 180
ctgagtCTCA aatgtttCTA gaatttatac tgtttttt 220

SEQ ID NO: 133      moltype = DNA length = 195
FEATURE          Location/Qualifiers
source           1..195
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 133
ctgttgetcc tgcacgctcc ccctccctG tgaccGCCc cccgcGCCCG cccttggTCC 60
cttggggccG ggtagctGCC tgaaggGGCG gggcattCTC ctcagatATA aggcttggAA 120
ggcagcAGCT gogactCCCG agacCCCCCC accagaAGGC catggCTCC ccacgatGT 180
ccgggGCTCT ctccc 195

SEQ ID NO: 134      moltype = DNA length = 216
FEATURE          Location/Qualifiers
source           1..216
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 134
gggttggggg ggcccattac tgaggCCCCC cccagAGAAT caggcactAA ctgcacAGG 60
gagtgggtgc tcacatcccT ggacGAGGA gctaggGGTGC agctGGGATG ggcagAGAGC 120
atcgcccccc tcgactGCGC caagctGGGG agagGGGTTG caggcaggat catagatAGT 180
agatttcacG gagattCCA tcggccGCGC ctttc 216

SEQ ID NO: 135      moltype = DNA length = 217
FEATURE          Location/Qualifiers
source           1..217
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 135
agcgtccaat agtttccaa cgaaatAGCT ttcttacctt gggaaatgtcc ccccccAGATA 60
gttgacactc aggaacAGCA cggyacaATA atggctCTGC ctctgtCTCA tcatTTCTT 120
ggaaaaaaATG tgagatgtCA caaaggGTCT cagaaACACA gggtagtCC CTGTATAACCC 180
tggaaaacAA caacagaATT tttaCTATGA atataAG 217

SEQ ID NO: 136      moltype = DNA length = 209
FEATURE          Location/Qualifiers
source           1..209
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 136
tctccctccc ctgtccccAG ccccccAGA gctgtttcG cttcgTCCAG accaacatCT 60
cccgccTCTC gcaggAGACc tccgAGcAGC tggtaggcGCT gaagccCTGG atcactCGCC 120
agaacttCTC ccggGCTGCTG gagtcAGCTG taaaggCTC caggCACCCC 180

```

---

-continued

---

cactccttcc ctcctgtcc tcacggccc	209
SEQ ID NO: 137	moltype = DNA length = 232
FEATURE	Location/Qualifiers
source	1..232
	mol_type = other DNA
	organism = synthetic construct
	note = Description of Artificial Sequence: Synthetic polynucleotide
SEQUENCE: 137	
acaagagccccaaagggttgtatggagcac	ctgccccgggg tgccaccagg cccccccac 60
cccggtcgac cgccacccccc	aaatgtcaga gctgggtgg gccagcagcc gcctatgcca 120
ctggggaggc ctctggggct	ggagggcag ggccaggc ttattcccc ggtcagctc 180
aaggggctct	gggacccag gcctacgggc aagggtttgg agggggcag gc 232
SEQ ID NO: 138	moltype = DNA length = 198
FEATURE	Location/Qualifiers
source	1..198
	mol_type = other DNA
	organism = synthetic construct
	note = Description of Artificial Sequence: Synthetic polynucleotide
SEQUENCE: 138	
ggaggtccgggtgcggccccc	caccgcacc cccgcctca gaggccctaga tcccagtct 60
ccagggtcctc aggcccccccg	ctgeatctcc agggctcccccc tccagcatga tccctgtctg 120
tccgcagtgg gccgcgttgc	cggcctggac gcaggcgtgc accagctgca cgtccgtctg 180
cacgcgttgg acacccgc	198
SEQ ID NO: 139	moltype = DNA length = 212
FEATURE	Location/Qualifiers
source	1..212
	mol_type = other DNA
	organism = synthetic construct
	note = Description of Artificial Sequence: Synthetic polynucleotide
SEQUENCE: 139	
gaggatcctccctgcgttgc	atcccctgcg tacctgtgc cagagcaac aggattaact 60
gcagaatctt catggcctgg	aggggggagg agcggggccc aggttcctct gggaaacaagg 120
agggacatgg gccgcgtatca	ctttacgggg aaatcgggag gggggggctg gctcatgccc 180
tctcctctctt ccctcacatgc	ctccgcgtcc cc 212
SEQ ID NO: 140	moltype = DNA length = 227
FEATURE	Location/Qualifiers
source	1..227
	mol_type = other DNA
	organism = synthetic construct
	note = Description of Artificial Sequence: Synthetic polynucleotide
SEQUENCE: 140	
cactgtcccc acatggccaa	aatctactgt tcccctggga gggccttat cccctgcata 60
caatcaactt atccagcag	ggtggaaatc gcagttacac aagtggttcg acagccacat 120
cccccggtgg gcctgtaccc	ctcctggcta tgtatttta tggggccac aaaaaataaa 180
actgcccctt gatggaaatgc	ctaaagataac tattcaacc cccccctg 227
SEQ ID NO: 141	moltype = DNA length = 227
FEATURE	Location/Qualifiers
source	1..227
	mol_type = other DNA
	organism = synthetic construct
	note = Description of Artificial Sequence: Synthetic polynucleotide
SEQUENCE: 141	
cactgtcccc acatggccaa	aatctactgt tcccctggga gggccttat cccctgcata 60
caatcaactt atccagcag	ggtggaaatc gcagttacac aagtggttcg acagccacat 120
cccccggtgg gcctgtaccc	ctcctggcta tgtatttta tggggccac aaaaaataaa 180
actgcccctt gatggaaatgc	ctaaagataac tattcaacc cccccctg 227
SEQ ID NO: 142	moltype = DNA length = 202
FEATURE	Location/Qualifiers
source	1..202
	mol_type = other DNA
	organism = synthetic construct
	note = Description of Artificial Sequence: Synthetic polynucleotide
SEQUENCE: 142	
acccacccaa atttttttt caattactcg atgcaatatac aaatacttac taagtaattt	60

---

-continued

---

```

gtagttgtac tctttcccc tctgcccattatctgtaga agcgatgatc tgctaccacc 120
ataaatttac acgtgttctt catggatct gggtcagtc ttctttcac tcgatgaaca 180
agctctaata tgaatttgta tg 202

SEQ ID NO: 143      moltype = DNA length = 225
FEATURE
source          Location/Qualifiers
1..225
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 143
gccagcttc aacagtgtct tgatctcggt gcgcagggtca gcctgacttag agaagtccag 60
tttggggatg ttcaattttt ggaagtattt agtgggtctt ccagggttaa gctgattgtt 120
tatcttgaca atcaacttcat tactaaagtc cagtgttattt tttctgtgt gtaaaacttgc 180
cactgtgtttt gattttccct caataggattt tccaaaaaac agcat 225

SEQ ID NO: 144      moltype = DNA length = 224
FEATURE
source          Location/Qualifiers
1..224
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 144
tttagcagggt ttccaaaaaaaaa aagcagccccgg tggcccttca tctgtgtccg tctttttctc 60
tttctgttcc caggcagcag aagcattttt agtgcatttctt tttggggacg octatcttct 120
caccttacat gcaaggccggat ttactcttcc cccaaaggat gtgcaactgg cccggaggat 180
ccggggccctt gaggaggac tcgggtgagc tctgtgcaccat agtgc 224

SEQ ID NO: 145      moltype = DNA length = 187
FEATURE
source          Location/Qualifiers
1..187
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 145
tgaggaggagt gattgtggaa gcaactgaagt tcttaggtt ctactttttt ttttattttt 60
cctcactgaa gatggcagtg ctggagattc agggcaatgg agagtctca tatgtgcac 120
aggaaaggccat cagcagtgcc atccagcttc tagaagacc attacagagg gtaaggatct 180
tcaggat 187

SEQ ID NO: 146      moltype = DNA length = 207
FEATURE
source          Location/Qualifiers
1..207
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 146
ctgtgctata atccccacca aatcaatgaa gattacctgt ggacgattca actggaaaac 60
aactgggtgc tgaggccatt cactcagggaa aaaaaaaaaa agcaatgatc cagtgtctcat 120
tggaaagcttgc tcggattttt gacactttgg gattgtctcg gcagggaaagca ggtcagtata 180
taggagccctt ttttttcttagt attaaaa 207

SEQ ID NO: 147      moltype = DNA length = 179
FEATURE
source          Location/Qualifiers
1..179
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 147
aaagtataac taggagatac ctgatcttcc ggaaggatgaaa aaaaaaaaaagt tctacagata 60
aagttaaag gctgtgtatc aaatcttccat cagcttaacag tacataagaa accaaaggatc 120
atctttatgtt gtttgacaga agtcacacaa agcaacttta cttacctgga gaagtgtcc 179

SEQ ID NO: 148      moltype = DNA length = 188
FEATURE
source          Location/Qualifiers
1..188
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 148

```

---

-continued

---

```

tcaattaatt tagtgctt caaatataaa tcaagcttc tccagttcca caaggactcc 60
accacagtct ttagatggaa gaaaaatcac tggttttcca ttgtgcctca ttttgactc 120
ttcacttaga ctgcggatct tttttttt caaatccatc acagctgcat taatattatc 180
caccttaa 188

SEQ ID NO: 149      moltype = DNA  length = 204
FEATURE
source
1..204
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 149
acaaccagag aaattttttt taaagtattt cttacccccc ccagtttat atcttgaat 60
ggaaattgtc ctaccaccc ttgttacatc ttttcactt tgactgtat aggactgtat 120
ttgtcagaat caagtatgtt tctataatg gaaggataga caaaggaaag ggaaggatag 180
acaaaagtggg gggtagggg aggg 204

SEQ ID NO: 150      moltype = DNA  length = 204
FEATURE
source
1..204
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 150
ccctgggtgt tgcactgccc taaccccgac ctcccaactca ccttaaccag gggctcagg 60
ttctgggggg ggactctgtg ttcagtcgc ttttccgc tttgaaccaa gttttcaag 120
tcagccaggat tatccaggat aaatctggta ctgtttaatc tttcaagctg gggcagaaga 180
caagaccatc ccagccatca gccg 204

SEQ ID NO: 151      moltype = DNA  length = 197
FEATURE
source
1..197
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 151
gctgcgcgcg gccccccgc acgacctgtt caacctggca gtcaaggcagg tacccacccc 60
atgcctttcc tccctctggc aactcagcac atgcttgc tgcacccccc ccccccacca 120
accagggtca gccgataaccc acgggtgcag ccttaccagg tgccaaactgc agattctgtt 180
ccaatggggc agcaccc 197

SEQ ID NO: 152      moltype = DNA  length = 189
FEATURE
source
1..189
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 152
ttcgagacc cccacccatgc cggcccgac tgaccctctc cccaaacccgt acgccccgg 60
accgagctcc tcaccctggt gattgaacag tctccatatt ctatgtaa gaattccat 120
tcccgccggc cggacccccc ccctccagac acccggccgg cctggctcc cccggctcag 180
gctgaggggg 189

SEQ ID NO: 153      moltype = DNA  length = 212
FEATURE
source
1..212
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 153
tggttctcca gtgtacccatgg ttagatgttt aaaagaattha ttatccaa tcactgacta 60
caaaaatttat attatgtttt aatctgtctg gtaaataaaaa ctgttgaattt attttgaat 120
gtacatttt ttttactaca aagaaaagct tattctactt gatttacaat gagaaaaaaaa 180
aatcatgttag cattttatgg taaatagatt ta 212

SEQ ID NO: 154      moltype = DNA  length = 177
FEATURE
source
1..177
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic

```

---

-continued

---

```

SEQUENCE: 154          polynucleotide
ctctactcat tgctctctgt ctgagggttgg gatcattcaag cataccctct gaaaggagat 60
aggaactaca acgcctttc ttgtgttatt gatgggtcg gcccctaaaaaaa aaaaatcaat 120
taatgtctta agaacagaat cctaataaat tggtaggtata agataaaatc ttgctta      177

SEQ ID NO: 155          moltype = DNA  length = 207
FEATURE
source
1..207
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 155
gaaaacctat ggaactaaag ctgcttattt attacttttt ttttcctcca tcaggtatga 60
agatgataaa aaagaatggg ctggatgtt gaaggaaata cgttatgctt caggagctag 120
ttgcctagca acacgtttaa atcttcaa actgtctaaa ctgtgaacaa ggtgacaaaa 180
cataatagat tgggaggtgg tttgttt 207

SEQ ID NO: 156          moltype = DNA  length = 223
FEATURE
source
1..223
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 156
cttattaaata aagtttttt ccttactct aatcaaaacc ctacccttat ttggtagtgt 60
tacctaattt caattattgc tactattacc tctccagtc tactagctc ttggcgagat 120
aacactgtatc cagaaaattgg tctgtatgc actctggatc cagctcgat ctattaatga 180
aaaaaaaaata aagatttcaa tattaagaaa gaatacaggta aag 223

SEQ ID NO: 157          moltype = DNA  length = 219
FEATURE
source
1..219
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 157
atttataact attattctac aaaatacagt gaaaatgccaa tatctgttcc tttttcttca 60
cagggtgtta gtaatgaagc cacctgcatt tttgttcaat cctggcgagg gacagatgtc 120
atgtatccggg atccagtttgaa cacccttac ccctccaaagg atgaaggacc caagggtttg 180
tgtgattttt gtttcaattt atgaaataact gaatttcgtt 219

SEQ ID NO: 158          moltype = DNA  length = 234
FEATURE
source
1..234
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 158
gctcatgggc atatagtctc cactgcccctt tcggccactg ggcactgggg ccacccctgg 60
ggacatgggc atgttagccat cttccgtgtt gagggtggag ctgtctggc ggtgtggcc 120
cccccgacgc tccaaagggtt gcatttccat accctctctt gggttaggac ggggtggcac 180
gaaggcggag tgcctgtgtc cggcagtcg gcctccatg ccacccctgg gtgg 234

SEQ ID NO: 159          moltype = DNA  length = 199
FEATURE
source
1..199
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 159
ccctgtcccc gccccgacca gcctcttcca gcccggagcc cagctgcagg gagttctgg 60
cctggggacc ccccccaggag ctgcaggcag acggccggcc gctctgcggc ttctatgccc 120
gcctgtcccc ggggtccccag ggcgcctgtg cggccctggc gcaccaggcag tgcctgcaggc 180
gccccggcgcc atgcgggtca 199

SEQ ID NO: 160          moltype = DNA  length = 209
FEATURE
source
1..209
mol_type = other DNA
organism = synthetic construct

```

-continued

note = Description of Artificial Sequence: Synthetic polynucleotide  
**SEQUENCE: 160**  
 ggccaaggcc agtcgcagcc caggtcccg agccggcagggt tgggtgcgga ccatggccctc 60  
 tcacaagctg ctggtgaccc ccccaaggc cttgtctaag ccccttcca tccccaaacca 120  
 gtcctctgtc gggcctggc ctccaacct gcctcctcgc atcatggcag cgggggggct 180  
 gcaaggatgatc gggccatgta gcaaggata 209

**SEQ ID NO: 161** moltype = DNA length = 221  
**FEATURE** Location/Qualifiers  
**source** 1..221  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 161**  
 ggtggccggc gccactcccc ccccacagcc aggttatccctt ctgtttttt tccttgtcca 60  
 caccgttca cacgaaggcc aaggccccac tgcccaaggcg gctcatggta ctgtactttt 120  
 gggagtagtc gcccacacag gcccacaacc cttccagttc cacagccttgggggctct 180  
 caaacccaggc tctggctctg agcacctgca atgcagaaga a 221

**SEQ ID NO: 162** moltype = DNA length = 203  
**FEATURE** Location/Qualifiers  
**source** 1..203  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 162**  
 tggatggaa tggccctggc ttggccagtt ggttggtagt agggtttggc tgaattcacc 60  
 aaacgtgctt tttttttctt ctctctctct cgctccctcc ctccctctct agccttcata 120  
 tgctgagttt gatattagt gaaatgtgt gggtttgatc ttcaaccaag gaatgtatctg 180  
 gtgagttatc catttcatctt ggt 203

**SEQ ID NO: 163** moltype = DNA length = 184  
**FEATURE** Location/Qualifiers  
**source** 1..184  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 163**  
 atcgggtatc aggccaaaatc atatgggtct tgacgtactg gggccccccc agcatggctc 60  
 ccagtaaaaaa gagacttttctt aagaatttcag gctttgaattt catatccctt tctgtctaaat 120  
 accacagcaa ttccagcccccc tcttgctgaa ggtgagttact aagacgattt cctctgttaat 180  
 ctgg 184

**SEQ ID NO: 164** moltype = DNA length = 224  
**FEATURE** Location/Qualifiers  
**source** 1..224  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 164**  
 ggaccaatca ggaacttgc ttctcaggaa cgtacctgtt tattgcgtgg atgggggggg 60  
 gtggggcaca ggacagccgt gcacaggcgt agtagtacc aatagactcg atgatgtctgg 120  
 cgaccacccgc actgagcatg ccgtatgacac cggccgcaga cacgggtggc agtccccactt 180  
 gaactgtggg aggaaaacaa ccatgaatgc tccatagagag gcag 224

**SEQ ID NO: 165** moltype = DNA length = 179  
**FEATURE** Location/Qualifiers  
**source** 1..179  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 165**  
 ggagatgtt gtgttaaccat aactgtatctt ttcaaggatt ttgtggaaat gcttcctgaa 60  
 aaagatgtc gctatgtttt gtatgtatc agctttgaaa caaaaagatc cagaaaagaa 120  
 gagttgtatgt tttttttgtt gtaagcatgt tagaaatatt gaggctctgt aaaactcat 179

**SEQ ID NO: 166** moltype = DNA length = 223  
**FEATURE** Location/Qualifiers  
**source** 1..223  
 mol\_type = other DNA

-continued

---

```

organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 166
agtcttggcc actgtttctc ataagaccag tcggatattt cctttgcc agaaatgctt 60
atctcatgt aggactggag gcaggcaggc tgacaatttc tgcaatagag gtgggttgt 120
tcgttcttt tgcaagaggc cctacaaattt ccactgttac ccatgcattc aaggatgtgg 180
cgtttgctg ccaaaggaaa aaaaatcaa atagtttac cac 223

SEQ ID NO: 167      moltype = DNA length = 215
FEATURE          Location/Qualifiers
source           1..215
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 167
atgactaca gacgcattga attccccca ggtgttgggt atattcccta caatgacttc 60
catgtcccgac gacccccc agtatatacc aacggcaaaa aacttgcattt tatttaccag 120
tatggtcaca tttagacccaa cgacaaacact gtcagttgg ggggcaata ccgatatgg 180
gagatccctg agtccgaggg aagcatcagg gacct 215

SEQ ID NO: 168      moltype = DNA length = 227
FEATURE          Location/Qualifiers
source           1..227
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 168
cctgagggtt gttgggggac gtgagtcctg ccccgctgc cagtcocca aaggccccca 60
ccggccgaccc cggcgcacct gccaggacat ccaccgactt gttgggggg ccgcgtaca 120
agatgcaggc acccccccacg cgtttctccc cacggggggg gcctccggc tgcacctgct 180
cctgggttta tttatgaaac cagaatacga tgtggaggcc cacgatc 227

SEQ ID NO: 169      moltype = DNA length = 202
FEATURE          Location/Qualifiers
source           1..202
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 169
aataattatc ccctggtaa atttttttt ttaaattcat caaactaaaa aaactacaca 60
ggctactaat ttcaatagg acctaccttc agcatctagg acaatatcta ctcccaaggcc 120
acctgtttct tccaaacagc tttagcaac atgaacttcc catttagata catcaatcac 180
tcgggctgtt aaggacagga aa 202

SEQ ID NO: 170      moltype = DNA length = 202
FEATURE          Location/Qualifiers
source           1..202
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 170
tgttcattct ggggcctcac attctgggc tgtaaaatgg aaataaaaaac acccacatca 60
tgcatgttgg gaggagata aagttagata cacacagtat gtgccaccaa atgattttc 120
ctctctgtcc ccccccataa ccccaactgt ggcacccggc atgtccctcc tgcaacctgc 180
ttctctgtat ccgttgcgtg ag 202

SEQ ID NO: 171      moltype = DNA length = 229
FEATURE          Location/Qualifiers
source           1..229
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 171
ccggccatcac ccagcggtatgttgc ctggcgtgaa ggacaagggtg tgcatgcctg 60
acccgttgtc agacctggaa aaaggggccgg ctgtggcag ggagggcatg cgcacttgc 120
cctccccacc aggtgttac accacgttca ctgaaaacc actatcacca ggccctcag 180
tgcttcccaag cctggggctg agggaaagacc ccccaagcag ctcagttag 229

SEQ ID NO: 172      moltype = DNA length = 183
FEATURE          Location/Qualifiers

```

-continued

- continued

SEQ ID NO: 178  
**FEATURE**  
**source**

moltype = DNA length = 180  
 Location/Qualifiers  
 1..180  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 178  
 ctgtgaacag cagcggagac gtgaggctgc ttatggcaat caggacctcc acctggaaaa 60  
 aaagcgcccc ttgcaggacc acggccggag cctggggacc caggacagcg ggctggcca 120  
 gggctggggg ggcgggcatg gccgggcaact catgggcata gtcccaagct ggttccctgc 180

SEQ ID NO: 179  
**FEATURE**  
**source**

moltype = DNA length = 218  
 Location/Qualifiers  
 1..218  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 179  
 gaacagccct ttcatttctg ttttttttcccagggtta ttccaaagat agtttcata 60  
 actccacaaa ctctgaattt gcagctgaag ctgagggtca gaatgtataca attgaggAAC 120  
 ccaacaagg ccagaaaagg aagagggata gacttcgaga ccagggctcc acaatgtatct 180  
 acctgaaggc tatccaggc atcctggga aatcgatg 218

SEQ ID NO: 180  
**FEATURE**  
**source**

moltype = DNA length = 223  
 Location/Qualifiers  
 1..223  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 180  
 cagtgcagg gagatcatgt ttgactgtat ggatgttctg tcagttagtc ctggggcaat 60  
 cctggatttc tacactgcga gtcggcttc ctgcgtctc caggagaaag ctctcaaAGC 120  
 atgttcaagt ggattgaccc aaaccgaatg gcagcatcgg cacactgtctc aatgttagtt 180  
 tatttttttc ctttcttcta ccaagaaaaaaa aaaaattgtc tct 223

SEQ ID NO: 181  
**FEATURE**  
**source**

moltype = DNA length = 220  
 Location/Qualifiers  
 1..220  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 181  
 aacaccaggc tagcttttca ttgttttact ttttttttccatggtag aggcttttac 60  
 ttaataacta ccatcttaaa aatctgaaag tattcgttgg tttaatcacct caagattttA 120  
 agttacgcctc ttcacatTTG tgacatctca tttaaaaagag tattgtgatt ctttcagca 180  
 ccaggaatat ctttagatgt tttagagaag cttaacttttgc 220

SEQ ID NO: 182  
**FEATURE**  
**source**

moltype = DNA length = 219  
 Location/Qualifiers  
 1..219  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 182  
 gctggggcagg gaacaggaaaa tataagggtt accactacag cagaaccagg aaactcctga 60  
 gcaccactta tctttttttt ctgttgattt aagttgtcaat ttatgaccccc aacaaggAGA 120  
 ttcaatgttca agaaggcctcc aaaatgtatg aagatgacaa agtacaaaata catgtacacg 180  
 ttgtccctccc acttgggttg catgttgacc tgtgaccagg 219

SEQ ID NO: 183  
**FEATURE**  
**source**

moltype = DNA length = 226  
 Location/Qualifiers  
 1..226  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 183  
 atgttagtgtt acaaccagggtt gtacggaaaa atattccgtt aataccactgt agtgacagt 60  
 ccccccccttta aagatggaaat cctggcaga aaccttggaa gcccctttttag gacggatTC 120  
 agggaaatgtatg agctaaaaaaa accattttgc tccatcatca aagcacttac agtttagca 180  
 atttaaaaaaatgactgttgc tcatcatctt accataaaag agaaaa 226

---

-continued

---

```

SEQ ID NO: 184      moltype = DNA  length = 179
FEATURE
source
1..179
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 184
agaattttgtt ttcattctgt ggctctaccc agaaccagcc tgagaatgtt caaagataac 60
acccctttttt ctcttaattt ttttttttt tccttatgcc aaaggcttca cccaaaaaac 120
tgccctgggtg tgccgcagtt tgctgagaca atgatgtgtc ctgtgtgtc cgacgctgc 179

SEQ ID NO: 185      moltype = DNA  length = 223
FEATURE
source
1..223
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 185
gtttgtgggtt ggggggggac cacaacccc ggcctgtccc tcttgcattac ataggcttcc 60
cgcccttagaaag cgcagccaga acatgcgcgt acggatccggg ttgttagtctg ggagctgtcc 120
aatggggccct gtggggatag ggtggagcgt cagggggggg tacctggcct gcacagggac 180
acctggccct gtgcgcacc cagtcggag atgctggcca gag 223

SEQ ID NO: 186      moltype = DNA  length = 233
FEATURE
source
1..233
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 186
gtgcccccttc atctctgtcc cagtttgcct gaaccagtcc ttgcagggtt ctggatgtt 60
ggcaggggccc aaaaaactt cactgggtt gtgtgggtt gttttggac ttggccatgt 120
ctgcacccctt aagacaaaag ttggatcatg ggttaaggcc tcggccaggctt gggccactc 180
tgccccccccc attctggtca gaatggcct gcaggcagag aatgtgcagc ctc 233

SEQ ID NO: 187      moltype = DNA  length = 192
FEATURE
source
1..192
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 187
ggggtgtggagg gggggggccga gagccggggg ggggcgttcc cgagagccca gaccgaggcc 60
ccctgcggcc cccgcattccg ccctcacccgtt ggtgcggcc ttggctggcc atggctggc 120
gttgcgtcc gtccgcgcgc tcggagccct ccgcgttgcacc ctggccggc aacgctcgct 180
ggaccaagag ag 192

SEQ ID NO: 188      moltype = DNA  length = 211
FEATURE
source
1..211
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 188
ctcacaaatc attcaggat aggttcttta gaaacctgtttt agatttaaac aaacaaacaa 60
acaaacaaac aaacctttt ttagaccatt aaataacctt cttcataaac tattttcttag 120
ataatgttct actcttataa taatgcagac acaagtcttct tgacacagaa gtcatacatt 180
ttcatgtatc ataaacatttta tgaaaaaaaaa c 211

SEQ ID NO: 189      moltype = DNA  length = 208
FEATURE
source
1..208
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 189
ttactttctt aatttgtcaa agtttagtatt taaaaatttc agatgactga tcattttttt 60
acaaatttca gatgactgac ccagtcacat taaatgttagg tggacacttg tataacaacgt 120
ctctcaccac attgacgcgtt taccggatt ccatgttgg agctatgtttt gggggggact 180

```

---

-continued

---

tccccacagc tcgagacct caaggcaa	208
SEQ ID NO: 190                    moltype = DNA length = 170 FEATURE                          Location/Qualifiers source                          1..170 mol_type = other DNA organism = synthetic construct note = Description of Artificial Sequence: Synthetic polynucleotide	
<b>SEQUENCE: 190</b> cgggggttctt gccatggctc agcttacgt gtcatccctg taagtgttag aaccaatgtat 60 tcaaatatca actgtggcag cagtttagtt accttaggtg caaaaaagag aaatattttt 120 ttttgttcat atatataatat tgtcgattt cataatgaaa tcagaagaga 170	
SEQ ID NO: 191                    moltype = DNA length = 221 FEATURE                          Location/Qualifiers source                          1..221 mol_type = other DNA organism = synthetic construct note = Description of Artificial Sequence: Synthetic polynucleotide	
<b>SEQUENCE: 191</b> tgcatctgca caggcttccc tggaggact tcatcaacag cacagttaca catcaaagtc 60 cttagtctt ccccgctgg cgcctatttt cacttctact ttcaatttca cagacagttt 120 tacagcactt tccatttcat tcttgacaat ctgagctacc taaggaaaaaaa aaaaatgttag 180 ttaacaaactt gcccagtcaa agaatgttcc ccaaaccatc c 221	
SEQ ID NO: 192                    moltype = DNA length = 216 FEATURE                          Location/Qualifiers source                          1..216 mol_type = other DNA organism = synthetic construct note = Description of Artificial Sequence: Synthetic polynucleotide	
<b>SEQUENCE: 192</b> tctttctaac cctactgtat ttgcattta cacctttcc ttcataactaa cctggatgtc 60 ccggggccgt tcataaggact gatacgcgtat ttttcttcc atgcttgcca gttctgttt 120 caagttcaag atttcattctt ggtggagctc tgtaggtca tttagtgtttt cttccatcg 180 ttcacatcta aggagaaaaaaa aaaatgttagt aattttt 216	
SEQ ID NO: 193                    moltype = DNA length = 212 FEATURE                          Location/Qualifiers source                          1..212 mol_type = other DNA organism = synthetic construct note = Description of Artificial Sequence: Synthetic polynucleotide	
<b>SEQUENCE: 193</b> aaggccctgg ctgatttcat ccgcaacaaa ctcttccca tcaaggcata tctgacaatc 60 cactcgtaatcccaatgtat gatctaccct tactcatatg cttacaaactt cggtgagaac 120 aatgtctgatg tggtaagtagt caaagtatgtt ggtatgacat ttactatgtt agatttttt 180 aattctaatac ctgaaaaaaa aatcattata ag 212	
SEQ ID NO: 194                    moltype = DNA length = 200 FEATURE                          Location/Qualifiers source                          1..200 mol_type = other DNA organism = synthetic construct note = Description of Artificial Sequence: Synthetic polynucleotide	
<b>SEQUENCE: 194</b> cattaattca agctgtgtctt ctcgtatttc ttccattttt tcaattttgtt gagtatattt 60 ttgattaattt agattgccaa catcacagag agccgtttagt aaactttttt ttttctgtt 120 taatgtatca ttaagcttccat taaaatactg gagaacatgtt tccttatcgc cttggatcat 180 tttctcagaa tgatgtttt 200	
SEQ ID NO: 195                    moltype = DNA length = 228 FEATURE                          Location/Qualifiers source                          1..228 mol_type = other DNA organism = synthetic construct note = Description of Artificial Sequence: Synthetic polynucleotide	
<b>SEQUENCE: 195</b> ttggaaaga ggttagcacat aactgcgtt aatttttttt ccaataatat tctctctagg 60 gccaagagcc tggctggcag agctcggtt ggtccgttccatccccca gaagaggccc 120	

---

-continued

---

```

cgtctgctgc cagccctggc actgagccca ctggcaccc agagccactg ggaacaggca 180
ctgtctcagc agaactgtcc agccccatggg ctgccaggaa ctccggac 228

SEQ ID NO: 196      moltype = DNA length = 218
FEATURE           Location/Qualifiers
source            1..218
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 196
tcataacctgt gatcattaca gtccttatec gggcttagat gagctttcc aagacagggt 60
ttgtctcttc cttcaatcga ttctccaaga tcagcagccc cagaataatc aggtctgatt 120
ctaccgttcc cctgaaaata catatataga tgtatttgaac agaaataaca catattacag 180
ctttttttt gagacggagt ctcgtctgt catccagg 218

SEQ ID NO: 197      moltype = DNA length = 220
FEATURE           Location/Qualifiers
source            1..220
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 197
catttcggc agtgtgtctt gacacagtc ctgaaaaaca cgaaatcccc gactctgcga 60
tgactcacct gtcgcgtgcc ctgagcttca cgcacgggtt ctgtgtgtc aaggcagcgt 120
gaggacagt ctgcgagtc cctgcagcag gatggtgcoc ttgtctgtcg octaccggtg 180
ggcgggggggg taagccaagg ctgtggctgc tcaagtccctg 220

SEQ ID NO: 198      moltype = DNA length = 193
FEATURE           Location/Qualifiers
source            1..193
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 198
gagagcagcg acagcccgcc caccagcccg ggctcagccct ccagcccccc tggacccct 60
ggggacgacc ccccccggca gaagtctccc agcggggccct tctgcactcc ccagtcccc 120
gtctccctcg cgcaggaggg caccggccag atctggaaga ggcagtctca ggaagtggag 180
ggccggggca tcc 193

SEQ ID NO: 199      moltype = DNA length = 229
FEATURE           Location/Qualifiers
source            1..229
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 199
cttagctgcc cttaccagggt tcttccagag atagtctctg cacatgtgt gtgggggggg 60
gccttcgtgt gcgggtgtgt ctgtgtctgt ctcggccacc cctgacagtc 120
ccacccctgc ccacaggctc ccgtgttggt cctcttcgcg ctggccctcc tcacccgt 180
cgtcttcctg gttgtctaca aggtgtacaa gtatgaccgc gcctgcccc 229

SEQ ID NO: 200      moltype = DNA length = 199
FEATURE           Location/Qualifiers
source            1..199
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 200
aaatgctgat gattgtcagg atttggaaat ccctgtccac tcagacaaac cctaaacacg 60
ttgtcataata ctcataatggg cctcgtaact ccagtcgtca cagtcgttgt agccattaca 120
ttgcagtttc cccggggatgc agattccact ggcacacaga aagttctcac cccttccaca 180
gagcactaaa aaaaaaaggg 199

SEQ ID NO: 201      moltype = DNA length = 191
FEATURE           Location/Qualifiers
source            1..191
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 201

```

-continued

---

-continued

---

```

SEQUENCE: 207
tggtgtcttt actgcagatg aaggatttg gggcagagca cttggcaggt catgaagggg 60
tccaacttct cgggttggta aacgtctacc tggacaaga agagagatc caacctcgag 120
aaaaaggggct gagtttgatt gaggctacco cgaggatcaa ttggcagaaa agataacttg 180
ctcaaggcgc tgatgggct tggaggaggt aacccccc cccc 224

SEQ ID NO: 208      moltype = DNA length = 212
FEATURE
source
1..212
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 208
gcmcagegcc ttcagcaacc ggagaagcat agttgcacgc gacctgggtgt gtgatctccg 60
agtgggtggg ggagggtcga ggaggggaaa aaaaataaga cgttcagaaa qagacccgga 120
aagggcctt tttttgggtg agctgggtgc ccagtgtgc ctccgatcct gagcctccga 180
gcctttgcag tgcaatgtcc cgcctgctgc at 212

SEQ ID NO: 209      moltype = DNA length = 217
FEATURE
source
1..217
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 209
gtttatgtt cacatgaaca aaatccaagg ttatagttt attctaaaat aactaatttt 60
catctaataca acttgtctgg aaaccttaca gactcatgaa aatttgtctt aacggtaaga 120
agtttcacgc tctctttatc agatcgaggat ctatttggaa aaaaaaacat gaaattaaca 180
ccacagtaat cagcattgtta tgtttggaa agtattt 217

SEQ ID NO: 210      moltype = DNA length = 218
FEATURE
source
1..218
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 210
atataattttt ttggagtgtc cagttccacat gttatTTTt cacttgactg aagtctttatg 60
tcttcTTTgtt gttttttcca cattttgggtt ggtgaacaga ttctcattga tcacttgtct 120
ctagaaggat tatagtaaaa aaaatacaag accgaaatat gagacatagt tatgtttaaa 180
caacaacggaa aattaatgtca cagttatcacat tttaggat 218

SEQ ID NO: 211      moltype = DNA length = 211
FEATURE
source
1..211
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 211
gaaaagggggg gggagactgg catattgtttt gttccagcga tgggtttt ctttcggaaag 60
ctcggtggaa aatgtggctg taattttttt gtcgtgtttt gtttacacgc cgcctgtcca 120
gtgttatgtt gtggcaacgg gcagttactcc aaggggccgtt gctgtgtttt cagcgggtgg 180
aaggggcaccg aatgtgtatgtt gcccactacc c 211

SEQ ID NO: 212      moltype = DNA length = 226
FEATURE
source
1..226
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 212
ggctggcgtg gggataagg cctcagaaaa gatgacacca caccggccca gacactagct 60
cttcagaaac gccccggttt acctctgaag ctggcggagg aagtgttcca ggagttgttg 120
aatagggtgtt ctctcattttt cggctgacaa tttaaggaaa aaaaatgttga atacaagact 180
ttggtatccc ttcatttagag agctccaaatg tcacagttt tggatgtt 226

SEQ ID NO: 213      moltype = DNA length = 199
FEATURE
source
1..199
mol_type = other DNA
organism = synthetic construct

```

---

-continued

---

```

note = Description of Artificial Sequence: Synthetic
      polynucleotide

SEQUENCE: 213
aaggcacttca tatgcctctg aaaaaatag attccaaatt tgacaattga ctcctttat 60
tgcagaatgc agcaactaac tggacatagt catagtgctg gatggttcca acagtattta 120
cccatgggac agtgttacag ctttttaaa tgaccttctt gaaagaatgg atattggcc 180
taaacagaca caggtatgt 199

SEQ ID NO: 214      moltype = DNA length = 169
FEATURE           Location/Qualifiers
source            1..169
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
      polynucleotide

SEQUENCE: 214
catcgaatgt attctcatcc acttcacaat gatccaaatt tggaaactgt gtatacgctt 60
tgtgaaaaaa aagcacaggt atggcagaaa ttgggtttt atagtagaat tctatatgtt 120
tcactgaata aatatgtgtt tgtcatttc ttatcctgtt ggtttaatt 169

SEQ ID NO: 215      moltype = DNA length = 215
FEATURE           Location/Qualifiers
source            1..215
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
      polynucleotide

SEQUENCE: 215
gtcctgcaag tcggtgagtc acggggcago cgcgagccgt ctgtccgtcc gtcagtccct 60
gggtgggggg ggggtgccccg cctttctct gcggacagcc cctagttccg cgcgcgccccg 120
caccctcgcc ggtgtcccta ctggcccgga cggtcctctt gaggcggagcg cccattttct 180
ccctcccttc ctggccccc gtcgcggccg acggg 215

SEQ ID NO: 216      moltype = DNA length = 168
FEATURE           Location/Qualifiers
source            1..168
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
      polynucleotide

SEQUENCE: 216
caagacagaa agacagacgt agatctaaga gcccacataa aaaacgctct aaatcaaggg 60
agagacggaa gtcaaggagt cgttgcatt cacggtagt ttttagagaaa ttaacaataa 120
ttttttttc ctcaaggttc tgtagtgct aaggataat attttaat 168

SEQ ID NO: 217      moltype = DNA length = 191
FEATURE           Location/Qualifiers
source            1..191
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
      polynucleotide

SEQUENCE: 217
atttggggaa atagatggga gaaaatgtga ttatatttac atgatttctt ttttttcag 60
gttaacagact ggggtgaccc atcatttgcatt gattttcttag agttagtgg cgtctctact 120
attactgcacca catcatttagg tgcataaaac tcaagtcata gaagaaaaaa tgggcctct 180
acatttagaaa g 191

SEQ ID NO: 218      moltype = DNA length = 207
FEATURE           Location/Qualifiers
source            1..207
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
      polynucleotide

SEQUENCE: 218
ttctgtgtct tactgttaacc tagggacattt gctgcaagaa actgccttgtt aggtgaaaat 60
aatgttctgttga aaatcagtgttga ctttggatgt tctcgtaac agatgggtgg agtgtattca 120
tcttctggct taaagcagat tcccataaaa tggacacac ccggaaatgttctt taattatgg 180
aagaatagac cacatttttt tttaat 207

SEQ ID NO: 219      moltype = DNA length = 186
FEATURE           Location/Qualifiers
source            1..186
mol_type = other DNA
organism = synthetic construct

```

---

-continued

---

note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 219**

```
gattgggcta aacactttgg gcagactgaa attgtggatc ttctagaatc ttacaggtaa 60
aactttgtac tattttaaat taattctacc gtctcttaaa aaagagctc accatataaa 120
ttatgggcta tc当地atgt a g c a t t t t a t t c a g a g a t t t t t t g t a 180
ttctaa 186
```

**SEQ ID NO: 220** moltype = DNA length = 204  
**FEATURE** Location/Qualifiers  
**source** 1..204  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 220**

```
gttataatt ttaatataa ttttatgc tctactagaaa gttgtaaaa tttccat 60
gttaactttgtt ct当地ccgata caagcagtgt gatgctgtgt gaaaccagggt ggacatttac 120
aggtaaaccc ccccagggtg ttgacacaga ggaactggca gttatgtgc tttgttgac 180
attcatcaag gtctgaaatt agag 204
```

**SEQ ID NO: 221** moltype = DNA length = 177  
**FEATURE** Location/Qualifiers  
**source** 1..177  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 221**

```
attgatggat tattttttcc atttctttaa caggaacatg aatctgaagg tggaagaaca 60
cctttatgta aagctgcaag agctggat tttgtgcact tgca gtttct tattagcaaa 120
ggtaaagaaa gggcaagtga tcattccaa gaggctacag tttatggaaa aaaaat 177
```

**SEQ ID NO: 222** moltype = DNA length = 183  
**FEATURE** Location/Qualifiers  
**source** 1..183  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 222**

```
ctgccaatcg cggagagatc aaaggatgc actacccact atggagggcc tggacttggg 60
agacacccctt aggccgtacc ttgttccctt ctccccatag ggtctgagtc tgccacccac 120
gtgccccgtgg cccccccac cccagectgg caaccagaga tcaaaccggc gctgcctgg 180
cac 183
```

**SEQ ID NO: 223** moltype = DNA length = 197  
**FEATURE** Location/Qualifiers  
**source** 1..197  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 223**

```
atttgccttc agagaaatata ttgctgtatgt gaagaagttt ctccaggtaa gaagtcattg 60
ctacatggttt agaacatcac tactgtctt ttttggttttaa gaatactgtt cttcttaaag 120
cgatataatgtt ttgttgccttca ctaccatgtac catgttaat atgatttattt atttatcctc 180
tggatccctt atcattt 197
```

**SEQ ID NO: 224** moltype = DNA length = 230  
**FEATURE** Location/Qualifiers  
**source** 1..230  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 224**

```
ttcttaactca ccagaaaaaaa acctctgcac ttcatcgccc gagcaagacc agggcaagcc 60
ccggaccccttc accacgaaatc cctctccacc ttccgtgcc aacatcatcg tctcttacgc 120
ggtcggcggtt cggaaacaaac tgcaaaaggccgg ggaggaccacg aactggagac gccaattaaag 180
ctgtccctcg cctccggaggc gcgccggc cccgcaccgc ccccccacggc 230
```

**SEQ ID NO: 225** moltype = DNA length = 220  
**FEATURE** Location/Qualifiers  
**source** 1..220  
mol\_type = other DNA

---

-continued

---

```

organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 225
aggaggctga ggcaggcaag ttgcttgaac ctgggaggca gaagttgcag tgagctgaga 60
tcgcgcagcc tgggtgacag agcgagactc tgcgtcgaa agaaaaaaaaa aagaaagtta 120
gcatgtgtaa ttcttattta acctttata ttgtatcaatc gccagaataa aagctacttg 180
gcgttaaggc taaaagacca agatctgaga gattcaaggc 220

SEQ ID NO: 226      moltype = DNA length = 231
FEATURE
source
1..231
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 226
aaacccaat accacaagcc cagttccgt gtcaaccaca caatctcagg gaaaactgaa 60
atgtacaatg caaaaaccgg caaacagggt ctatccctcg ggggggcgtc caagtccact 120
ggggccgcgc tcaatggcct caaggctgtt ggcagggtga accccaaatgc atgcactaag 180
gagggtggggg ggccggcagct gcggggaggcctt ctgcagctgc gggaggggctt g 231

SEQ ID NO: 227      moltype = DNA length = 228
FEATURE
source
1..228
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 227
aataacatat tattatagtt gaaagatttt ttttttagtgtt agtgatttca aactgtgttc 60
atttatatgc aacccatgaa cgcctccctt ctccatgtc tcctttttagg tttagggcc 120
atgttgaagc tttatatac atttaccaa cttagggtgg tgagttgggtg ctgtggcag 180
cctgggtttaa aaatctccac tggagagaag acacagcagg attcagat 228

SEQ ID NO: 228      moltype = DNA length = 225
FEATURE
source
1..225
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 228
ctcagcgcaca gaaaaaggga agggtgtctc atcccccttc ctccctctct ccctgtctcg 60
agcccttagcc atggccgagg caggggctgg gctgagccgg accgtcaactg agacaacgg 120
taccgtgaca accgagccccg tgagaaaggc gggggggcgg tgctgttttag gggctggga 180
gatactggga gggaggggac agggattaga agagtttttggag 225

SEQ ID NO: 229      moltype = DNA length = 190
FEATURE
source
1..190
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 229
gtctggccccc ccaacaccct tccctgtgc ataccctgtc tgaagttaaga agctctcaag 60
ccgggtcaagc tgacccttgg cctcagagcc aaaatcctca gggtgagagg ccagccagg 120
tgacagtaca gagatggcta ccctggaga agggaatcag ccaagggtga gaggttaagc 180
tgcagcctgg 190

SEQ ID NO: 230      moltype = DNA length = 229
FEATURE
source
1..229
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 230
acttgcttggc ggctgaggta agggaaaggc gaggttttttcc ctagccggcc tggccctgg 60
gggttacggc cgatccctgc cccttggatg gccagactat atcctggac ctctctctgg 120
ctgtatagccg tgggggggggg gtgtcggttgg ggtatcttggt gacatggggg gcagtggcc 180
cagaacaccc ctttctccat ctctcccaa tactccagg 229

SEQ ID NO: 231      moltype = DNA length = 220
FEATURE
source
1..220
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

```

---

-continued

---

```

source          1..220
               mol_type = other DNA
               organism = synthetic construct
               note = Description of Artificial Sequence: Synthetic
                      polynucleotide

SEQUENCE: 231
aacacaatga tctgaaggct tttcttggct gtaaaatgac taaatgaact gggtgactta 60
caagtcctt tcaagcaaaa aaaaaattct atgatttgat gatgcactta atttttagtt 120
gaaaacatca tgacatatac tttgttttc cttcaaaggaa tatttgacata tctgccat 180
aggagaaggaa tatcgaagcc aagctgactc aggtgattga                           220

SEQ ID NO: 232      moltype = DNA  length = 218
FEATURE           Location/Qualifiers
source            1..218
               mol_type = other DNA
               organism = synthetic construct
               note = Description of Artificial Sequence: Synthetic
                      polynucleotide

SEQUENCE: 232
gggctgcccc gaaaaaaaaat acaagacgtc caattaaattt gaatatcaga taaacaacg 60
ataatattttt agcataagtc tgtcccatac aatattttggg acagacttac actattttaa 120
gtatccaggta taactggaaag tctttgtttt ttatggcca gatctggtaa cccttaggtc 180
ctccagttt aggaggaggaa gaggacaccc cctccacc                           218

SEQ ID NO: 233      moltype = DNA  length = 224
FEATURE           Location/Qualifiers
source            1..224
               mol_type = other DNA
               organism = synthetic construct
               note = Description of Artificial Sequence: Synthetic
                      polynucleotide

SEQUENCE: 233
acggacacgca cctcggtcaa atccaaggcg gagaaggatg ttagaccaaa ggggttccag 60
tttgcattttt actgcaagta actctcaat ggttgggttta atctattaca aaggaaaaaa 120
atgccaggaa ttatatatcat gaggataaaag accacccccc gttctccac tcccctccct 180
cattttttga agaagtctcc caccagatgg gctgtggcat ttgc                           224

SEQ ID NO: 234      moltype = DNA  length = 215
FEATURE           Location/Qualifiers
source            1..215
               mol_type = other DNA
               organism = synthetic construct
               note = Description of Artificial Sequence: Synthetic
                      polynucleotide

SEQUENCE: 234
ataattttagt atgttttagt atgttttagt aatatttctt tttttttcaa tttacccat 60
ctgaaattac ttcatatgtc ttccagctg tctttccatc gttgatacat aattgccaaa 120
gtagccaaatg tgaactccct acttttaga ttcttgatgc actactttgg atttttcaaa 180
ggcccttcga ttctatgcaa tctggatga tacag                           215

SEQ ID NO: 235      moltype = DNA  length = 205
FEATURE           Location/Qualifiers
source            1..205
               mol_type = other DNA
               organism = synthetic construct
               note = Description of Artificial Sequence: Synthetic
                      polynucleotide

SEQUENCE: 235
ctgaccatag gcaaacgcct agctaatgt ctacatccag cattaccagg tggagttcat 60
cggaaggcgc ttgaaacata tggaaatatac ttggaccta gcgacttgcc 120
aaagatcttt ttatataatg gtaaaatataa ttttactata tatataacaaa taatataaag 180
aaaatgtatt gctaaaatctt tttat                           205

SEQ ID NO: 236      moltype = DNA  length = 193
FEATURE           Location/Qualifiers
source            1..193
               mol_type = other DNA
               organism = synthetic construct
               note = Description of Artificial Sequence: Synthetic
                      polynucleotide

SEQUENCE: 236
cacatacata cacaaaacctt ataaaaaaaaa ataaaatggaa gaggctatga gaaagagaat 60
tcaatgtcaa taaagaacaa gcatcacccaa aaatgtcagc cagttccat attcacacca 120
cctgtccatg ttccccacac acatagacag cactaggacg tatccgtcgc tttgtgtgac 180
caggagatcg cggtg                           193

```

-continued

---

```

SEQ ID NO: 237      moltype = DNA length = 208
FEATURE          Location/Qualifiers
source           1..208
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 237
tgcaatgc aaatttaggg gtatgcgtt ggtccaggt tgattctagg gaattgcgt 60
attcacttaa cttccgatt gatgacaaat agtttcttt tatattattt gcattatctt 120
ctatcttga ggacatacca gaagatctt gtgttattctg tgctaccgt gataaatgt 180
tggaaaggtc aaatatattt ttttgaat 208

SEQ ID NO: 238      moltype = DNA length = 183
FEATURE          Location/Qualifiers
source           1..183
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 238
aatctgtgcc aagcaaccag gggaaatctg caatgaagct gacccctgtg acccacaca 60
agggctgtat ttttgcgtact cagtagacag gcctaggatc gagactggag tttttttttt 120
taagtgtctt cttctggacc tgctggaaaa gatttgcgtt agacagaatt tttttttttt 180
gca                                         183

SEQ ID NO: 239      moltype = DNA length = 223
FEATURE          Location/Qualifiers
source           1..223
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 239
caatgcggccaa aattcagaat agtctgatta acaatcttgc ctggagatat ggtcaattt 60
ataaaggattt ttatgatatac gagtgcgtat atcatctaat ttgaagcac cgtcatctt 120
cgaggatttca cttgcgttgc ccaggcttcc attttcatgtt gtgtcgttgc ctgcctttt 180
attaaatcctt gaaagaaaagg agaaaaaaaaa agcagacactt ttc                                         223

SEQ ID NO: 240      moltype = DNA length = 190
FEATURE          Location/Qualifiers
source           1..190
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 240
caacatgtat tcccacttac tggaaatccctg atagccaggtt gaataactttt gacgagatgt 60
tccaggtaaa acctgcagaa acatgtttac acaattttttt tactaaataa tttaaagaaaa 120
atattttat attatcatat tttcccaaag atttatattt aaaaaaaaaa actccctttt 180
tgatTTTTA                                         190

SEQ ID NO: 241      moltype = DNA length = 185
FEATURE          Location/Qualifiers
source           1..185
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 241
gtaccttagc cacatggttt tagaataattt ttcacaccttca tcaatgtctc tgccaggaa 60
aaactctgttgc acatcatcatc ataactctca tgaagaaggat aaacatccct taggcctgaa 120
tagccttccat tcactctgca atgggtttcc aaggcctaaat ttatagagggaa ggaaaaaaaa 180
aagtcc                                         185

SEQ ID NO: 242      moltype = DNA length = 190
FEATURE          Location/Qualifiers
source           1..190
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 242
ctgaatccag tgctttccctc cgcatgttgt tgctttctca aagaagatataa tttttttttt 60
cttcattttt ttttaccttgc ctccactgtc ttgggtttttt tttttttttt cttccaggat 120
acagctgcaaa aacaaagatataa aggtttaaag aagagttccaa gacagaaatataa ctgtttctc 180

```

---

-continued

---

tcaaaaaaaag	190
SEQ ID NO: 243	moltype = DNA length = 218
FEATURE	Location/Qualifiers
source	1..218
	mol_type = other DNA
	organism = synthetic construct
	note = Description of Artificial Sequence: Synthetic polynucleotide
SEQUENCE: 243	
atatacatgcc tggtgtccat	gtttatTTTt gtgttatatat atatatatat gaatattctg 60
agttaaaaaa aataattcaa	cctatTTac ctgttcaGG tttgttgata ccggataaAG 120
cttcactttg atgggtattc	tgattgtat gacttctggg tgaatgcaga cgctctggat 180
atccacccag ttgggtggta	tgagaaaacc ggccacaa 218
SEQ ID NO: 244	moltype = DNA length = 206
FEATURE	Location/Qualifiers
source	1..206
	mol_type = other DNA
	organism = synthetic construct
	note = Description of Artificial Sequence: Synthetic polynucleotide
SEQUENCE: 244	
caGATTGATA ggggtatgtc	atTTTTTTT tcagattggg gtaagacAGC taccaatATC 60
atcaagaaaa gttctgtataa	tcttagaaaaa tctttgtctt caagctccat tggttccaaAC 120
tcaacctatc ttacatccaa	atctaaatCC agctctacCA cctatttCAA aaggAAatAGC 180
cacacaggtg agtctaaaga	tgtctt 206
SEQ ID NO: 245	moltype = DNA length = 226
FEATURE	Location/Qualifiers
source	1..226
	mol_type = other DNA
	organism = synthetic construct
	note = Description of Artificial Sequence: Synthetic polynucleotide
SEQUENCE: 245	
ctgggggtat tggtaggcaat	ctgcttgaac tcattgtcc agttggcct cccatagaAG 60
gttcttctgt ttaaggctgt	aatcacgtca gtatTTCTGT cccagtgtaa agctatgtGA 120
agagccctaa gtaaaatttGT	ggcatttagat gacatttaga aaaaaaaaaat aagacttcta 180
taataaaaga ataaagtaat	catgttcatt tggtagattA attaaaa 226
SEQ ID NO: 246	moltype = DNA length = 180
FEATURE	Location/Qualifiers
source	1..180
	mol_type = other DNA
	organism = synthetic construct
	note = Description of Artificial Sequence: Synthetic polynucleotide
SEQUENCE: 246	
ttactgaacc caagcaaaaa aaagtgtctt	gtcctaaata gtaattataa ttatgttact 60
tttcagtatt gtacagtcta aaactgctac	agcgTTTCTt tggttgat gtgattactG 120
aacttttCG tggcactca	aggatttGat ttctacaAG caattgaacc 180
SEQ ID NO: 247	moltype = DNA length = 220
FEATURE	Location/Qualifiers
source	1..220
	mol_type = other DNA
	organism = synthetic construct
	note = Description of Artificial Sequence: Synthetic polynucleotide
SEQUENCE: 247	
gaagtggaga atgcagttac	tctcaaaacc tccaaAGAAT ttactagtat taaaactcca 60
gttaattctc tccagtgtt	agtcacaggG ttttcatTC tggccaAGC tgactgaagg 120
tcttgcataa catctctgtA	gttatttca tagtgcac tccaggtaa cagaaaatca 180
taggcaggTT tccacatcat	ctataaAGAA aaaaaataga 220
SEQ ID NO: 248	moltype = DNA length = 168
FEATURE	Location/Qualifiers
source	1..168
	mol_type = other DNA
	organism = synthetic construct
	note = Description of Artificial Sequence: Synthetic polynucleotide
SEQUENCE: 248	
aaatacatac tttagaatcc	catttgcTCC ggaataAGt tctattgcat tgatactggT 60
gggcagcatt gtaataattc	tgcagcatac tctgctggg aagataactac 120

-continued

---

```

ctttaaaaaa aattgcaaag gcacattatt taaattaaaa tgtaagca          168
SEQ ID NO: 249      moltype = DNA  length = 216
FEATURE           Location/Qualifiers
source            1..216
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide
SEQUENCE: 249
aatctgtgga agtccatccc ttcaaccagg ttgcccagt ctccaggc tc aaacgaggc  60
agccctgggt cacagattt cctgggggc aaatccaat gaggaggggc cccgtggacc 120
cctgactctg agcaggcccc ccaagggtc ccacatcctt gtccctgcaca gccccccccc 180
agccccaggc actcacgcgt aggccctcaa gtcacc                      216

SEQ ID NO: 250      moltype = DNA  length = 208
FEATURE           Location/Qualifiers
source            1..208
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide
SEQUENCE: 250
tctgtgttct tgcgtccccc agcttggaa gccttata cccagtggag aagctccaa  60
ccaagctctc ttgaggatct tgaaggaaac tgaattcaa aagatcaaag tgctgggctc 120
cggtgcgttc ggcacgggtt ataaggtaag gtcctggca caggcctctg ggctggggcg 180
caggccctct catggctcg tggggagc                      208

SEQ ID NO: 251      moltype = DNA  length = 214
FEATURE           Location/Qualifiers
source            1..214
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide
SEQUENCE: 251
ggggccacgc ggctggggcg ctgttaggagg acctcgggg tcggagaaa ggggtgtcca  60
gggcgtgatc ggagcctccg ggagccgcg gaccgggctc ggggtgggtg ttgaggggcc 120
caacgcgcac ttctgccccg gaccccccac ctggaaagcca aggccgtcag 180
ggcccgaaac cgcctcacc tggtaggcac gctc                      214

SEQ ID NO: 252      moltype = DNA  length = 206
FEATURE           Location/Qualifiers
source            1..206
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide
SEQUENCE: 252
ggagatgtat tacttagaaat tcagtcttc acaaaaaaa cattatatcc gagttctgtt  60
aaagccccac ccagaacggc ccataaaggaa atgaagacat aggtataggat catggtaga 120
aactgttcca gtttagcaca gatgtcgagg cagaaggcta atttaagtaa cattgcaatg 180
aggtagccagg cttttttttt aatatt                      206

SEQ ID NO: 253      moltype = DNA  length = 176
FEATURE           Location/Qualifiers
source            1..176
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide
SEQUENCE: 253
gcagcctggc caacactgaa agctccaatt aggattgaaa aaaaaacctg agcaaaataa  60
catgaggaaa agtttaagt acatcttgc accattat atctaaaatg tgatgtat 120
gttaacatagt acataaaggaa ttaagtttag gtttatcctt tcctttttt ttttgt 176

SEQ ID NO: 254      moltype = DNA  length = 178
FEATURE           Location/Qualifiers
source            1..178
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
       polynucleotide
SEQUENCE: 254
tgctttccc acagccactg tttccaaatca gggccaccgt ctgccactc tgcacctca  60
ggttcagacc cttcaagatc taccaggacg agtgagaaaa aaacttcaag gcaattcaca 120

```

-continued

---

```

gacacaggat ataggaactg actgttcaact aggtttaaat atacatgcac ttttttat    178
SEQ ID NO: 255      moltype = DNA  length = 194
FEATURE
source
1..194
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide
SEQUENCE: 255
aagtctacat acctcggttc ataagctccg atggaaatag ctgcaaggc aaaaggc 60
aatcttttc ctatcttc aaaaggcagg caataaccg tatctctgc gaaaaaaaaat 120
cgattccaag gccccaagac agaccagctg cccatagca cttgttgc atccattaga 180
gtcctttac acca                                         194

SEQ ID NO: 256      moltype = DNA  length = 227
FEATURE
source
1..227
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide
SEQUENCE: 256
ggggccagac ccattggcatt aggcttttc tgcttcccg gggctggaga ataggac 60
tttcctatct tcacctctcc cactggagg gaacaatctt cccctgagt ctggggctg 120
gagcgggcta gagttctgtc ttgtggga gcccggcgc gggggggccgg gcacagtctc 180
ccagcatcag ccccggtggc ttctcagag actgctggcg gtttctt 227

SEQ ID NO: 257      moltype = DNA  length = 224
FEATURE
source
1..224
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide
SEQUENCE: 257
ccatctgtgg cacattgttg tctccataaa aaaaaatctt acactttcc aatagcatta 60
agtttgc 60
tttgcattc ttcataggg ccgatgtgtc gtgcccacaa gcgattggac 120
ctgtggagg tctgttgtgg gtgggtcccg gaactgtggg cggggcggtt taccaggc 180
tccctctgt ttgtgtctt tagatgtcg tactctcagg aggc 224

SEQ ID NO: 258      moltype = DNA  length = 206
FEATURE
source
1..206
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide
SEQUENCE: 258
ggtaaatttc caagtgtatct acaaacttgc aagtctttagt ggaataatca ttgacatgtc 60
tctttttt ccccaactt aacttgc 60
aagaatcttgc tgatttactt ggagaatggg tgctatctgg 120
caaaaaggta agaattttt gtccttgc gatgaaatac tagtcatttta attatactta 180
acatacagca gtaatcactt atatat 206

SEQ ID NO: 259      moltype = DNA  length = 225
FEATURE
source
1..225
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide
SEQUENCE: 259
ggggagttct gagtacccgc aggcccttctt cccgctgtttt cctgaagctc cgtcaatac 60
attccaaatctt tttctgtatctt actttctgtc ctgtgtccaa cagcattgtc acttcttcat 120
ctgaatctgt ttctgtatctt aaggaaatctt cagttatctt ttcagaatgt tcaagaatgt 180
atgaaacata aaaaaatccc ttcaatggaa agaaaatgg aatat 225

SEQ ID NO: 260      moltype = DNA  length = 201
FEATURE
source
1..201
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide
SEQUENCE: 260
tcttcttcc atccatatctt gcatgaattt tccgagccaa gaatccactt ttgtacacag 60

```

---

-continued

---

```

cagcattgg atcatgagga atgtccaaaa atgggttagt agtactcca atacgactga 120
tggctttgg atgtgtccg ttatgtttct cctcagttact ttctgaggaa gactttttt 180
tcttttcata atctctaaaaa g 201

SEQ ID NO: 261      moltype = DNA length = 186
FEATURE
source
1..186
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 261
agaagttata ttcttataaa ccatgtgata agatttcttc tgggagtaac atttctaaaa 60
aaaggatcac agttccatct ttctatgttc tatacttgct ttatgatc tttttttct 120
aaagagaaaag aactgtcaga tggggctca ttccatggc aaaaggaaat taaatttaaa 180
acataaa 186

SEQ ID NO: 262      moltype = DNA length = 221
FEATURE
source
1..221
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 262
tttgaggagc ttccaaaatt tcatttcaa tgggggtgtc aaatgtcttc agaataagcc 60
acaaatgc aaataatctc cgaaaccaat ctgtggcaat ggcagattgg agggaaatga 120
aatctgtat tgggtactg aggctgttactg tttttttttt attaataat 180
ttatgtttt aatttatctt ttccatataat gtcatatatac t 221

SEQ ID NO: 263      moltype = DNA length = 197
FEATURE
source
1..197
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 263
atgatagtct tctttccaa ttattagacc acttaggctt atctttggta ttttgttct 60
tattgatttt agggtctta tatttgcata atttgaaatg catggtaaaa aaatgtatct 120
actgtccatt ttttttaggt tgatacccaa agatcgtat tactgtggg tttttttttt 180
cactgggagt gatattt 197

SEQ ID NO: 264      moltype = DNA length = 214
FEATURE
source
1..214
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 264
caccataatg cataattcaa tcagaaaata tgaagagctt cttacggagt cccacagaat 60
ctgggttgc attatcatc ttacttgggtt qcacaagtgg agcaaatgaa acagcaagga 120
tattttact ttcccaagtg ttctgtccag ttgcataat ctgtgttaac tatttagaga 180
aaaaaaaaaa agaaatgtaa caacttaatc ccaa 214

SEQ ID NO: 265      moltype = DNA length = 217
FEATURE
source
1..217
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 265
ttggggccgg acgttactcg gtcggacta cagcgatgaa aatgactata ctcacacatc 60
ctataatgca atgcacgtga agcatacgtc ctgcataatc gagagagaaat gctcacccaa 120
gctgtatgaa aggttaaggcc tgggttatca gagccccccc cccagtcaga ggtcaatgg 180
attttttttattt tggacattttt gtgaattttt aattttttt 217

SEQ ID NO: 266      moltype = DNA length = 218
FEATURE
source
1..218
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

```

-continued

---

-continued

---

```

SEQUENCE: 272          polynucleotide
gccagggtgg agatcgcccc ccaggtctg aatctgaaac caggagatgg gtgcagagt 60
ggcacagggg agcagggggc agagcgggtt gggggggggg tggagggtga gggaggggg 120
tggaaagggtt aggccggggg agggccacct ggaactgcac cgtctctgc atggcacca 180
ggtgacagg gatgtggggg gcattggaca ccagcccc 218

SEQ ID NO: 273          moltype = DNA length = 228
FEATURE           Location/Qualifiers
source            1..228
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 273
caactgtcac tcccactgga ggcgtgtgt atgcagggcc cgccagtaag cccagctgcc 60
tacagccaag gcccagcgac agacatccctc ctccgttagc agcagcagcc tgcctcctc 120
cagccggcgc aggaggcctg ccacacccag catgcggccc ccccacatcc cataagtgtt 180
ggccgtgaag acagctacaa tgaggatgtt cactgcgacc accaacag 228

SEQ ID NO: 274          moltype = DNA length = 221
FEATURE           Location/Qualifiers
source            1..221
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 274
catgcccccc ccaccaccca aggtgagcag ggtgccagca gcctgtttcc tcttcctcg 60
ctggccatcg gggtttctgc tcccgagttt ggactgtggta atgggtggg tgcgtgtgtc 120
tgcgtccatcg ctggctctc cctgggtggg accttgcgtt ccaagactg gtgagccctc 180
cccaagtgtac ctggccacc cacttgtacc tgagctgcc c 221

SEQ ID NO: 275          moltype = DNA length = 208
FEATURE           Location/Qualifiers
source            1..208
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 275
ccaccatact gtgggatgag accacggtag tcacctggc ggaactccct tgagtatcaa 60
agtaattccccc tccagtattt tggctgtaca tctgtgtctc tgcgtaaaggta tacgttgg 120
ttcggcttta aaagaagaag aaaaagaaaa ttacttagct tatttatgtc atcttattaa 180
ttttttctta atatgaatgtt aattttaaa 208

SEQ ID NO: 276          moltype = DNA length = 225
FEATURE           Location/Qualifiers
source            1..225
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 276
agacaacaga aagttagggg tgctttcaat ttattattca acctgagctc tcctgtgact 60
cttggttctt ttggaaacat gttaagttt gaaagcttac tttctggaaa gcttcagaaaa 120
caagtgtatgt tgactggaca acatgaggag gttcctcaag ctctaaacctt acgtgcattt 180
tttttcaag atcagattgtt gtggagattc tttttttttt tttttttttt 225

SEQ ID NO: 277          moltype = DNA length = 181
FEATURE           Location/Qualifiers
source            1..181
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 277
tccacaagtc aatactatca gcaaatttca acaagctttt gtgtatgttt ctatgttttac 60
cttccaaacta tatatactta aaaaaaaaaa caatcgctca tatggatataa aagtttagga 120
gtttaggtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 180
a 181

SEQ ID NO: 278          moltype = DNA length = 215
FEATURE           Location/Qualifiers
source            1..215
mol_type = other DNA

```

---

-continued

---

```

organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 278
cagaagcgta gaggtcaccc aggctgtttt ggggggatg tccccccacc gaaacccgt 60
ccgctgcctg ctgtggagac gcccacatgt caaactcacc aggcaccatc cctgtcageg 120
tcggtgctga gtatgtccc tgtccagcg gggggacatg tgggggtac ccggggagg 180
tcgtgctcgc caagtcacgg cctgagaaat caaaag 215

SEQ ID NO: 279      moltype = DNA length = 221
FEATURE
source
1..221
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 279
acatcactca actcatattt acagataaaa aaaaagtctg atagatattt tcccttata 60
cccaacaaat tattgatttgc tgctatttgc atataaaagg ataaatattaa gttatcttt 120
tttaactaacat ctgaatgtat atgaacaaat taaaagaat gttatctcta cttatagatc 180
tcgatttattt aatgatgtcat accaggaatg actggatctc c 221

SEQ ID NO: 280      moltype = DNA length = 223
FEATURE
source
1..223
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 280
aatgtccaat ttctgtcaag cttaccttag agaagatggg tttgctgctt ggtggctgtt 60
gagaagagca cttttctgtt attaaggctt cactcctgag ctcaagatcc tctgcagcca 120
ttaaatagca gttctgtctgtt attacttca aatcttttc attgaatttc taatgaaata 180
aaacaaatcc taagggaaattt gtcatcaaga aaaaaatataat taa 223

SEQ ID NO: 281      moltype = DNA length = 219
FEATURE
source
1..219
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 281
agcatttgaa tttagaatga catctttggg ctctgaatca atggcattta tacaggagcc 60
tatgggtcoca caagaccagg gagaataatg gggaaatgtatc tcttcttcaa attttgtacc 120
actcacatc tctgagaact ggtttaaaaaa aaataaaataa agccaatttag taagaagctt 180
aaatttacaa gtttgcactt aatataaagt aacttattat 219

SEQ ID NO: 282      moltype = DNA length = 202
FEATURE
source
1..202
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 282
catcgtgtc agcatgggg ctgctaaagcg tgggtgcaca ctccctcccc ggggtgggg 60
ggggggcaggg tggggcccgag gctgactcg ccccccacc accgtgtcg actgtccggg 120
tgtctgtcc catcatcatg ggtgtcaacg taggcacatc catcaccaggc accctggct 180
caatggcgcgac gtcaggggac cg 202

SEQ ID NO: 283      moltype = DNA length = 192
FEATURE
source
1..192
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 283
gagctgatct atcaatctt gatccat aagattgtat tttatccaa catttaaaat 60
tactctacag ttttcttca gtttgcgtat tcaatgttta aatggccatc atccctgtt 120
ttttccctaa tctttttttt aaatggattt agcatatctt actggaaatc tgagtccatt 180
agtgacttca tt 192

SEQ ID NO: 284      moltype = DNA length = 222
FEATURE

```

---

-continued

---

```

source          1..222
               mol_type = other DNA
               organism = synthetic construct
               note = Description of Artificial Sequence: Synthetic
                      polynucleotide

SEQUENCE: 284
ctacctgttt tggtagacatc tcatccaaa gcttggaaatt gtctttttt ttctttccca 60
aaagtaccag aataggagtt ttctctgtg gacctgaac ctggctgaa accctgagta 120
aacaaggcat ctccactct gagtctggcc ctcggggagt gcatttcatt ttcaacaagg 180
aaaacttctc acttgtctc tccatgagga aataaatgtg gg 222

SEQ ID NO: 285      moltype = DNA  length = 193
FEATURE          Location/Qualifiers
source           1..193
               mol_type = other DNA
               organism = synthetic construct
               note = Description of Artificial Sequence: Synthetic
                      polynucleotide

SEQUENCE: 285
tccagcccg acaccacaga gcctggctac atcaacatca tggagttggc agcatctgtt 60
tgccgcgtatg accttagatga catggacatc ttctggcttc aggaactcaa tgaagacctt 120
gcagaaaatgg gtaagttctt ttctcacaag taagttgtga cttattcat gtttttttt 180
taagttatca gaa 193

SEQ ID NO: 286      moltype = DNA  length = 184
FEATURE          Location/Qualifiers
source           1..184
               mol_type = other DNA
               organism = synthetic construct
               note = Description of Artificial Sequence: Synthetic
                      polynucleotide

SEQUENCE: 286
ctgttgacta acccaactgct tctgttatctg aatgctgtct ccaggtcaga gcccgggctt 60
cgcccttcgtc gagtttagtc acttgacgga cgctacacga tggatggaa ccaatcagg 120
tgcttgcgcg cacttgaacc cccccccaaa caaataactac tttgtaatcg agcgtcccc 180
atcg 184

SEQ ID NO: 287      moltype = DNA  length = 174
FEATURE          Location/Qualifiers
source           1..174
               mol_type = other DNA
               organism = synthetic construct
               note = Description of Artificial Sequence: Synthetic
                      polynucleotide

SEQUENCE: 287
tgcttaaatgt ggacagcaact agtggagtg acttatacaga cataccctc atctgcagga 60
agaaaggatt aatggtttag ttttattaaac atttgaaaat aggcaatata agtggaaagtt 120
tttttttcc tactagcaaa cttattaata atttacttt gatatggcaa atta 174

SEQ ID NO: 288      moltype = DNA  length = 217
FEATURE          Location/Qualifiers
source           1..217
               mol_type = other DNA
               organism = synthetic construct
               note = Description of Artificial Sequence: Synthetic
                      polynucleotide

SEQUENCE: 288
gcactgccta tcctcatccc ccccatgagc cttggttttt atgcaaccta tggtcatgca 60
gggatgcct tacaccctcc cacatgtca tgactggcag gcagctgacc agttccatca 120
tagcagcgcg ctccgaagca cctggcccca ccctcaggta agagtctgtat agtggaaagaa 180
agacactgtgg agaaaggatggc cacaggatct gtgtgtc 217

SEQ ID NO: 289      moltype = DNA  length = 193
FEATURE          Location/Qualifiers
source           1..193
               mol_type = other DNA
               organism = synthetic construct
               note = Description of Artificial Sequence: Synthetic
                      polynucleotide

SEQUENCE: 289
ctcccacactg tccccatctg gaaagctcg acgcaaggta gaggccaaaa aggtgacacc 60
cctgatgcgc acaggcccccc ccgcacagca gcctgtggc cagccaggcc ccccgaaacgt 120
atgctctgtg cgcactcaca cctgtgcaca cacaatctg cacacaagca tacctctgt 180
cacacccctccg cac 193

SEQ ID NO: 290      moltype = DNA  length = 194

```

-continued

FEATURE source Location/Qualifiers  
1..194  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 290  
cgcgagttac tcggtgacaa aggttcttt ggccgtactt tggggggct gtcttcatca 60  
ctatagggtc tcgaagtaac ttctgtgtac catcttat aatctggctt tactgtgtca 120  
ggaggtatgg gaactctggg gggaaacctca ggtttgtctt catcgaggtt aggagaagct 180  
ctgtgtatac aaca 194

SEQ ID NO: 291 moltype = DNA length = 187  
FEATURE source Location/Qualifiers  
1..187  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 291  
gtctgttatt atgttctggc agaactgtgt atgctgacgc tggactttt ctacaagttc 60  
aggacgcgc cctctctgtg tgagtatgg gaccgccttc tggcagatgc tctaccagca 120  
gcagggggca gttgtgggg ggggcgcctg gagtgtggaa ggaagggggt ttccatggc 180  
tgggtga 187

SEQ ID NO: 292 moltype = DNA length = 174  
FEATURE source Location/Qualifiers  
1..174  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 292  
ttaccttggaa atatcccccc agtaccattc tgcattcttga agagaaaacag aactgttc 60  
cattccattt ggaactgtc aagtcatggg ctttaggtggc ttgggtggaa gagctagaag 120  
agaaatgaat attactctgt agatgttaat tatttactgt ttcttaattt cctc 174

SEQ ID NO: 293 moltype = DNA length = 193  
FEATURE source Location/Qualifiers  
1..193  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 293  
ctcacctgtt ccccacagaa cagataccag ttttcatcg agatgggttc cacaaccact 60  
ggggccgcgc tgatccatgg ggggttcttc ctttgtgtt caacagagcc tggggggctc 120  
atcccatagc ggtaatgggg gggcagagtg accttggatg ccaaggccag cagggcagaa 180  
agacgcatgc ctg 193

SEQ ID NO: 294 moltype = DNA length = 204  
FEATURE source Location/Qualifiers  
1..204  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 294  
ggcattaaat ttaatctgac aaaggccatca agtcgggggg gattaaatgt ttacttactc 60  
atgcagatt ttgttttgtt tgtaaaactt gtgtgtgtg gcaaaaacc gagaataata 120  
cttggcattt tcgatatacat ctttggaaaa gaagagacag agaaaaaaaaat aaacatttg 180  
tgaattagttt atacatttca ttta 204

SEQ ID NO: 295 moltype = DNA length = 210  
FEATURE source Location/Qualifiers  
1..210  
mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

SEQUENCE: 295  
ccatggcaca ttttgttaat acgtgtgtat ttttcccccc cccaggggact actggcattc 60  
agtgtatgtgg tcatagaat ctctccagag gaggggcat gccttgaccc tgcccaaggca 120  
aattttgtata gggatgtgtat ttgcgaaac tacagaaacc tggttcttcc ggttgaggat 180  
aacttcataa cacaatttat ttgttccata 210

-continued

---

```

SEQ ID NO: 296      moltype = DNA length = 203
FEATURE           Location/Qualifiers
source            1..203
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 296
ggtaacttga tgcttcacaca caccaattgt gtgtggccct tggtgccat aaggtacagg 60
gggctggcga gagagaccac tggaaactca acaagattca tcccatgaat tagaacccacc 120
actaactcgcc cgctacgaag gcacatgtaa tagccagtga cgcgtcaccag gcagtgcact 180
tgttagaacc cccccatgg tgg 203

SEQ ID NO: 297      moltype = DNA length = 225
FEATURE           Location/Qualifiers
source            1..225
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 297
tctctgtttc cttgaggacc tctctcacgg agacggacca cagcaagcag aggctgggg 60
ggggaaaagac gaggaaagag gaggaaaaca aaagctgtca cttatggaa atacaagg 120
tagacaaaaa attgatttga tctgaatctt taagaaaaaa aaaataagaag gaaaataaaa 180
agcccgaccc aaaaagcaag tacttggaaag ctggaggcga aaagc 225

SEQ ID NO: 298      moltype = DNA length = 198
FEATURE           Location/Qualifiers
source            1..198
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 298
ggtttctgaa ccaccatggaa aggtggactt aagtgttataa gtgtctttctt gctatagccc 60
agattgttttgc tttttctgtt taaaaccccc cacatttttc tgctgtttaa ggaagacta 120
gtaccctttaa ttgggttaccat tttatgttgc cgcccttcgat gatgggacaa gttacttgc 180
tcactgcagc ggaaggaa 198

SEQ ID NO: 299      moltype = DNA length = 176
FEATURE           Location/Qualifiers
source            1..176
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 299
cttttagata cagaacactt catgttaacat aattttttc taactttcta taagtaattt 60
agtgcaaaaaa taacctttaaa aaactacttc tgacaacaaac acagagatcaa aaatctaagt 120
attcttttc ccccccaagaa ctaaggaaag tagaaactta cctaagatgt aagagt 176

SEQ ID NO: 300      moltype = DNA length = 194
FEATURE           Location/Qualifiers
source            1..194
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 300
catcttttctc tttttttttc tccccagccg gcgatgtatgt gatcttcatttt gaccctgtatg 60
acatcatcac caacatcgag atgattgttgc acggctgttgc ggcgggggtt tgcaaggcc 120
ggtaacgggtt ctteccagcc aactatgtgg agtgcggca gtagggccccc cagccccccc 180
ccggagctgc gccc 194

SEQ ID NO: 301      moltype = DNA length = 201
FEATURE           Location/Qualifiers
source            1..201
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 301
ggaggatcatgc gttttttca agctggtacc tcgacagaaaa ggcctggAAC tcctgttctc 60
ttgtggcaga ggcagccctc agctctgcac ggtcaaaaggc cttgggtgagg tcaatagctc 120
ggacttggttt ctggaaagggg agggggaggc cccccccact ggactcacaa ctgcgttgt 180
tccaaggccag cagcctgtgg 201

```

---

-continued

---

```

SEQ ID NO: 302      moltype = DNA  length = 221
FEATURE
source
1..221
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 302
gtcctgtcag aatgactgtc tccgagctgg ggggggagtc cactccacgt gtgctcgct 60
cacctgacgc gtagtgc tccgaggaga agaacccgaa gacatcgaaa gaaaaggct 120
tcagaaagt ctggtaaaggc gagagacagg cagtcacagg gctttgcatt tacgcttcgc 180
actgttcaga gaatatgtta ctttctacaa ttacataagt t 221

SEQ ID NO: 303      moltype = DNA  length = 185
FEATURE
source
1..185
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 303
ccaggtacta aaaaagaaaa taacatgaaa acgcccagggt ttacttgaat gttttataa 60
gatagaaata tatgtcttca ccatgggggg ggtctcgatt tcactaacgt tggatataa 120
aatgtctgcataaaaaagta cttttaact ttgtaaacct cttttgtac aagtacttat 180
tttgc 185

SEQ ID NO: 304      moltype = DNA  length = 204
FEATURE
source
1..204
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 304
cccgagcccc cccccagaac tgctcgccgaa gtaactgc aa gatctccat cgggatggg 60
cgggggccccc tgtgcagatg ccaggagggg tcgagctggaa cactccgtgc tggacgcag 120
gccagtgccg cctacgtggaa gactgtccac acagctgcct cagaagcgcc gggcttcct 180
cacacctgtag cggggtgccgaa ggag 204

SEQ ID NO: 305      moltype = DNA  length = 192
FEATURE
source
1..192
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 305
attcgggtgg agcacgttgg cggagttggat gtacttgggg ccccgccagga tctggtagag 60
gaagtagcag atatggtcat tgctcagctg ctggcttttca agcaacttgtt acaggctcgt 120
ctccatcagg ttctgcacaa tggatgttgg gatggctccg cagacccccc caggcagggg 180
gcagtgggag gc 192

SEQ ID NO: 306      moltype = DNA  length = 195
FEATURE
source
1..195
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 306
aacagttctg tgactgtcaa tctttgtttt ctttctgtgg cttttccccccc ccccttcaat 60
caatgccttt tctttttctt aggcaagact ggtgtttttt tactttgtata ccaagaggta 120
ccaggaaacca ttgcatttgg gtaagtaacg tggataataag tcacatctcagc tagctaaatg 180
aatgtattctt aaacc 195

SEQ ID NO: 307      moltype = DNA  length = 188
FEATURE
source
1..188
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 307
gagacagggt gacgtgtcag agacaatcg agtcttctttt gaggcaggca agtctttccc 60
cccaagctcc aagaggctcc ttaccatcca ggaagtggat gagttcttc tggggctgtc 120

```

---

-continued

---

```

caagctcacc aaggaggatg agcagcaaca ggcctacag gacattgcct ccaggtgggg 180
gagctgcc 188

SEQ ID NO: 308 moltype = DNA length = 187
FEATURE Location/Qualifiers
source 1..187
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 308
tttcggcgtt ccagcctcag gttggttca tcctggagcc tggagtgggg ggacacattc 60
ctgggtccag gctttctga ggacacttga atacctgcct cagtctccct tgaggcagct 120
gtcacagaga ccaccccccac ccaggaccag tagagcagca ggaggattaa gggttgggg 180
gccttag 187

SEQ ID NO: 309 moltype = DNA length = 209
FEATURE Location/Qualifiers
source 1..209
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 309
cgcgcagtgc atgttaagg aactcggctc gagcccatag cggccgcgt agacgcccgc 60
cgcgctctgg cccgccccatgc ccccccgcg ggggtacaaag ccctgcacatcg aggccgcgaa 120
gaaaggccccc gaaacccgcctc catagcccg ggcgtgggg ttggaaagcaa acgcacaaga 180
agtttgttct gggaaaggctc cggtagcga 209

SEQ ID NO: 310 moltype = DNA length = 190
FEATURE Location/Qualifiers
source 1..190
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 310
agcctgggtgc tccttcttgg gggggctgac atgactaagt cgggtccaca gcctttcc 60
cacactcacc cattccctgg ccacccctgc ctatgcacg acccatgtca cctgtgggtc 120
cctgtgggca ctgcgtcgtg gccggcccg aggggagagg ccattgtatc tgcatcgta 180
gcaggctca 190

SEQ ID NO: 311 moltype = DNA length = 231
FEATURE Location/Qualifiers
source 1..231
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 311
gttagcctgc accaatgctc actagggtt cactagggtc ccagagtgg cctcagagcc 60
agaacacctaa ggtatggaaa gtcttacc agggggccag gaggggccat ctgtccagg 120
gtcccccagga ggcctgggtc acccttgaaaa tggAACAGAA catgggggtgg gtcacccct 180
gtcagaacac acgtgtggacc cccccccgac catgcaccaag gggcagcccc t 231

SEQ ID NO: 312 moltype = DNA length = 221
FEATURE Location/Qualifiers
source 1..221
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 312
cttttgcatt tgctgtgaga actcacccag aacttgacca ggaagaaggc atgggggggg 60
ccacgatcat atagctctcg gagggccacc tttttctcaag ggaattttgc gtagatctgc 120
cggaacgtca cactctcgag cggccgcgt ccggggctgg ggcagtgcgt gctgtatgtc 180
acgaacacgtt ggcctctggta ctgaggaggg ttggccgtgg 9 221

SEQ ID NO: 313 moltype = DNA length = 201
FEATURE Location/Qualifiers
source 1..201
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 313

```

---

-continued

---

```

tttctgtggc ccctccctt cctagtgacg tctctcagct ctaccctctc aaaccgtcat 60
atgacgtcac tatggcgta tggtgatggc cacctaccc gatgacgtaa tcatggccg 120
tgaggtaaa ccagaccccc ccaagaagga aggagactgc ggggagctt gggattccg 180
agcacaggat gatgtactgg g 201

SEQ ID NO: 314      moltype = DNA  length = 181
FEATURE
source
1..181
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 314
agtgggtgg tcttgtcccc cccatccag gatcctggtg aacatggacg acaaactgt 60
gaagcattac tccaatggg acaccttcca gctgcaggat gaagaagccg gggggctta 120
caagctacc ctgacggaga tctaaaggcc tgcggccac agtcggccag gagttcagtg 180
c 181

SEQ ID NO: 315      moltype = DNA  length = 203
FEATURE
source
1..203
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 315
gacagcgtgt cttgccttag gtatcatata aagttgtga atatgtaaaat ctaagtaaca 60
agtagctcca aggttagaa ggccaacaaa tctacattca gcacggacat cttcatgagg 120
ccagtaatcc catttataag gtgcattggc tgcgtgaagaa agagggagga aaagatccc 180
ccaaacaaaa ctatgaaaca caa 203

SEQ ID NO: 316      moltype = DNA  length = 175
FEATURE
source
1..175
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 316
caaagaatgc aagagtcaag atgcggagc aggaatatga attccctgag ccgcgaaatc 60
tcaacaacaa tatgcgtcg cctcttccc cccggaaatg gtactctcca atcaagggt 120
gtctctttac tcaaaaaatc tatcatcagt cctgataaaa cacttacata gtgaa 175

SEQ ID NO: 317      moltype = DNA  length = 197
FEATURE
source
1..197
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 317
tcttcgtggc actggatgcc gcgtagacag ggctggcac acagctcacc actgcaggc 60
cccccccaagg qcacctggcg gtgtcattgg cacgatgaaa ctccaggaa cagagggaga 120
ggtgggcctc cgtgcccacg cacgcaccc catgcagacc aaaggagtg tgctggcg 180
gggcttagcag cctaggg 197

SEQ ID NO: 318      moltype = DNA  length = 213
FEATURE
source
1..213
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 318
ctttaccaaa gcggattggc attatgtta agttcttaa ttacagatca agaaaatgca 60
tacagaagat gggggggggg cacacctaatt taatttttat atttagatta aagaaaataa 120
ttaaatgtgt tttttgtgg gattgatttt cagaagctaa atgcaactag ttcatctgaa 180
ggcagcacgg ttgatattgg agctcccgcc gag 213

SEQ ID NO: 319      moltype = DNA  length = 188
FEATURE
source
1..188
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

```

---

-continued

---

```

SEQUENCE: 319
tgtccctcg ctggaggcca gaggccaaga ttagtgcatt gctggggcgc ttgcagagg 60
aaggggggca gtggacccc ccaggtggac aaagcgcttc atctgtggct attctgagag 120
acggacagga gtgggttttgc ttgtcttc gaaggaatgt acttggattt ggaggctgtt 180
cctactca 188

SEQ ID NO: 320      moltype = DNA  length = 187
FEATURE
source
1..187
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 320
tgggtgggtc ttggcaagtt cccttgagc caacaagggtt accacttag gctttctct 60
tctccagggg tttccagta agatgctgc ttgtggca gcacctggg cggcagagc 120
tcggcatggg tgactccacc agccagtcgc ccccaattaa gaggtcatgc ccagatgtgc 180
aggtgag 187

SEQ ID NO: 321      moltype = DNA  length = 172
FEATURE
source
1..172
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 321
gtatgaccggg gtcatttcctc cttggaccaa acccccctag attcaagtgc taaaatct 60
tctgttagtg tcaccagccc ctctggagtc tcctccctcta catctggagg agtacccctc 120
acatccaata tgcattgggtc actgttacaa gagaaggacc ggatttgca ca 172

SEQ ID NO: 322      moltype = DNA  length = 228
FEATURE
source
1..228
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 322
gagggtgggg ggggcctacc tggcagcaac ggtgtacccc acaccccat ccctctagct 60
cggtgtgcgt gtacctggta ggggtctgtt tgtagatcaa agtactctag gaaggcgtt 120
gcaaattcac agaagaggaa attgtgagtc tcattgtatgg tcctcatgca ccagtacgt 180
ttattgttgg cgctggtgca ggcacagaaa ggcccccactg ccaaggaa 228

SEQ ID NO: 323      moltype = DNA  length = 204
FEATURE
source
1..204
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 323
cgtgccacgt gccaaggact gttctgggc tggggaaacc acgaaggcatg ggaggcgggg 60
actgcggaca agtcccagcc caagcgtgg gggggggctca cgcactgaa gatgtggccg 120
tgggggtgca ccacgggggtt gatgtcaagt ggcagcttct cttgttttc ctccagaatg 180
gcctaaaagg aaacacacccg gaag 204

SEQ ID NO: 324      moltype = DNA  length = 233
FEATURE
source
1..233
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 324
cagccccctgc cacccacaga cacgggctga ctgggggtgtc tgccccctt gggggggggc 60
agoacacagggc ctcaggctg ggtgcacactt ggcacctaga agatgcctgt gcccgggttc 120
ttgctgtctt tggcactggg ccgaagccca gtggccctt ctctggagag gcttgggggg 180
cctcaggacgc ctacccactg ctctccgggtc agtctggaaac cctggggaga cga 233

SEQ ID NO: 325      moltype = DNA  length = 217
FEATURE
source
1..217
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic

```

---

-continued

---

```

SEQUENCE: 325          polynucleotide
caggtatccc aacagggtgct ggctcttgg ggggagaggat gatgcctaga cctggctcaa 60
taccagcacc aacgacctct ctccctcaac tggacactca cctgattccc cagagggta 120
cacctctcca ggagcatctg cacagaaagg aaatcagtga tcataccatcc atgatcc 180
ggggggggag aagtcatgct tcaggaaagg tggatcc                           217

SEQ ID NO: 326          moltype = DNA length = 225
FEATURE
source
1..225
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 326
cagaagatgg aagagaacag tttaagtttg qaatgcctgc tgatgcctaaa taccactcc 60
tttttcacta gtatgaattc agcgtatct ttgatacaag ccacctgtct ggggaaggagg 120
aagttctcgat cttcattgtt actgtcaga ggtaaagggg gggtttaaac actttttca 180
aagatgagag atttatggga ccaggaaagg atgtgtcttg gttaa                           225

SEQ ID NO: 327          moltype = DNA length = 194
FEATURE
source
1..194
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 327
atttgaacct tgaatctctg cagaataact ctattcccc cgtatggtaa agcagatcat 60
atttaaaatc aacattnaaa aacatctata tcctaattttt atttttaaa cataaaacat 120
caggctactt gcctttggc aatttctgtg catttccag ggctagtgtt atacttccac 180
agagaagtag acac                           194

SEQ ID NO: 328          moltype = DNA length = 215
FEATURE
source
1..215
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 328
atatgtttca ttttctgaca tacattgttt ctaatgtgct ttgtatcttg agttcaatg 60
agtgttcccccc ccggcatatt gctgtatgaa tttagatgat tgccctaaata aatataaaa 120
tatgtggggg ggggtgttgtt ttattattgtt ttatattttc tattttagat gcagcagg 180
ttaggaagac ctcttctgtgaa agaatcttcc tttag                           215

SEQ ID NO: 329          moltype = DNA length = 198
FEATURE
source
1..198
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 329
atttattttat ttgttaggtt ttcaagactgt ttccctttaac agccagtttag gatatcttagat 60
attaactta ccaggctca tgccacatcc atcatccagg aaacacaaca tgaatcccc 120
ctgcagttt tcattatcca ctgtaaagaga aagcagttat tactaataaa tattcggctt 180
aatttgttagt tcagaaag                           198

SEQ ID NO: 330          moltype = DNA length = 193
FEATURE
source
1..193
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 330
ggccacaaga ccaaagatac ttacatcccg attgccaggta aagagcagaa gcactacatc 60
ttctctgcagg gggggcacga gccccacaaa agaaagtaag gagcattttt ttgacctta 120
cagaaattcc atagaattct ttgaggagat tagtgtatgag gcttccaaat tagtggatag 180
gctgacacag tca                           193

SEQ ID NO: 331          moltype = DNA length = 182
FEATURE
source
1..182
mol_type = other DNA

```

-continued

organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic  
 polynucleotide

**SEQUENCE: 331**  
 atttgttct tacagtggc gggatgccta ctgggtggag aaccttatgg tccaggagct 60  
 ggagaacttc aatccccctc tcaagtttg tcttcataag cgggacttca ttccctggcaa 120  
 gtgtatcatt gacaatatca ttgactccat tgaaaagagc cacaaaactg tctttgtct 180  
 tt 182

**SEQ ID NO: 332** moltype = DNA length = 229  
**FEATURE** Location/Qualifiers  
**source** 1..229  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic  
 polynucleotide

**SEQUENCE: 332**  
 cactgttgg ccgcgtgacc ccccccagg agtagagcac ggcacacgca caacttactt 60  
 cttccacagg gccgggggct gctcggttc ctctccca gagcccaagg atgcgacaag 120  
 tgcgaggggc caggccagca gcatacgcc agcgcaacc aggagccgga ccggaaagag 180  
 cgtcagtgcc atgaggccca cctgcagaca gagggggca ttatccaga 229

**SEQ ID NO: 333** moltype = DNA length = 190  
**FEATURE** Location/Qualifiers  
**source** 1..190  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic  
 polynucleotide

**SEQUENCE: 333**  
 ctcttcgtc cagagccctc gcattggaaa cagctgcata gggcccaacg acaagagggg 60  
 gaggtagggg accctgttca cgatggaccc ggtgtgtacc ccccaatgg ggactgtcat 120  
 ggtatgtctt aaggggagaca atcgcttagt gtaatttagt ccattttccgg ttggagcttc 180  
 tgcccaagtgg 190

**SEQ ID NO: 334** moltype = DNA length = 203  
**FEATURE** Location/Qualifiers  
**source** 1..203  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic  
 polynucleotide

**SEQUENCE: 334**  
 ccctccggc ccccccac acggccccac cgctgttgc gctcgagct tctccacc 60  
 acatccataac catggctgtg ttctcgacg tgctaccgc gctgtctcg agggcccaag 120  
 ggaaccttgg gggtaagcga tccctggag agttgtata gacgcagagg ggctgaagca 180  
 agataaggcc cgcctagtag ggt 203

**SEQ ID NO: 335** moltype = DNA length = 214  
**FEATURE** Location/Qualifiers  
**source** 1..214  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic  
 polynucleotide

**SEQUENCE: 335**  
 ggagaaaaag cgggaaattag aaaacgggtt acaagatgtc agcggacagc tcaatttctac 60  
 taaaaagccc cccaaagaaag gtgagttat actttcatgc cactacatg tgactccatc 120  
 ctgccttctt gactgtctt tattgacaaa tgaagattca gacttgaaacg tcttaactt 180  
 tctgatattgt tctgcagcga atgagaaaaac agag 214

**SEQ ID NO: 336** moltype = DNA length = 213  
**FEATURE** Location/Qualifiers  
**source** 1..213  
 mol\_type = other DNA  
 organism = synthetic construct  
 note = Description of Artificial Sequence: Synthetic  
 polynucleotide

**SEQUENCE: 336**  
 caattgaagt ttatgttca aacccttggt aaactcggtc ttttccccccc cccaggagag 60  
 tggaaatcgatc gacctgatc cagggttctgt cattgtatgc acaatgttca atccttgttc 120  
 gtatcgatc aatggaaatgt aatcggtatgt gatgtatgtt atatgttca aataatttt 180  
 taaaatgtgaa tagtctttcc attcttaaca ctg 213

**SEQ ID NO: 337** moltype = DNA length = 206  
**FEATURE** Location/Qualifiers

---

-continued

---

```

source          1..206
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 337
aaaaataact gattacctgg gctttctgg tacatcagaaa ggattactgc aaaatggtc 60
cttgtaaata gtctcagaca ggtcatgatc caacaaagt gccataattt cttccctact 120
tggggggggac ataagtggtt tcagaatgtt agcagttcga ggggtgaagc tgccatccc 180
agattgtgaa ggagagctag tagatc 206

SEQ ID NO: 338      moltype = DNA length = 179
FEATURE           Location/Qualifiers
source            1..179
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 338
gtgactgtttt gaaccctttt ggccacatt caccttggga gcacccctagg tttttttttt 60
tctggcctga aaggcagcaa ctgtgaactc agtgtgtaat cttaatttc tagctgggt 120
gatccccccc aattcagtgc ttcattttga tgtacttctg atggatattt ggaattctg 179

SEQ ID NO: 339      moltype = DNA length = 182
FEATURE           Location/Qualifiers
source            1..182
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 339
aacacagtgg tggccaatgc ctcccttacc tacctcaccc ccctgacggc ctccaaatc 60
gttagccatct gggggaggat gttagctac accatcgaa aggaccaggc acacagctaa 120
ctaacttattt cagctttggg ccgatggggg tggggggggg tcttcagtg aatattttt 180
tt 182

SEQ ID NO: 340      moltype = DNA length = 204
FEATURE           Location/Qualifiers
source            1..204
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 340
gaatctctcc cccccccctc ttttttcccttccca gacatTTTaaat gtggatgcac 60
catgccagac cgtggaaagat ttaacgaatg ggggtgtat ggcccaagggtt cttcaaaaga 120
tgttggaaat aattgtttat ctttatgtgtt attcttatgc atggatattt gtttaggttg 180
aaacatggat taaaaaaatgtt 204

SEQ ID NO: 341      moltype = DNA length = 229
FEATURE           Location/Qualifiers
source            1..229
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 341
gttagtcattt gtttttccccc cccccccggaa gctcagaaga agcaatttggaa agagagtctt 60
gtgtgtatcc aagaggaaatg tgggtcgatc atggcagatc actctacaga ggagccctgt 120
gaaggaggggc cagatgccga tggagaagatc agacagatc ggatagaggaa ggacttcgtat 180
gaagatggggt gtcatgtatc ggttggaaat aaacatggat tattcacac 229

SEQ ID NO: 342      moltype = DNA length = 218
FEATURE           Location/Qualifiers
source            1..218
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 342
aagcttccgc cctcttgcacc ttcttccttc caacctgatc cccactccaa cccctgcagg 60
ctgtcttggg gggcaggacc ctcttccttc agcacaaggggg cccctggatg gcctgcagg 120
aggctccctt gacatgtatc gagggtgatc accgggtgggg ggggggtcagg ggggtccctgc 180
tgggggttggat ctgacccgttgc gttttggggc acggctgt 218

SEQ ID NO: 343      moltype = DNA length = 213

```

---

-continued

---

```

FEATURE          Location/Qualifiers
source           1..213
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 343
gtctgggtt ggagaaagt ttcatttc tggctacttc cccactatgt ggaccattc  60
atgaaactcc attgcatagt tctcatgtgg gggggcagaa ggggtgccca agcctgccc 120
tctctgcctc caaaaatat ggcttgaa aggagttccc aaaaatgaag aggaccctct 180
gaatcagcag gcctaccct cacaggatg aga                           213

SEQ ID NO: 344      moltype = DNA length = 175
FEATURE          Location/Qualifiers
source           1..175
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 344
cgagtcttct gattctaaa ctccaaactg tgtataaaagg tggggggaga atcactcata  60
atctcatcg ccagagataa ccaatgtaca aaatccaaact atggaaagaa agggaaacta 120
gacccagggc tcagggaccc agagttactg atttcagtc acaaattgtc tcatac 175

SEQ ID NO: 345      moltype = DNA length = 221
FEATURE          Location/Qualifiers
source           1..221
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 345
caaaaaaact ggtatactgt ttgtgttatt taggcatatt ttgatccaa tatttacact  60
atctaaatat tggggggggg ttgtttcta acactctagt ggcataatgtt gtttgcatt 120
gtcttattga aagtgtatca tcatttttagg aagctaaatgtt ttatgttgc ttgtttttt 180
taggaagctg ttgataactc ttccctgagct ggtggatgt a                           221

SEQ ID NO: 346      moltype = DNA length = 220
FEATURE          Location/Qualifiers
source           1..220
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 346
ccagccggga atttcgtgcg ggtgtgaag gagctgcggg agccggagaa gaatgaaact  60
gcgtggagtc agcctggctc cggcgttctt cttaactggcc ctgagtcgtt gggggcagcc 120
tgcagaggct ggggtaaatc cttctcccccc tccccggcgc ccatcacgc tcttcacgc 180
tgaattacc tttttttgg ggggggggtc cctttatctt                           220

SEQ ID NO: 347      moltype = DNA length = 182
FEATURE          Location/Qualifiers
source           1..182
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 347
cgaaaaactc ccaaagacat ttcaagatc cgctgtctgt gcaagaaaa ggaaaacgta  60
ctgaatatca actgtgagaa caaaggattt acaacagtta gcctgtcca gcccccccg 120
tacgaatct atcagctttt tctcaatgg aaccttctgtt caagactgtt tccaaacgaa 180
tt                           182

SEQ ID NO: 348      moltype = DNA length = 200
FEATURE          Location/Qualifiers
source           1..200
                  mol_type = other DNA
                  organism = synthetic construct
                  note = Description of Artificial Sequence: Synthetic
                         polynucleotide

SEQUENCE: 348
agttgaaatg aaatgtttt gtttcttta tgtttaaac tgttacaggg accatatgtat  60
gtgggtggttc taccaggagg taatctgggc gcacagaattt tatctgaggt aaaaattct 120
actcaattat acctcaataa agctgggggg ggggaaaaac taaagaattt cagcatctgc 180
ttatgttctt gtaatttttgtt                           200

```

-continued

---

```

SEQ ID NO: 349      moltype = DNA length = 199
FEATURE          Location/Qualifiers
source           1..199
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 349
tcccggtctcc gccttgtccc cccccccggc cccgccccctg ccctgegcgg ggcgtggctc 60
cgtctgccc tcagctcggt tcccgctcc agccctggcc cccgcccccc gctctccct 120
ttcccgcccc tgcctgggt cccgccccggc ctgccccggc accgcagaag gcctctggta 180
acttccttgaa cagtttctc 199

SEQ ID NO: 350      moltype = DNA length = 177
FEATURE          Location/Qualifiers
source           1..177
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 350
atgccttaac ttggactta gagatttatgt ttctttccc cccccccatt 60
aatcaggcaa aacacattct tggtcttaag agatttcgat gggcttggtc aagttatcat 120
tccccaggat gaggtaatag aaaattttctt gtattatctt aaaaacctt ttgatgc 177

SEQ ID NO: 351      moltype = DNA length = 205
FEATURE          Location/Qualifiers
source           1..205
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 351
tgctggggct ggcagcaaaag ctgggtaaca cttcccccctcc ctaggaccat tgccccccccc 60
ccacacctgc tcccccacctt ccttccatccca gacccctttc ctagcttgc ctcagagg 120
aggccttggc cgggtatgtt tgctgtcgccg gggggcggag ggttacctca gtcctgggg 180
tggaggaggagg ctctctggca ggcca 205

SEQ ID NO: 352      moltype = DNA length = 211
FEATURE          Location/Qualifiers
source           1..211
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 352
caagacacac agaaacttgt gactgtggga gagaacgcac atggagtgtt ccccaagtat 60
gtggaaactcc cgtccccccc ccggagaagg ggtggggcc tcagagccag agggccggcc 120
acccggggcgc cacagagggc acggcgagct caactcacag catcagctca gcacgcact 180
ccttgcgttcc ctctctggc tgactcagca c 211

SEQ ID NO: 353      moltype = DNA length = 202
FEATURE          Location/Qualifiers
source           1..202
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 353
ggccagegggtt gageccggca ccctggggag gggcgagggc cggggccacac agtgtgacct 60
ccccacacggg ccatgtctga cctggggccag ggctgggggg gggttgggtt ggcaggcagc 120
caggagagcg gggcccaggg agagaccccg ctgtctgcgc agggaggagc tgccttacag 180
ccgcactggc ctccctgggtt ag 202

SEQ ID NO: 354      moltype = DNA length = 195
FEATURE          Location/Qualifiers
source           1..195
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 354
gtgtggccgc ggcgcagcgc ggcgcacac gccccggccg gcccggccggc aggaaggcgc 60
ggcgccccgg cgggaacggg gagggacccc ccccccggc cacaggtgcg agcgcggccg 120
ccccactcac tcactcgatg aagttagctgg cgcctctc cgtgaagccc tcctccacgc 180
cgcgccccggcgtctg 195

```

---

-continued

---

```

SEQ ID NO: 355      moltype = DNA  length = 208
FEATURE
source
1..208
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 355
taccagctgc tacagttag agagtgtct ctcaggcggg ctggtecgaa gagctggct 60
gtctctgtc agcttctgc gaggatacg gggggggggg gggggacgtg tcctgcacat 120
ggtcacaagg cctgaccctt ttctctggcc ctccaggaga ggctgaggag ggggtggag 180
tccggcgctc actagggtcg ctaagggc 208

SEQ ID NO: 356      moltype = DNA  length = 185
FEATURE
source
1..185
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 356
gggggtgggg gggggaaaggaa gagatcattt agggagtgcc gggaggagggg gggggctggg 60
agccccggacg cttgggtccc ggctgggggtt ggggggttga atgggggggg ttcagggttg 120
aggtcgtac ggggagaaga ggtatgggtcc gtgcgcgtt cggatgggtc cggggccgg 180
gtgaa 185

SEQ ID NO: 357      moltype = DNA  length = 207
FEATURE
source
1..207
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 357
gtcctccact gggccaccc cgagttggcc ttggacaggaa cgcgctgcag ctcggccctt 60
gcctctgtct ctcctcgta ctgtcccgcc agcagggtcgc agtcatgcgg gggactgc 120
agtgcgtggg ccaggcggtt cttegcctgg ggaggggggg gggcaccagg aggtggggagg 180
gactccctgtt gccccatttc cttagtt 207

SEQ ID NO: 358      moltype = DNA  length = 191
FEATURE
source
1..191
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 358
ctgcagtgtt ttatgttctt acttgtccct aagatgagca cttacccagg catacgatgt 60
gcctctgtt gggggccgggg ggggggtgtt gtcccaaattt tgatgttggaa gcagcagggg 120
atccccctgtt ggttagagacc aacaggcagg ctctgtggcc tgggtggctt ggtgtaccca 180
ggagacactca g 191

SEQ ID NO: 359      moltype = DNA  length = 215
FEATURE
source
1..215
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 359
ggaggggctt gagctggctt gggctgggtt cctcagccac atgccaacgg ctgccttc 60
ctgctctcca cagcctcagt aacaacgggg tttctgtggc cgggggtcat tgggtgtca 120
ggccctgttgg tgcgtgttgg accctggcag agtgcacat cagggtggggat ctccctcaga 180
ccacggtacc catccccccc cccatcatgc tctct 215

SEQ ID NO: 360      moltype = DNA  length = 195
FEATURE
source
1..195
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 360
gccccatcgcg gggcgccggaa gccccatcgcg gctggccatg gcctccaaact ttaacgacat 60
agtcaaggcag ggctacgtga aaatcccgcc cagggatgtt ggggtggatg gctcgctgg 120

```

---

-continued

---

```

cttgctcctt cccggcgct cgttcgcccc ggctggctgc ctgggggggg ggcaggaga 180
ggtgaccggc ggggg 195

SEQ ID NO: 361 moltype = DNA length = 185
FEATURE Location/Qualifiers
source 1..185
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 361
gaccacccgg ctgagcggagg cgccggagaa ggccgttctt ctggacgagg acgtactt 60
gtgggtggag ctgcgcaca tgcataatcg agatgtgtcc aagtgcgtgc acacggggac 120
cgatccccccc ccccacccgc cactgtggc ctggtagcgg cttgggatc cttggctgtc 180
gcca 185

SEQ ID NO: 362 moltype = DNA length = 220
FEATURE Location/Qualifiers
source 1..220
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 362
atggatggcg tggtgcattgc tatcgcaacc cctgttctac agataaggaa actgtatggcc 60
agagagagac ggaggactct tggtaatatt gggggggacc gcccctgcac ctcggggacg 120
tgacccccc cccaggctga cccgttcttc gctctgtcgt ccccaagggtgg aggaggagct 180
gacgcaccc cagaagaaac taaaaggac agaggacgag 220

SEQ ID NO: 363 moltype = DNA length = 226
FEATURE Location/Qualifiers
source 1..226
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 363
aggcctcagg ccgaccccaag ccagccagcc ggggcttctt gcccctggct gagagtgtgt 60
tggtggggc ggggttctcag ccgtggcact tgcccccata acttcacccca cgcggacacg 120
tgccctatg gcacatgtca ccttcacca cccgcggacca ccagctggag aaggtcatgt 180
gtggacacc ccccccgc ccccttctt tttggaaacaa aaccaa 226

SEQ ID NO: 364 moltype = DNA length = 230
FEATURE Location/Qualifiers
source 1..230
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 364
ccccccccc acttcctgtat ggcagggtgt gcaggagaaaa ttctcagagt ttggcagcga 60
gacaggatg gcaggcgccg aacggcttgcg agctgtgaac cagatgtgtt atgagctgtat 120
cgagtgtgc catacagcag cggccaccaat ggccgagttgg aaggacggac tgaacggggc 180
ctgggcttagt ctgctggagc tcatggcaca acggggccaa ctgctggccg 230

SEQ ID NO: 365 moltype = DNA length = 216
FEATURE Location/Qualifiers
source 1..216
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 365
ctggaggcac accacatgg aagagatcta ggtgggggac ctaagtagtc agcctgttt 60
atcttggaga cacggaaacc ccccccac ccaagagat ctgtgaatgt tcttagggcag 120
agcctgggtg gtgggtgggg catggagacc ctaccggcct ttgtgttagt tcagaaggct 180
cttcacatct tctctccct tggggaccgg gccggc 216

SEQ ID NO: 366 moltype = DNA length = 209
FEATURE Location/Qualifiers
source 1..209
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 366

```

---

-continued

---

```

aagctgtccc accccccccc cttccagtt tgctctgaaa acgagtccga agccgaagct 60
gaccagcaga tggacaactt gtacctgaaa gccttgagg gtttcattgc cgtggtgacc 120
caagatggcg acatgatctt tctgtcagaa aacatcagca agttcatggg acttacacag 180
gtgacacccct cttcttatctc tttcaaaaag 209

SEQ ID NO: 367      moltype = DNA length = 232
FEATURE
source
1..232
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 367
agcacgcccc cttccccaca gaggacccccc ccccccgcctc accgggagac actgggcgtc 60
ctcccgcccc cggccctctg agaccaggcgc tccttaggaac ttcggccgg agggcggggt 120
cctaagccctt cctgagctgg gcgagggtccg cgccggacccg gcagacacccg ccattgcct 180
gcccgggggg cagggtgggtc acagctgcgg cgtctgaacg gaggccgcga gg 232

SEQ ID NO: 368      moltype = DNA length = 224
FEATURE
source
1..224
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 368
taatcagaca aaaggtagag gttgtctaa attaaaaatg tctaataacct aagttttatt 60
atctccccc cccccccctt ttgttttga tccttaggact tcaaggacag caaggaagaa 120
aaggagaatc aggagttcca ggtaaaggggc aggctgcatttccatccctc ggagtgtatgt 180
ccccggggggc ctgccttc agaggggcagg ttgtgttgc attg 224

SEQ ID NO: 369      moltype = DNA length = 226
FEATURE
source
1..226
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 369
gaggagagaa caaggggggg gggggaaaaa cagagaagag cagagtgaag aggttaggag 60
ttcccccgggagaa gagagtgcata ggttagagag ggtgaggaga aaggtgggggaa aaggggggg 120
ccacggggatgggggggc ctgggtcago atcctttgtc gcccaggagg ggcctgggtc 180
tcatgctcggc ggtctactcc agggcgtgtc ctgagctgag ggacgg 226

SEQ ID NO: 370      moltype = DNA length = 222
FEATURE
source
1..222
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 370
ctaagtttc tcttcatgtat ttgttgtgg ttgtatactg agacctctga gggattttct 60
tccttttcgc agagggcagt gatgactgtt cgaaatgtg tgatggagaa gagcgccagg 120
aaggcctctag ctccctagca agacgtgggt caggtaagga cggggggggg gatttctggc 180
ctggggctgc agagtgtcag gaggttgaat ttggagctcg tg 222

SEQ ID NO: 371      moltype = DNA length = 225
FEATURE
source
1..225
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 371
ctgcttttagg ttaattttt ctgaagacag ttccacatcca agtcagccgt gtcacccctt 60
ccaccctcaa gttttggcga catggagttc tggatgttaa atgactccctc ctcttcagg 120
caggacttgc gaaacctccgt gattttgtgg ggagcactag gatcatccgt gaagaggaa 180
aaagttgggg gggggggagac aatgcataca gcaaccaaat aataa 225

SEQ ID NO: 372      moltype = DNA length = 222
FEATURE
source
1..222
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic

```

---

-continued

---

```

SEQUENCE: 372          polynucleotide
cctcccaacta tggggaaaag aatagtcac ctctcgctc cagagtgcgt ggagatctcc 60
tcattcatct tgccgcacc cttggatttgc cccctgtccc ccgcagatc ggcctggaa 120
ttatacggtc tttttttgg tcagggttcc agaccaccc ggcattgacc aaggcccccc 180
ccccctcggt tccctcagaa ccccaaggaat ggggaaggcg gt                      222

SEQ ID NO: 373          moltype = DNA length = 200
FEATURE
source
1..200
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 373
ggtgctgcc aacctgacag tccaggagat ggactggctg gtgctgggg agtgcagat 60
ggccctggag tggggcaggca ggccagggt gcgcacatgg ctggcaggaa 120
gagctgcac ggtgaggccg gggggggcc ttggcgcggc ggcacatggcaca actgagagac 180
acagacaagt gccaagggtgg 200

SEQ ID NO: 374          moltype = DNA length = 206
FEATURE
source
1..206
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 374
gcacagcagg tcagcaacaa gtttatttttgc cagcttagcaa ggtaacaggg tagggcatgg 60
ttacatgttc aggtcaacctt cctttgtcggtt ggttgcatttgc ttgggggggg 120
gtagggggaaa ggcggcaga agtacatgg agtgggtgcg ggcctccctgt agaacctgg 180
tacgagactt tggggcagtt cacccgtt 206

SEQ ID NO: 375          moltype = DNA length = 218
FEATURE
source
1..218
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 375
ggcgccatc agctccatga agatcattttgc tcgtacggct gccccgcagc 60
ctatgggtga caggcggtgt aaggcccttataaggcatgc cgggggggggg cggggcaggaa 120
ccccccccccccc ccccccggcc tggggccctt aaggccgcgc ctggcagetc tgctcacctt 180
gtcgccgtc acagccagca ccgcgcggcg atatgtc 218

SEQ ID NO: 376          moltype = DNA length = 169
FEATURE
source
1..169
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 376
catgatagta atcatagcct gatgttccaa gaaattcaaa tggcaatac cctattatgg 60
gtggtgccctt aaattacaca gaaaacaatc attattataa gttttccat aataagaat 120
attcaacta gaaataaaactt ttgggggggg ggctttttt ttcaaaaaa 169

SEQ ID NO: 377          moltype = DNA length = 178
FEATURE
source
1..178
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

SEQUENCE: 377
cagtggcatg aatcccatcc agactatgac cgaagtccga gccagacagg tgaaggagat 60
gaaccctgtt cttggcatcg actgtttgc caaggggaca aatggtgagg agctgtcact 120
cctctgcgtt gaggggggggg gatgtctgtt ttaaactgaa taatcatccca ctgttatgt 178

SEQ ID NO: 378          moltype = DNA length = 214
FEATURE
source
1..214
mol_type = other DNA
organism = synthetic construct
note = Description of Artificial Sequence: Synthetic
polynucleotide

```

-continued

polynucleotide

**SEQUENCE: 378**

```
ttgcatacc acccttgac ccactccacc cctggccacac agccccctcca cacccccccc 60
caacaacctt ctcaccttag atgctaagcc cacggcgaaa cacctgttac atggccgggg 120
catcatcata gtagtggtt agtagctgag ggccagaccc aatcacatg cgccgtgccc 180
cgccactgtc ctacaaggcca agcaccggcc agag 214
```

**SEQ ID NO: 379**

moltype = DNA length = 176

**FEATURE**  
source

1..176

mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 379**

```
tcaagttatca ttttaggcaa atgtgacccg acaagaataatc aggttgagatt tattttctgt 60
cctccataatt gttgggtgtc tggcacaagt ctttaaagga gatgttttt gaggctgtta 120
tggaaactgt tggtggggggg ggtggggggg catatatatta gttcatatat atttag 176
```

**SEQ ID NO: 380**

moltype = DNA length = 200

**FEATURE**  
source

1..200

mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 380**

```
gaagacagac tcttgcggag ccgtctccgg cagctcagtg ggccgttctt ggtcaggaa 60
ggccggctgt cccgcttcgc cggcaaggt gtggaggagc aggggctgag gacggagagc 120
ccggccggccg ctggccggaa tctagctcgc ccggcgcctt ctggcttctt cttcccccc 180
ccacttctt aaccctccggg 200
```

**SEQ ID NO: 381**

moltype = DNA length = 213

**FEATURE**  
source

1..213

mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 381**

```
accttcttctt cctcatagaa gtaaagctcc cggcccccac cagaggcgaa gcgaaaaggaa 60
atgaactctt gggcgtggaa gccatagat ggtctgttca ggagagaggg gcatgggtaa 120
aaaggaggcc ggaggagacc tggctgaagg atggctctc acccccccc cgcctgttcc 180
cccaaatcac ctgcacattt ttttagcttca agg 213
```

**SEQ ID NO: 382**

moltype = DNA length = 220

**FEATURE**  
source

1..220

mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 382**

```
tctgactctt ttcttcttaac tctagcggat cttggaaaaag tcaaaagcgag taggtgaagg 60
aagaacaagca actcacagac aagacttaggg gggggggcca ttgggtcttc acacccgttca 120
gtaaatgggg aagaacgacg agaaagcgcc agaataaccc cccgccccca gtcacacactc 180
gcacacttggg atgaacttgg ctgcgtcccc aagtgcggcc 220
```

**SEQ ID NO: 383**

moltype = DNA length = 196

**FEATURE**  
source

1..196

mol\_type = other DNA  
organism = synthetic construct  
note = Description of Artificial Sequence: Synthetic polynucleotide

**SEQUENCE: 383**

```
ggcagccaga gggcgagat gagtaacagg ggtaaacacct ccctgtgtcc cctccctctc 60
tctcccttctc tctccctctc tctttaccc ccccadggct ccacatccgc ctcaggggct 120
tctccacccc aagtctggct ccattctgg cttctgttg gtgacagacc cccccccctaa 180
ggtgtctgtt tgggggg 196
```

**SEQ ID NO: 384**

moltype = DNA length = 229

**FEATURE**  
source

1..229

mol\_type = other DNA  
organism = synthetic construct

-continued

---

note = Description of Artificial Sequence: Synthetic  
polynucleotide

SEQUENCE: 384

aagaaggccctt ctccattgtc tcagtcactt cagggtgcag gaagcggctg gccaaccgtt	60
cgtatgtcgctc cagcgctcagg cgactcaggg acctctcgta atccctggga gagaaggcga	120
ggaagattca ggggctgcctt gttttttttttt ggcatacact gcaacagagg taggggctgg	180
gagcaactgg ggctgaagag gatggccagg gttacccttc tcaccctct	229

---

1. A method for nucleic acid sequencing, comprising:
  - (a) exposing a plurality of template polynucleotide strands, including a set of homopolymer recalibration template polynucleotide strands, to a series of flows of nucleotide species flowed one nucleotide species at a time according to a flow ordering; and
  - (b) determining sequences of bases for the plurality of the template polynucleotide strands, including the set of homopolymer recalibration template polynucleotide strands, using sequencing by synthesis to generate a plurality of sequencing reads.

wherein each homopolymer recalibration template polynucleotide strand comprises an amplicon sequence including at least one homopolymer, the homopolymer having a homopolymer type, A, C, T, or G, wherein a homopolymer length of the homopolymer is between a minimal homopolymer length and a maximal homopolymer length.

wherein the set of homopolymer recalibration template polynucleotide strands comprises a set of amplicon sequences having at least a minimal threshold number of occurrences of each homopolymer length between the minimal homopolymer length and the maximal homopolymer length for one or more of the homopolymer types A, T, C, and G,

wherein the sequencing reads corresponding to the set of homopolymer recalibration template polynucleotide strands include the homopolymers having the homopolymer lengths between the minimal homopolymer length and the maximal homopolymer length for the one or more homopolymer types.
2. The method of claim 1, wherein the minimal threshold number of occurrences of each homopolymer length is 10, 25, or 50 occurrences for the set of amplicon sequences of the set of homopolymer recalibration template polynucleotide strands.
3. The method of claim 1, wherein a number of the amplicon sequences in the set of amplicon sequences is at least 288 amplicon sequences.
4. The method of claim 1, wherein the amplicon sequences of the homopolymer recalibration template polynucleotide strands include at most one homopolymer of length 6, 7, 8, 9, or 10 bases per amplicon sequence.
5. The method of claim 1, wherein the amplicon sequences of the homopolymer recalibration template polynucleotide strands have a minimal distance of 7 bases between separate homopolymers within the amplicon sequence when the homopolymer lengths of the separate homopolymers are 4, 5, 6, 7, 8, 9, or 10 bases.
6. The method of claim 1, wherein the amplicon sequences of the homopolymer recalibration template polynucleotide strands do not overlap.
7. The method of claim 1, wherein the minimal homopolymer length is 5 bases for the amplicon sequences of the set of amplicon sequences of the set of homopolymer recalibration template polynucleotide strands.
8. The method of claim 1, wherein the maximal homopolymer length is 10 bases for the amplicon sequences of the set of amplicon sequences of the set of homopolymer recalibration template polynucleotide strands.
9. The method of claim 1, wherein the set of amplicon sequences further comprises a group of amplicon sequences including additional n-mers of bases C and G.
10. A method for preparing a homopolymer recalibration panel, comprising:
  - extracting, from a set of amplicons used in sequencing-by-synthesis, a set of candidate amplicons satisfying a first set of criteria, wherein the first set of criteria includes amplicons known to belong to regions of a reference genome with no variants; and
  - selecting, from the set of candidate amplicons, a reduced set of amplicons satisfying a second set of criteria, wherein the second set of criteria includes selecting amplicon sequences that together comprise at least a minimal threshold number of homopolymers of each homopolymer length between a minimal homopolymer length and a maximal homopolymer length for one or more of homopolymer types A, T, C, and G.
11. The method of claim 10, wherein the minimal threshold number of occurrences of each homopolymer length is 10, 25, or 50 occurrences for the reduced set of amplicons.
12. The method of claim 10, wherein the minimal homopolymer length and the maximal homopolymer length is for each of homopolymer types A, T, C, and G.
13. The method of claim 10, further comprising:
  - determining underrepresented homopolymers of the set of candidate amplicons; and
  - augmenting the set of candidate amplicons with amplicons including the underrepresented homopolymers.
14. The method of claim 10, wherein the reference genome is NIST NA12878.
15. The method of claim 10, wherein the reduced set of amplicons comprises at most one homopolymer of length 6, 7, 8, 9, or 10 bases per amplicon sequence.
16. The method of claim 10, wherein the reduced set of amplicons comprises amplicon sequences having a minimal distance of 7 bases between separate homopolymers within the amplicon sequence when the homopolymer lengths of the separate homopolymers are 4, 5, 6, 7, 8, 9, or 10 bases.
17. The method of claim 10, wherein the amplicon sequences of the reduced set of amplicons do not overlap.
18. The method of claim 10, wherein the predetermined minimal homopolymer length is 5 bases for the amplicon sequences of reduced set of amplicons.
19. The method of claim 10, wherein the maximal homopolymer length is 10 bases for the amplicon sequences of the reduced set of amplicons.

**20.** The method of claim 10, wherein the reduced set of amplicons further comprises a group of amplicon sequences including additional n-mers of bases C and G.

\* \* \* \* \*