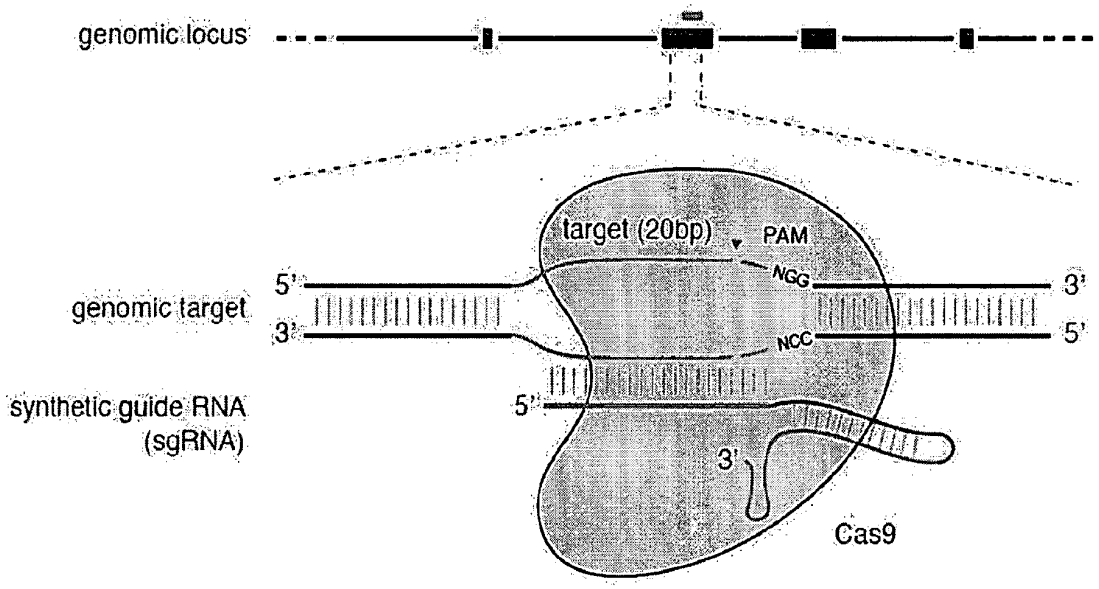

genomic locus
target (20bp)
PAM
NGG
NCC
5'
3'
genomic target
3'
5'
synthetic guide RNA
(sgRNA)
5'
3'
Cas9

FIG. 1

A

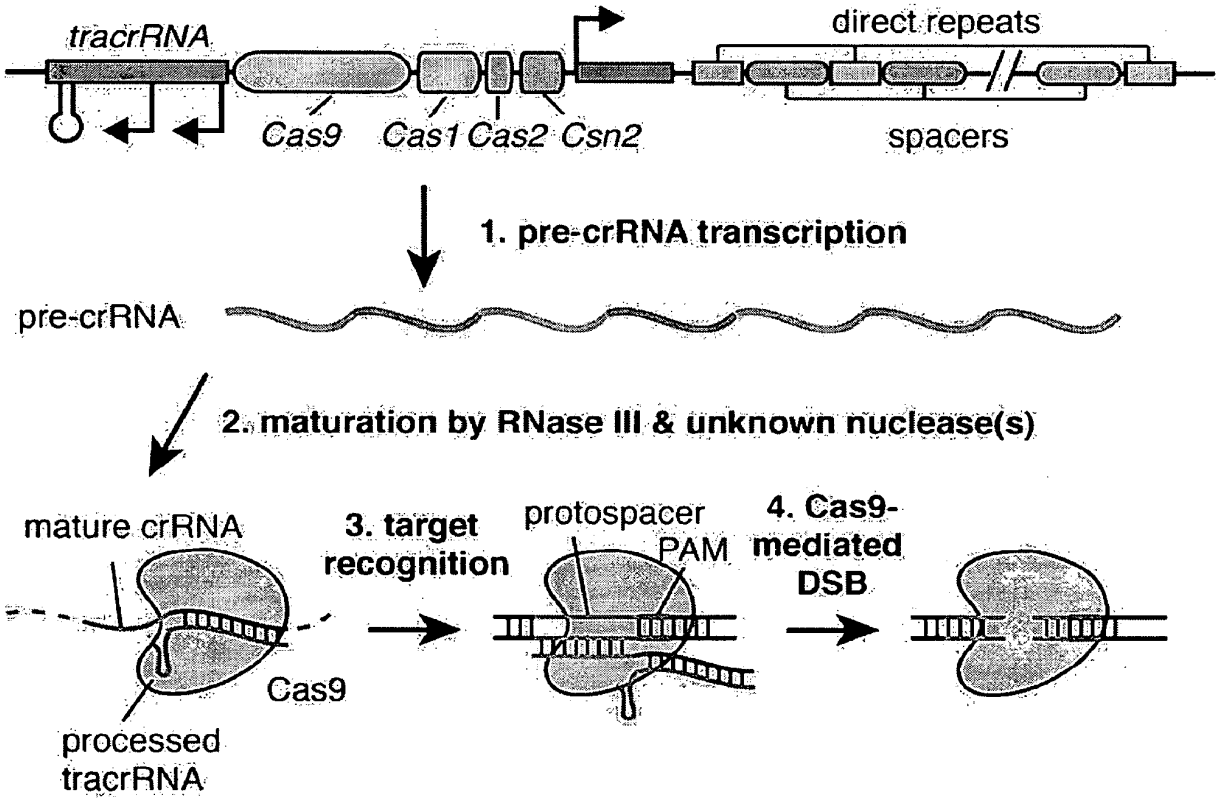

Streptococcus pyogenes SF370 CRISPR locus 1
tracrRNA
direct repeats
Cas9
Cas1
Cas2
Csn2
spacers
1. pre-crRNA transcription
pre-crRNA
2. maturation by RNase III & unknown nuclease(s)
mature crRNA
3. target recognition
protospacer
PAM
4. Cas9-mediated DSB
Cas9
processed tracrRNA

FIG. 2A

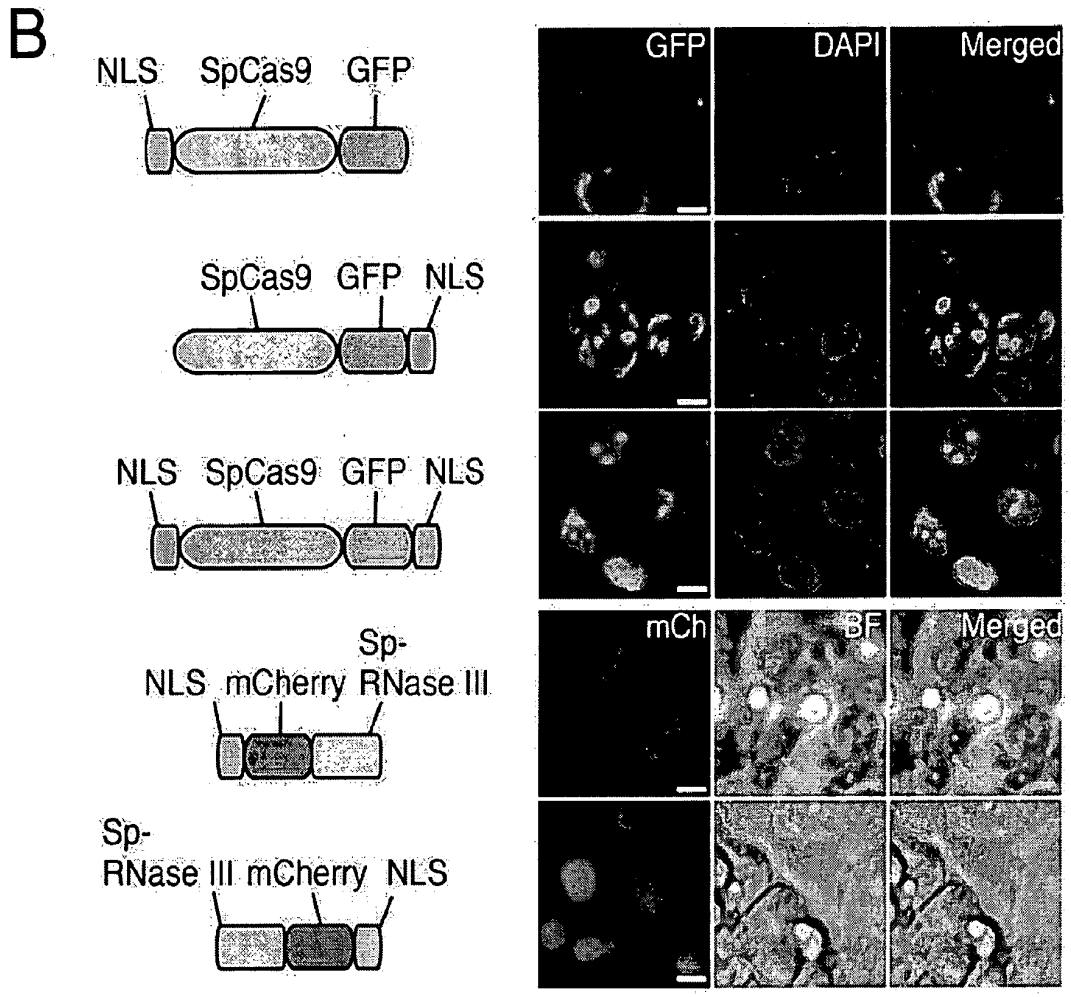

B
NLS SpCas9 GFP
SpCas9 GFP NLS
NLS SpCas9 GFP NLS
Sp-
NLS mCherry RNase III
Sp-
RNase III mCherry NLS
GFP
DAPI
Merged
mCh
BF
Merged

FIG. 2B

C

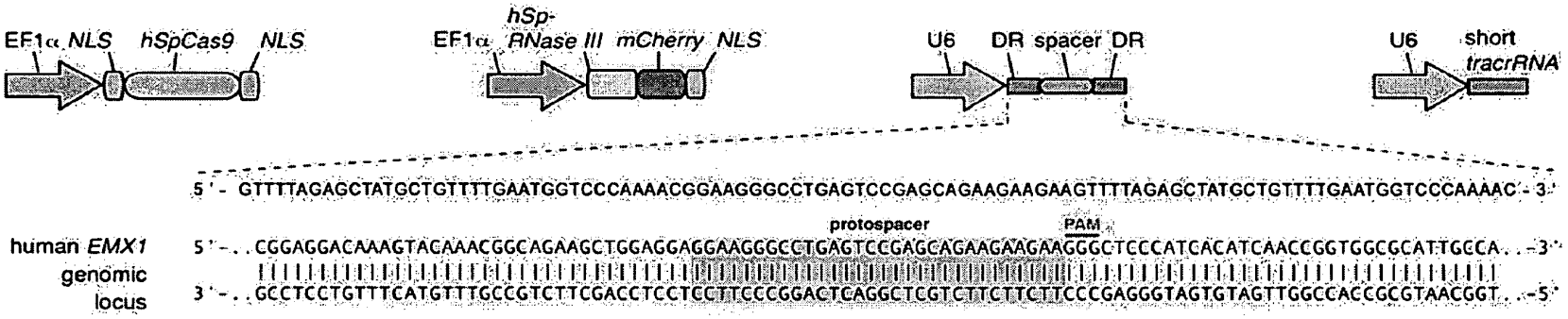

EF1α NLS hSpCas9 NLS
EF1α hSp-RNase III mCherry NLS
U6 DR spacer DR
U6 short tracrRNA
5'- GTTTTAGAGCTATGCTGTTTTGAATGGTCCCAAAACGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGTTTTAGAGCTATGCTGTTTTGAATGGTCCCAAAAC-3'
protospacer
PAM
human EMX1 genomic locus
5'-..CGGAGGACAAAGTACAAACGGCAGAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCACATCAACCGGTGGCGCATTGCCA...-3'
3'-..GCCTCCTGTTTCATGTTTGCCGTCTTCGACCTCCTCCTTCCCGGACTCAGGCTCGTCTTCTTCTTCCCGAGGGTAGTGTAGTTGGCCACCGCGTAACGGT...-5'

FIG. 2C

D

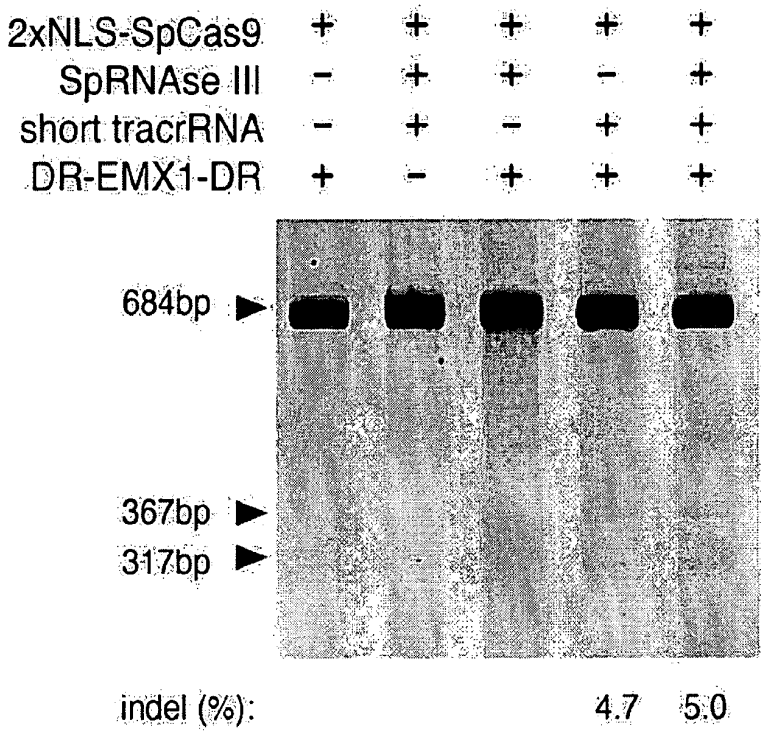

2xNLS-SpCas9 + + + + +
SpRNAse III - + + - +
short tracrRNA - + - + +
DR-EMX1-DR + - + + +
684bp
367bp
317bp
indel (%): 4.7 5.0

FIG. 2D

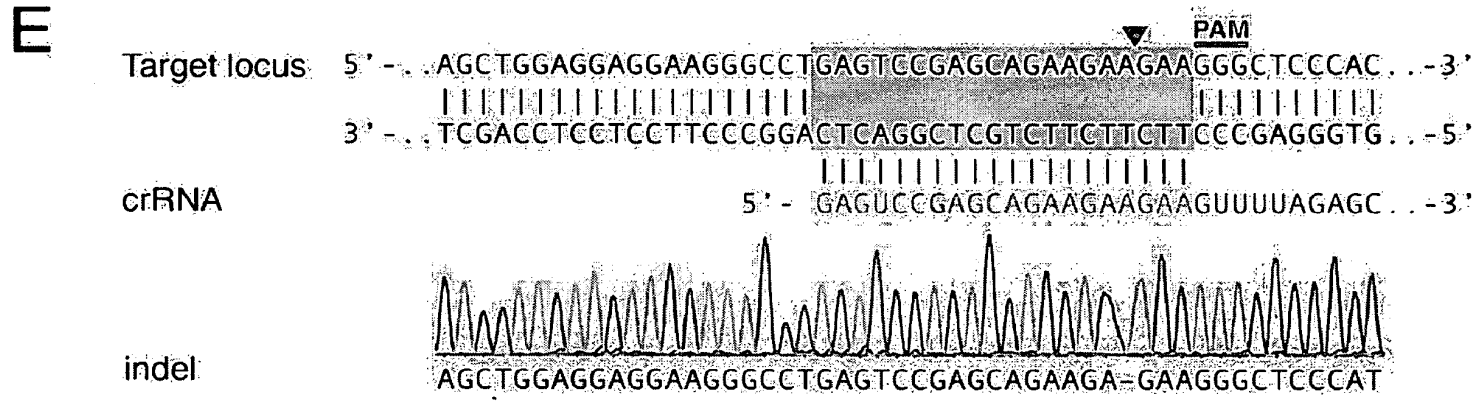
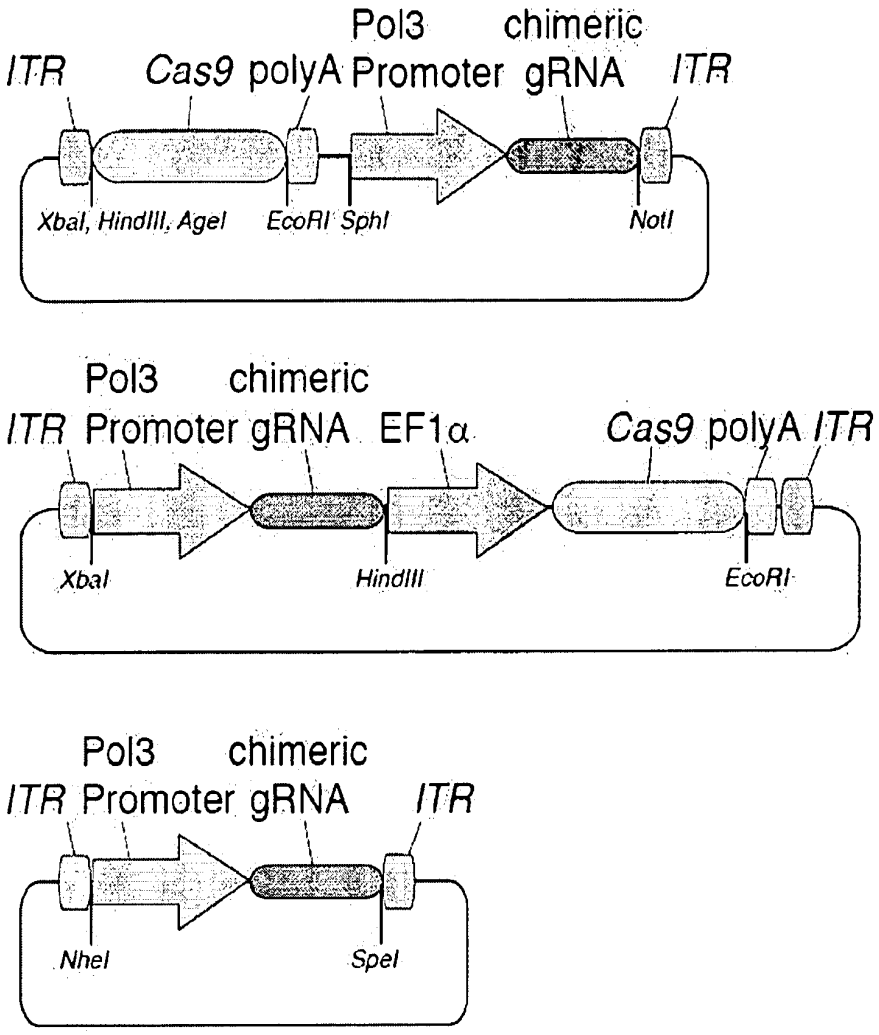

E
PAM
Target locus 5'-..AGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCAC..-3'
3'-..TCGACCTCCTCCTTCCCGGACTCAGGCTCGTCTTCTTCTTCCCGAGGGTG..-5'
crRNA
5'- GAGUCCGAGCAGAAGAAGAAGUUUUAGAGC..-3'
indel
AGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGA-GAAGGGCTCCCAT
F
human EMX1 protospacer target (mutation in 5 of 43 sequenced clones = 11.6%)
WT 5'-..CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAG AAGAAGGGCTCCCATCACAT..-3'
Δ1 CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAG--AGAAGGGCTCCCATCACAT
+1 CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAAGAAGGGCTCCCATCACAT
Δ3 CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAG----AAGGGCTCCCATCACAT
m1, Δ6 CTGGAGGAGGAAGGGCCTGAGCCCGAGCAGAAG-------GGCTCCCATCACAT

FIG. 2E-F



* depending on size of the entire construct,
can contain one or more Pol3-chimeric guide RNA (gRNA) cassette

FIG. 3

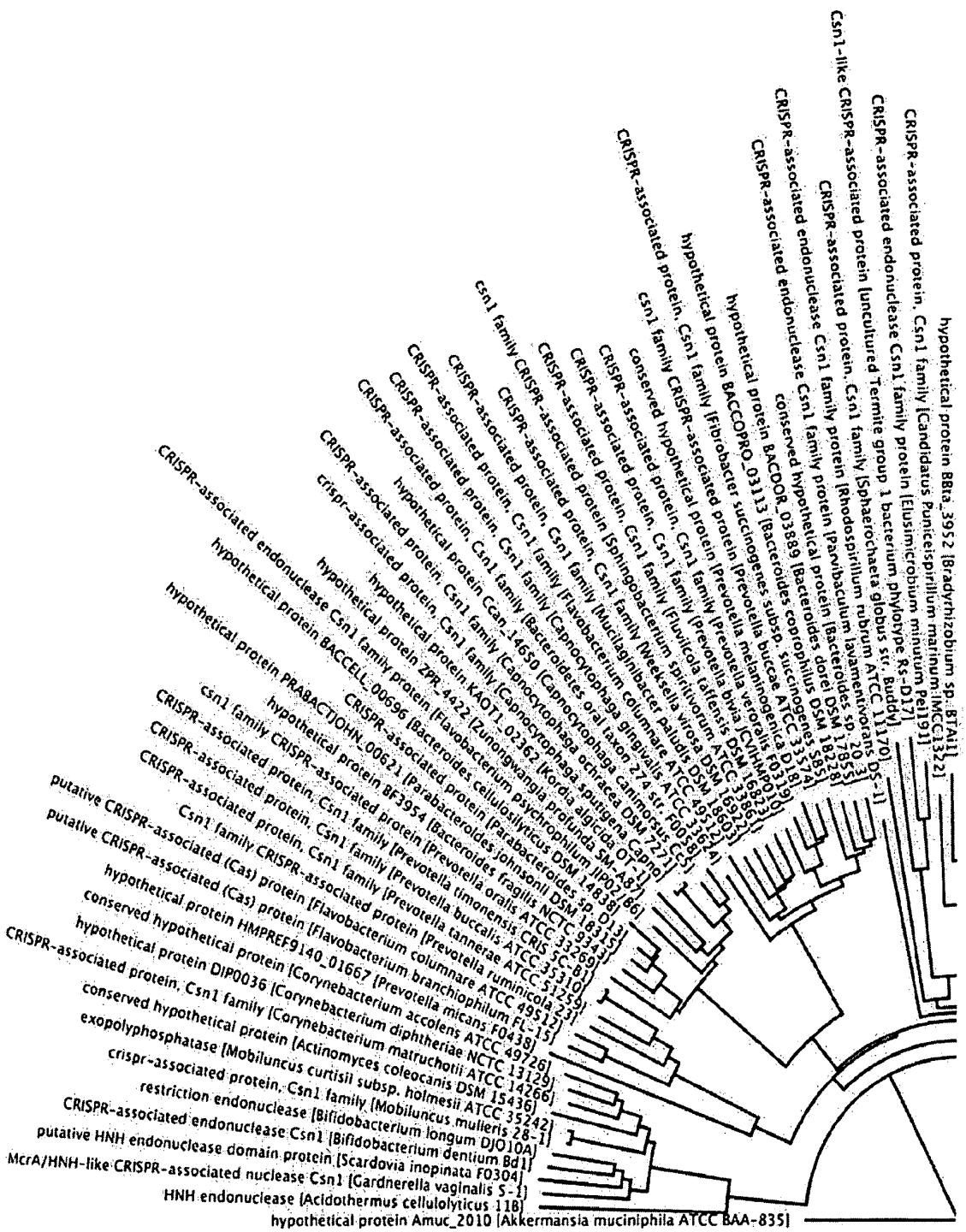

hypothetical protein BBta_3952 [Bradyrhizobium sp. BTAi1]
CRISPR-associated protein, Csn1 family [Candidatus Puniceispirillum marinum IMCC1322]
CRISPR-associated endonuclease Csn1 family protein [Elusimicrobium minutum Pei191]
CRISPR-associated protein [uncultured Termite group 1 bacterium phylotype Rs-D17]
Csn1-like CRISPR-associated protein, Csn1 family [Sphaerochaeta globus str. Buddy]
CRISPR-associated protein, Csn1 family protein [Rhodospirillum lavamentivorans DS-1]
CRISPR-associated endonuclease Csn1 family protein [Parvibaculum lavamentivorans DS-1]
CRISPR-associated endonuclease Csn1 family protein [Bacteroides dorei DSM 17855]
conserved hypothetical protein BACDOR_03889 [Bacteroides coprophilus DSM 18228]
hypothetical protein BACCOPRO_03113 [Bacteroides succinogenes subsp. succinogenes DSM 20539]
CRISPR-associated endonuclease Csn1 family protein [Bacteroides buccae ATCC 33574]
conserved hypothetical protein [Prevotella melaninogenica D18]
hypothetical protein [Fibrobacter succinogenes subsp. succinogenes S85]
CRISPR-associated protein, Csn1 family [Prevotella oris F0302]
csn1 family CRISPR-associated protein [Prevotella bivia JCVIHMP010]
CRISPR-associated protein, Csn1 family [Fluviicola taffensis DSM 16823]
CRISPR-associated protein, Csn1 family [Flavobacterium spiritivorum ATCC 33861]
CRISPR-associated protein, Csn1 family [Mucilaginibacter paludis DSM 18603]
CRISPR-associated protein, Csn1 family [Sphingobacterium thalpophilum ATCC 33952]
CRISPR-associated protein, Csn1 family [Weeksella virosa DSM 16922]
conserved hypothetical protein [Prevotella buccalis ATCC 35310]
CRISPR-associated protein, Csn1 family [Capnocytophaga gingivalis ATCC 33624]
CRISPR-associated protein, Csn1 family [Capnocytophaga canimorsus Cc5]
CRISPR-associated protein, Csn1 family [Capnocytophaga ochracea DSM 7271]
csn1 family CRISPR-associated protein [Capnocytophaga sputigena Capno]
CRISPR-associated protein, Csn1 family [Flavobacteriales oral taxon 274 str. F0058]
CRISPR-associated protein, Csn1 family [Bacteroidetes oral taxon 274 str. F0058]
CRISPR-associated protein, Csn1 family [Capnocytophaga ochracea DSM 7271]
hypothetical protein Csan_14650 [Capnocytophaga sputigena Capno]
CRISPR-associated protein, Csn1 family [Capnocytophaga ochracea DSM 7271]
hypothetical protein KAOT1_02362 [Kordia algicida OT-1]
CRISPR-associated protein, Csn1 family [Zunongwangia profunda SM-A87]
hypothetical protein ZPR_4422 [Zunongwangia profunda SM-A87]
CRISPR-associated protein, Csn1 family [Flavobacterium psychrophilum JIP02/86]
hypothetical protein BACCELL_00696 [Bacteroides cellulosilyticus DSM 14838]
CRISPR-associated protein [Parabacteroides johnsonii DSM 18315]
hypothetical protein PRABACTJOHN_00621 [Parabacteroides johnsonii DSM 18315]
CRISPR-associated endonuclease Csn1 family protein [Flavobacterium psychrophilum JIP02/86]
hypothetical protein BF3954 [Bacteroides fragilis NCTC 9343]
csn1 family CRISPR-associated protein [Prevotella oralis ATCC 33269]
CRISPR-associated protein, Csn1 family [Prevotella buccalis ATCC 35310]
CRISPR-associated protein, Csn1 family [Prevotella timonensis CRIS 5C-B1]
CRISPR-associated protein [Flavobacterium columnare ATCC 49512]
hypothetical protein [Prevotella tannerae ATCC 51259]
hypothetical protein HMPREF9140_01667 [Prevotella ruminicola 23]
putative CRISPR-associated (Cas) protein [Prevotella micans F0438]
CRISPR-associated protein, Csn1 family [Flavobacterium branchiophilum FL-15]
Csn1 family CRISPR-associated protein [Corynebacterium accolens ATCC 49726]
putative CRISPR-associated (Cas) protein [Corynebacterium diphtheriae NCTC 13129]
hypothetical protein DIP0036 [Corynebacterium matruchotii ATCC 14266]
conserved hypothetical protein [Actinomyces coleocanis DSM 15436]
hypothetical protein [Mobiluncus curtisii subsp. holmesii ATCC 35242]
exopolyphosphatase [Bifidobacterium mulieris 28-1]
crispr-associated protein, Csn1 family [Bifidobacterium longum DJO10A]
restriction endonuclease [Bifidobacterium dentium Bd1]
CRISPR-associated endonuclease Csn1 [Scardovia inopinata F0304]
putative HNH endonuclease domain protein [Gardnerella vaginalis S-1]
McrA/HNH-like CRISPR-associated nuclease Csn1 [Acidothermus cellulolyticus 11B]
HNH endonuclease [Akkermansia muciniphila ATCC BAA-835]
hypothetical protein Amuc_2010 [Akkermansia muciniphila ATCC BAA-835]

FIG. 4A


hypothetical protein Nham_4054 [Nitrobacter hamburgensis X14]
CRISPR-associated protein, Csn1 family [Azospirillum sp. B510]
hypothetical protein [Dinoroseobacter shibae DFL 12]
conserved hypothetical protein SU517_3938 [Sphingomonas sp. S17]
hypothetical protein HDN1F_34120 [gamma proteobacterium HdN1]
CRISPR-associated protein [Ralstonia syzygii R24]
CRISPR-associated endonuclease Csn1 family [Nitrosomonas sp. AL212]
CRISPR-associated protein, csn1 family [Acidovorax avenae subsp. avenae ATCC 19860]
CRISPR-associated protein WS1445 [Wolinella succinogenes DSM 1740]
CRISPR-associated protein [Helicobacter cinaedi CCUG 18818]
CRISPR-associated protein CDL2123 [Gluconacetobacter diazotrophicus PAL 5]
CRISPR-associated protein RPC_4489 [Rhodopseudomonas palustris BisB18]
CRISPR-associated protein, Csn1 family [Methylocystis sp. ATCC 49242]
CRISPR-associated protein, Csn1 family [Clostridium cellulolyticum H10]
CRISPR-associated protein, Csn1 family [Bacillus cereus Rock1-15]
CRISPR-associated protein, Csn1 family [Roseburia intestinalis L1-82]
CRISPR-associated protein, Csn1 family [Lactobacillus coryniformis subsp. torquens KCTC 3535]
CRISPR-associated protein, Csn1 family [Aminomonas paucivorans DSM 12260]
CRISPR-associated protein, Csn1 family [Simonsiella muelleri ATCC 29453]
CRISPR-associated protein, Csn1 family [Neisseria cinerea ATCC 14685]
CRISPR-associated protein, Csn1 family [Neisseria lactamica 020-06]
CRISPR-associated protein, Csn1 family [Neisseria meningitidis 053442]
CRISPR-associated protein, Csn1 family [Neisseria meningitidis Z2491]
CRISPR-associated protein, Csn1 family [Neisseria meningitidis alpha14]
CRISPR-associated protein, Csn1 family [Neisseria meningitidis ATCC BAA-1200]
CRISPR-associated protein, Csn1 family [Neisseria bacilliformis ATCC 23330]
CRISPR-associated protein, Csn1 family [Kingella kingae ATCC 23330]
CRISPR-associated protein, Csn1 family [Actinobacillus minor NM305]
CRISPR-associated protein, Csn1 family [Actinobacillus pleuropneumoniae serovar 10 str. D13039]
CRISPR-associated protein, Csn1 family [Haemophilus parainfluenzae T3T1]
CRISPR-associated protein, Csn1 family [Ruminococcus albus 8]
CRISPR-associated protein, csn1 family [Nitratifractor salsuginis DSM 16511]
hypothetical protein MMO80330 [Mycoplasma mobile 163K]

FIG. 4B

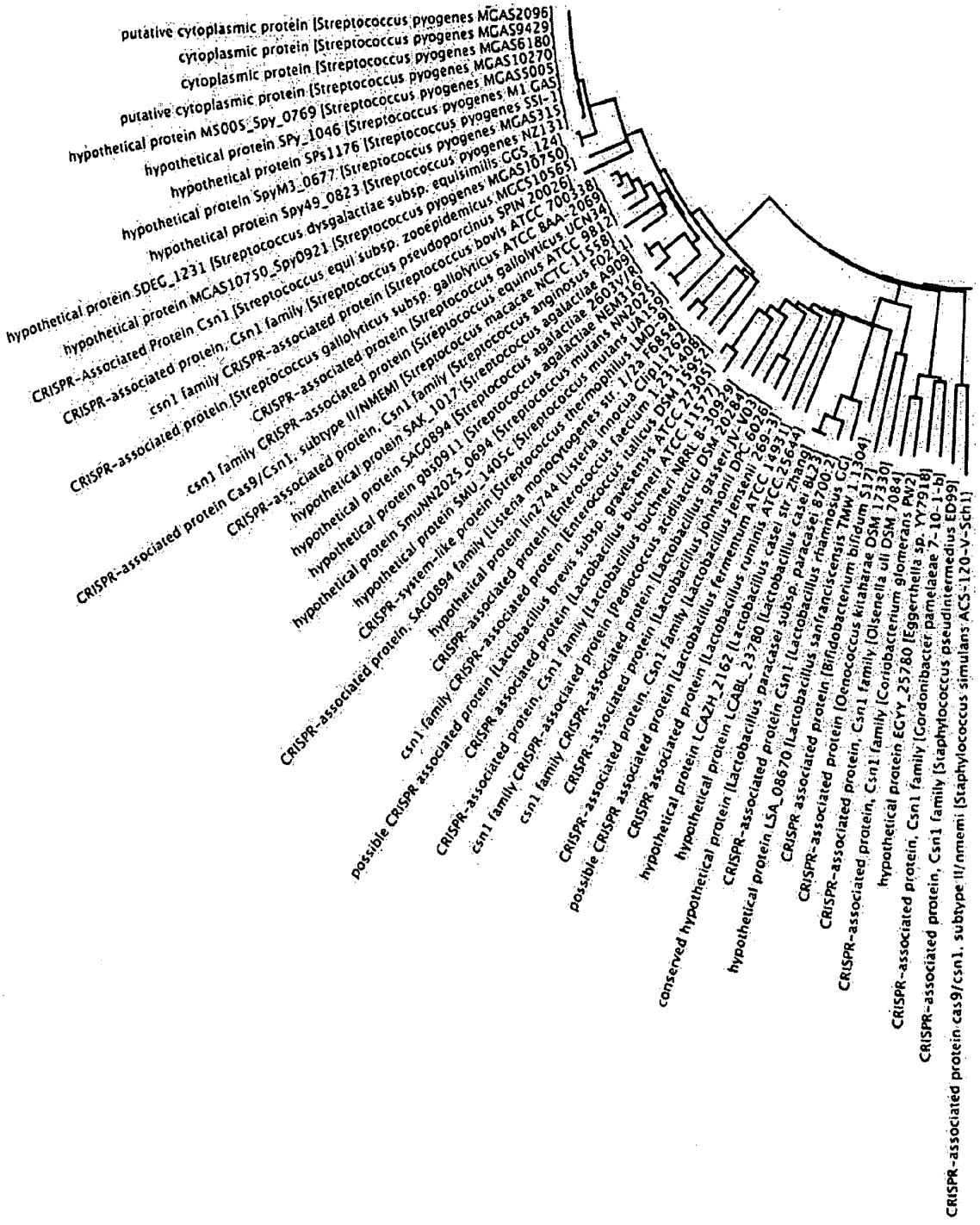

putative cytoplasmic protein [Streptococcus pyogenes MGAS2096]
cytoplasmic protein [Streptococcus pyogenes MGAS9429]
cytoplasmic protein [Streptococcus pyogenes MGAS6180]
putative cytoplasmic protein [Streptococcus pyogenes MGAS10270]
hypothetical protein MS00S_Spy_0769 [Streptococcus pyogenes MGAS5005]
hypothetical protein SPy_1046 [Streptococcus pyogenes M1 GAS]
hypothetical protein SPs1176 [Streptococcus pyogenes SSI-1]
hypothetical protein SpyM3_0677 [Streptococcus pyogenes MGAS315]
hypothetical protein Spy49_0823 [Streptococcus pyogenes NZ131]
hypothetical protein SDEG_1231 [Streptococcus pyogenes MGAS10750]
hypothetical protein MGAS10750_Spy0921 [Streptococcus pyogenes MGAS10056]
CRISPR-Associated Protein Csn1 [Streptococcus dysgalactiae subsp. equisimilis GGS_124]
CRISPR-associated protein, Csn1 family [Streptococcus equi subsp. zooepidemicus SPIN 20026]
csn1 family CRISPR-associated protein [Streptococcus pseudoporcinus SPIN 20026]
CRISPR-associated protein [Streptococcus gallolyticus subsp. gallolyticus ATCC 700338]
CRISPR-associated protein Cas9/Csn1, subtype II/NMEMI [Streptococcus bovis ATCC BAA-2069]
csn1 family CRISPR-associated protein [Streptococcus macacae NCTC 11558]
CRISPR-associated protein, Csn1 family [Streptococcus equinus ATCC 9812]
hypothetical protein SAK_1017 [Streptococcus gallolyticus UCN34]
hypothetical protein gbs0911 [Streptococcus anginosus F0211]
hypothetical protein SMU_1405c [Streptococcus agalactiae A909]
hypothetical protein SAG0894 [Streptococcus agalactiae 2603V/R]
hypothetical protein SmuNN2025_0694 [Streptococcus agalactiae NEM316]
CRISPR-system-like protein [Streptococcus mutans UA159]
hypothetical protein lin2744 [Listeria monocytogenes str. 1/2a F6854]
CRISPR-associated protein [Enterococcus faecium 1,231,408]
csn1 family CRISPR-associated protein [Lactobacillus brevis subsp. gravesensis ATCC 27305]
CRISPR-associated protein, Csn1 family [Lactobacillus buchneri NRRL B-30929]
possible CRISPR associated protein [Pediococcus acidilactici DSM 20284]
CRISPR-associated protein, Csn1 family [Lactobacillus johnsonii DPC 6026]
CRISPR associated protein [Lactobacillus gasseri JV-V03]
hypothetical protein LCAZH_2162 [Lactobacillus fermentum ATCC 14931]
hypothetical protein LCABL_23780 [Lactobacillus ruminis ATCC 25644]
CRISPR-associated protein Csn1 [Lactobacillus casei str. Zhang]
conserved hypothetical protein [Lactobacillus paracasei subsp. paracasei 8]
hypothetical protein LSA_08670 [Lactobacillus casei BL23]
CRISPR associated protein [Lactobacillus sanfranciscensis TMW 1.1304]
CRISPR-associated protein [Lactobacillus rhamnosus GG]
CRISPR-associated protein, Csn1 family [Bifidobacterium bifidum S17]
CRISPR-associated protein [Oenococcus kitaharae DSM 17330]
hypothetical protein, Csn1 family [Olsenella uli DSM 7084]
CRISPR-associated protein, Csn1 family [Coriobacterium glomerans PW2]
hypothetical protein EGYY_25780 [Eggerthella sp. YY7918]
CRISPR-associated protein, Csn1 family [Gordonibacter pamelaeae 7-10-1-b]
CRISPR-associated protein Cas9/csn1, subtype II/nmemi [Staphylococcus pseudintermedius ED99]
[Staphylococcus simulans ACS-120-V-Sch1]

FIG. 4C

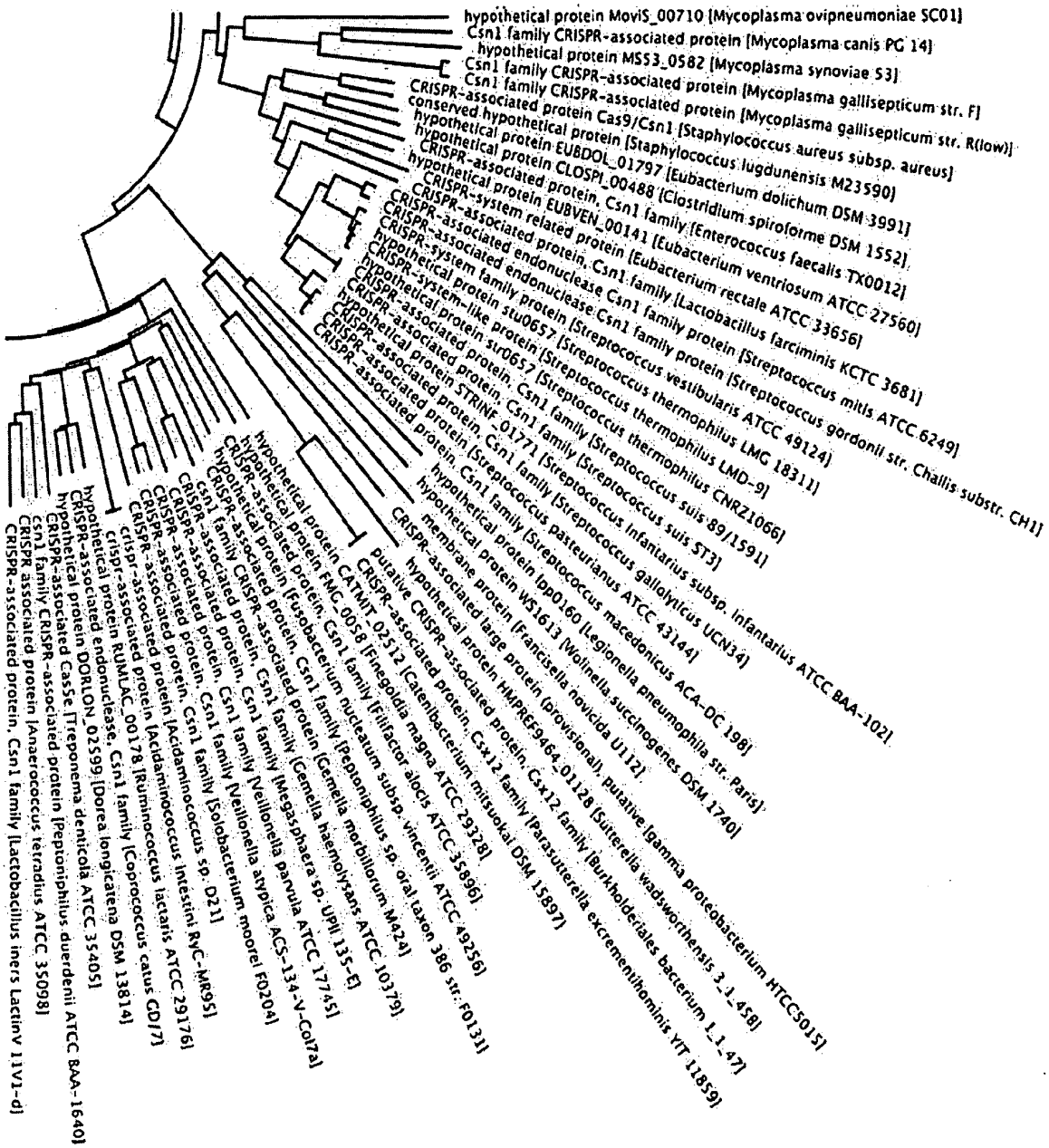

hypothetical protein MoviS_00710 [Mycoplasma ovipneumoniae SC01]
Csn1 family CRISPR-associated protein [Mycoplasma canis PG 14]
hypothetical protein MS53_0582 [Mycoplasma synoviae 53]
Csn1 family CRISPR-associated protein [Mycoplasma gallisepticum str. F]
CRISPR-associated protein Cas9/Csn1 [Staphylococcus aureus subsp. aureus]
conserved hypothetical protein [Staphylococcus lugdunensis M23590]
hypothetical protein EUBDOL_01797 [Eubacterium dolichum DSM 3991]
CRISPR-associated protein, Csn1 family [Enterococcus faecalis TX0012]
hypothetical protein CLOSPI_00488 [Clostridium spiroforme DSM 1552]
CRISPR-system related protein [Eubacterium rectale ATCC 33656]
CRISPR-associated protein, Csn1 family [Lactobacillus farciminis KCTC 3681]
CRISPR-associated endonuclease Csn1 family protein [Streptococcus mitis ATCC 6249]
CRISPR-system family protein [Streptococcus vestibularis ATCC 49124]
hypothetical protein stu0657 [Streptococcus thermophilus LMG 18311]
CRISPR-system-like protein [Streptococcus thermophilus CNRZ1066]
CRISPR-associated protein, Csn1 family [Streptococcus thermophilus LMD-9]
hypothetical protein STRINF_01771 [Streptococcus suis 89/1591]
CRISPR-associated protein, Csn1 family [Streptococcus suis ST3]
CRISPR-associated protein, Csn1 family [Streptococcus gallolyticus UCN34]
hypothetical protein lpp0160 [Legionella pneumophila str. Paris]
hypothetical protein WS1613 [Wolinella succinogenes DSM 1740]
membrane protein [Francisella novicida U112]
hypothetical large protein (provisional), putative [gamma proteobacterium HTCC5015]
CRISPR-associated protein, Csx12 family [Sutterella wadsworthensis 3_1_45]
putative CRISPR-associated protein, Csx12 family [Burkholderiales bacterium 1_1_47]
hypothetical protein HMPREF9464_01128 [Sutterella wadsworthensis 3_1_45]
CRISPR-associated protein, Csx12 family [Parasutterella excrementihominis YIT 11859]
hypothetical protein CATMIT_02512 [Catenibacterium mitsuokai DSM 15897]
CRISPR-associated protein, Csn1 family [Filifactor alocis ATCC 35896]
hypothetical protein FMG_0058 [Finegoldia magna ATCC 29328]
hypothetical protein [Fusobacterium nucleatum subsp. vincentii ATCC 49256]
CRISPR-associated protein, Csn1 family [Peptoniphilus sp. oral taxon 386 str. F0131]
hypothetical protein [Fusobacterium nucleatum subsp. nucleatum ATCC 10379]
CRISPR-associated protein, Csn1 family [Peptoniphilus sp. oral taxon 386 str. F0131]
CRISPR-associated protein, Csn1 family [Gemella morbillorum M424]
hypothetical protein [Fusobacterium nucleatum subsp. nucleatum ATCC 10379]
CRISPR-associated protein, Csn1 family [Gemella haemolysans ATCC 10379]
CRISPR-associated protein, Csn1 family [Megasphaera sp. UPII 135-E]
CRISPR-associated protein, Csn1 family [Veillonella atypica ACS-134-V-Col7a]
CRISPR-associated protein, Csn1 family [Veillonella parvula ATCC 17745]
Csn1 family CRISPR-associated protein [Veillonella atypica ACS-134-V-Col7a]
CRISPR-associated protein, Csn1 family [Veillonella parvula ATCC 17745]
CRISPR-associated protein, Csn1 family [Solobacterium moorei F0204]
CRISPR-associated protein, Csn1 family [Solobacterium moorei F0204]
CRISPR-associated protein, Csn1 family [Acidaminococcus sp. D21]
CRISPR-associated protein [Acidaminococcus intestini RyC-MR95]
CRISPR-associated protein [Ruminococcus lactaris ATCC 29176]
crispr-associated protein RUMLAC_00178 [Ruminococcus lactaris ATCC 29176]
crispr-associated protein [Coprococcus catus GD/7]
hypothetical protein RUMLAC_02599 [Dorea longicatena DSM 13814]
CRISPR-associated endonuclease, Csn1 family [Coprococcus catus GD/7]
hypothetical protein DORLON_02599 [Dorea longicatena DSM 13814]
CRISPR-associated protein [Treponema denticola ATCC 35405]
hypothetical protein TREDEN_0405 [Treponema denticola ATCC 35405]
CRISPR-associated CasSe [Treponema denticola ATCC BAA-1640]
CRISPR-associated protein [Peptoniphilus duerdenii ATCC BAA-1640]
Csn1 family CRISPR-associated protein [Anaerococcus tetradius ATCC 35098]
CRISPR-associated protein [Anaerococcus tetradius ATCC 35098]
CRISPR-associated protein, Csn1 family [Lactobacillus iners LactinV 11V1-d]

FIG. 4D

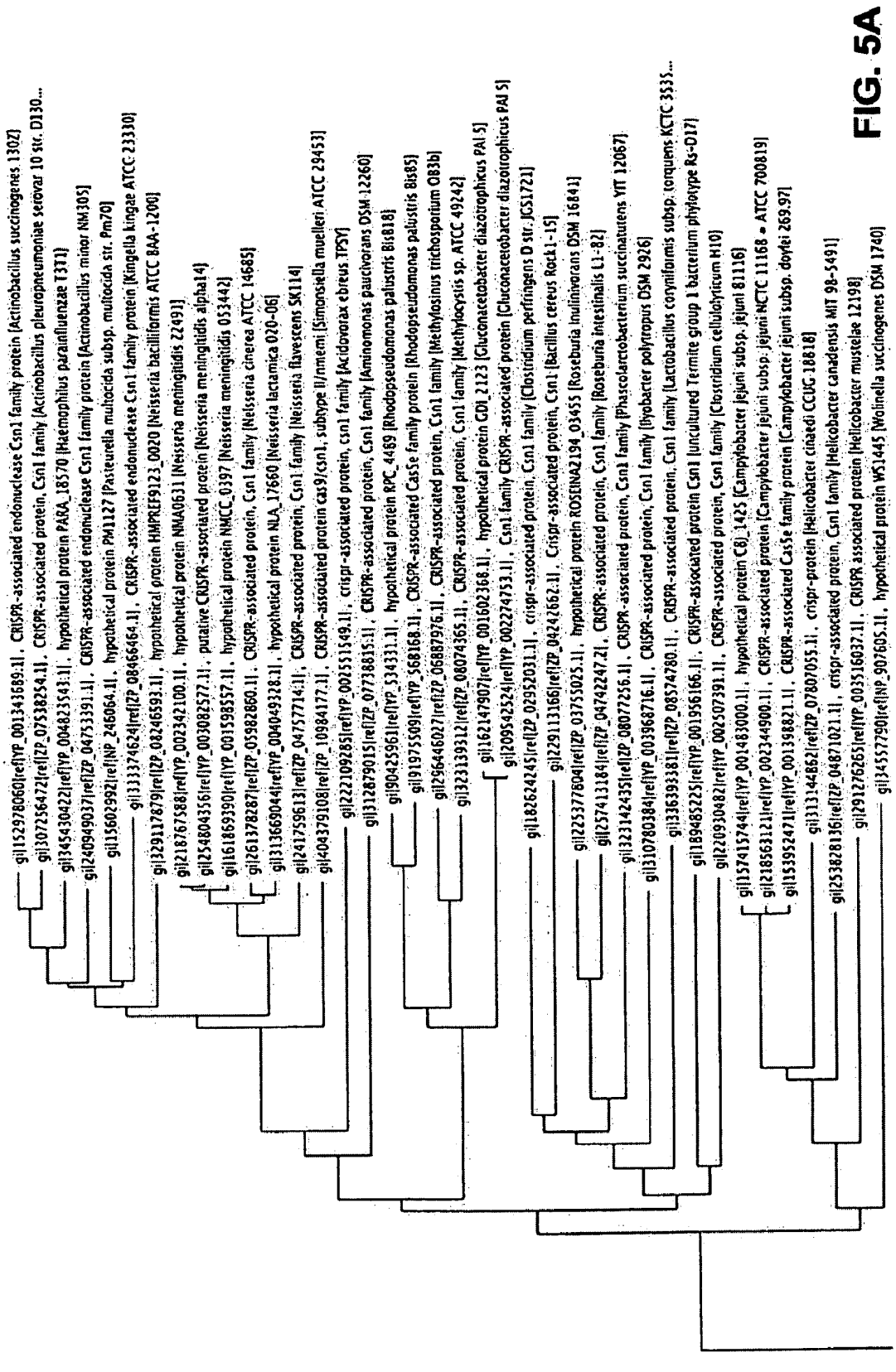

gi|152978060|ref|YP_001343689.1|, CRISPR-associated endonuclease Csn1 family protein [Actinobacillus succinogenes 130Z]
gi|307256472|ref|ZP_07538254.1|, CRISPR-associated protein, Csn1 family [Actinobacillus pleuropneumoniae serovar 10 str. D130...
gi|345430422|ref|YP_004823541.1|, hypothetical protein PARA_18570 [Haemophilus parainfluenzae T3T1]
gi|240949037|ref|ZP_04753391.1|, CRISPR-associated endonuclease Csn1 family protein [Actinobacillus minor NM305]
gi|156022992|ref|NP_246064.1|, hypothetical protein PM1127 [Pasteurella multocida subsp. multocida str. Pm70]
gi|333374624|ref|ZP_08466464.1|, CRISPR-associated endonuclease Csn1 family protein [Kingella kingae ATCC 23330]
gi|329117879|ref|ZP_08246593.1|, hypothetical protein HMPREF9123_0020 [Neisseria bacilliformis ATCC BAA-1200]
gi|218767588|ref|YP_002342100.1|, hypothetical protein NMA0631 [Neisseria meningitidis Z2491]
gi|254804356|ref|YP_003082577.1|, putative CRISPR-associated protein [Neisseria meningitidis alpha14]
gi|161869390|ref|YP_001598557.1|, hypothetical protein NMCC_0397 [Neisseria meningitidis 053442]
gi|261378287|ref|ZP_05982860.1|, CRISPR-associated protein, Csn1 family [Neisseria cinerea ATCC 14685]
gi|313669044|ref|YP_004049328.1|, hypothetical protein NLA_17660 [Neisseria lactamica 020-06]
gi|241759613|ref|ZP_04757714.1|, CRISPR-associated protein, Csn1 family [Neisseria flavescens SK114]
gi|404379108|ref|ZP_10984177.1|, CRISPR-associated protein cas9/csn1, subtype II/nmemi [Simonsiella muelleri ATCC 29453]
gi|222109285|ref|YP_002551549.1|, crispr-associated protein, csn1 family [Acidovorax ebreus TPSY]
gi|312879015|ref|ZP_07738815.1|, CRISPR-associated protein, Csn1 family [Aminomonas paucivorans DSM 12260]
gi|90425961|ref|YP_534331.1|, hypothetical protein RPC_4489 [Rhodopseudomonas palustris BisB18]
gi|91975509|ref|YP_568168.1|, CRISPR-associated Cas5e family protein [Rhodopseudomonas palustris BisB5]
gi|296446027|ref|ZP_06887976.1|, CRISPR-associated protein, Csn1 family [Methylosinus trichosporium OB3b]
gi|323139312|ref|ZP_08074365.1|, CRISPR-associated protein, Csn1 family [Methylocystis sp. ATCC 49242]
gi|162147907|ref|YP_001602368.1|, hypothetical protein GDI_2123 [Gluconacetobacter diazotrophicus PAl 5]
gi|209542524|ref|YP_002274753.1|, Csn1 family CRISPR-associated protein [Gluconacetobacter diazotrophicus PAl 5]
gi|182624245|ref|ZP_02952031.1|, crispr-associated protein, Csn1 family [Clostridium perfringens D str. JGS1721]
gi|229113166|ref|ZP_04243662.1|, Crispr-associated protein, Csn1 [Bacillus cereus Rock1-15]
gi|225377604|ref|ZP_03755025.1|, hypothetical protein ROSEINA2194_03455 [Roseburia inulinivorans DSM 16841]
gi|257413184|ref|ZP_04742247.2|, CRISPR-associated protein, Csn1 family [Roseburia intestinalis L1-82]
gi|323142435|ref|ZP_08077256.1|, CRISPR-associated protein, Csn1 family [Phascolarctobacterium succinatutens YIT 12067]
gi|310780384|ref|YP_003968716.1|, CRISPR-associated protein, Csn1 family [Hyobacter polytropus DSM 2926]
gi|336393381|ref|ZP_08574780.1|, CRISPR-associated protein, Csn1 family [Lactobacillus coryniformis subsp. torquens KCTC 3535...
gi|189485225|ref|YP_001956166.1|, CRISPR-associated protein Csn1 [uncultured Termite group 1 bacterium phylotype Rs-D17]
gi|220930482|ref|YP_002507391.1|, CRISPR-associated protein, Csn1 family [Clostridium cellulolyticum H10]
gi|157415744|ref|YP_001483000.1|, hypothetical protein CJ0_1425 [Campylobacter jejuni subsp. jejuni 81116]
gi|218563121|ref|YP_002344900.1|, CRISPR-associated protein [Campylobacter jejuni subsp. jejuni NCTC 11168 = ATCC 700819]
gi|153952471|ref|YP_001398821.1|, CRISPR-associated Cas5e family protein [Campylobacter jejuni subsp. doylei 269.97]
gi|313144862|ref|ZP_07807055.1|, crispr-protein [Helicobacter cinaedi CCUG 18818]
gi|253828136|ref|ZP_04871021.1|, crispr-associated protein, Csn1 family [Helicobacter canadensis MIT 98-5491]
gi|291276265|ref|YP_003516037.1|, CRISPR associated protein [Helicobacter mustelae 12198]
gi|345577901|ref|NP_907605.1|, hypothetical protein WS1445 [Wolinella succinogenes DSM 1740]

FIG. 5A

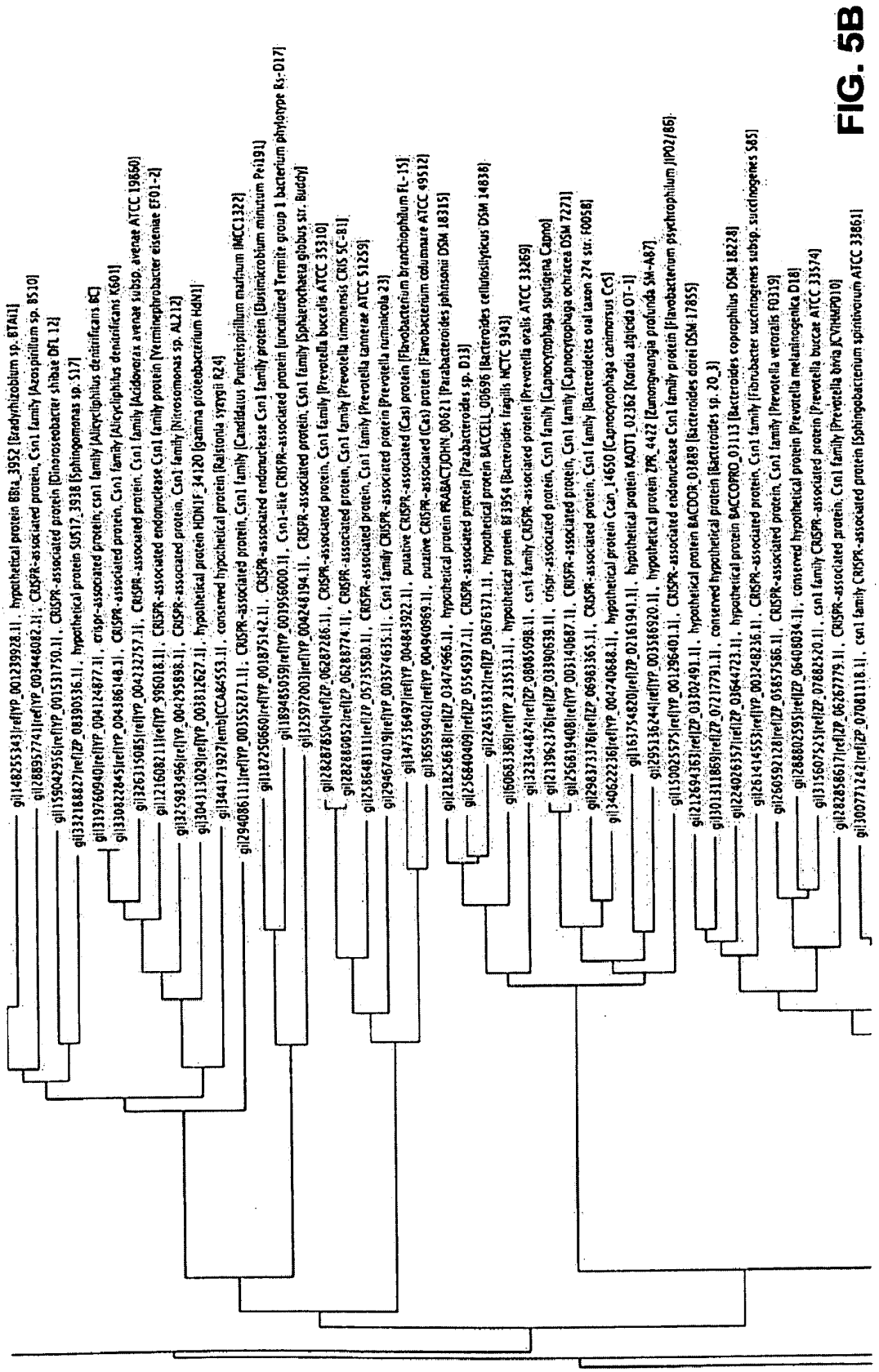

gi|148255343|ref|YP_001239928.1|, hypothetical protein BBta_3952 [Bradyrhizobium sp. BTAi1]
gi|288957741|ref|YP_003446082.1|, CRISPR-associated protein, Csn1 family [Azospirillum sp. B510]
gi|159042956|ref|YP_001531750.1|, CRISPR-associated protein [Dinoroseobacter shibae DFL 12]
gi|332188827|ref|ZP_08390536.1|, hypothetical protein SJS17_3938 [Sphingomonas sp. S17]
gi|319760940|ref|YP_004124877.1|, crispr-associated protein, csn1 family [Alicycliphilus denitrificans BC]
gi|330822845|ref|YP_004186148.1|, CRISPR-associated protein, Csn1 family [Alicycliphilus denitrificans K601]
gi|326315085|ref|YP_004232757.1|, CRISPR-associated protein, Csn1 family [Acidovorax avenae subsp. avenae ATCC 19860]
gi|121608211|ref|YP_996018.1|, CRISPR-associated endonuclease Csn1 family protein [Verminephrobacter eiseniae EF01-2]
gi|325983496|ref|YP_004295898.1|, CRISPR-associated protein, Csn1 family [Nitrosomonas sp. AL212]
gi|304311029|ref|YP_003812627.1|, hypothetical protein HDN1F_34120 [gamma proteobacterium HdN1]
gi|344171927|emb|CCA84553.1|, conserved hypothetical protein [Ralstonia syzygii R24]
gi|294086111|ref|YP_003552871.1|, CRISPR-associated protein, Csn1 family [Candidatus Puniceispirillum marinum IMCC1322]
gi|118725060|ref|YP_001875142.1|, CRISPR-associated endonuclease Csn1 family protein [Burkholderia minutum P15191]
gi|189485059|ref|YP_001956000.1|, Csn1-like CRISPR-associated protein [uncultured Termite group 1 bacterium phylotype Rs-D17]
gi|312597200|ref|YP_004244194.1|, CRISPR-associated protein, Csn1 family [Spherochaeta globus str. Buddy]
gi|282878504|ref|ZP_06287286.1|, CRISPR-associated protein, Csn1 family [Prevotella buccalis ATCC 35310]
gi|282886052|ref|ZP_06288774.1|, CRISPR-associated protein, Csn1 family [Prevotella timonensis CRIS 5C-B1]
gi|258648111|ref|ZP_05735580.1|, CRISPR-associated protein, Csn1 family [Prevotella tannerae ATCC 51259]
gi|294674019|ref|YP_003574635.1|, Csn1 family CRISPR-associated protein [Prevotella ruminicola 23]
gi|347536497|ref|YP_004843922.1|, putative CRISPR-associated (Cas) protein [Flavobacterium branchiophilum FL-15]
gi|365959402|ref|YP_004960969.1|, putative CRISPR-associated (Cas) protein [Flavobacterium columnare ATCC 49512]
gi|218256638|ref|ZP_03474966.1|, hypothetical protein PRABACTJOHN_00621 [Parabacteroides johnsonii DSM 18315]
gi|256840409|ref|ZP_05545917.1|, CRISPR-associated protein [Parabacteroides sp. D13]
gi|224535382|ref|ZP_03676371.1|, hypothetical protein BACCELL_00696 [Bacteroides cellulosilyticus DSM 14838]
gi|150683389|ref|YP_213533.1|, hypothetical protein BF3954 [Bacteroides fragilis NCTC 9343]
gi|323344874|ref|ZP_08085098.1|, csn1 family CRISPR-associated protein [Prevotella oralis ATCC 33269]
gi|213962376|ref|ZP_03390639.1|, crispr-associated protein, Csn1 family [Capnocytophaga sputigena Capno]
gi|256819408|ref|YP_003140687.1|, CRISPR-associated protein, Csn1 family [Capnocytophaga ochracea DSM 7271]
gi|298373376|ref|ZP_06983365.1|, CRISPR-associated protein, Csn1 family [Bacteroidetes oral taxon 274 str. F0058]
gi|340622236|ref|YP_004740688.1|, hypothetical protein Ccan_14650 [Capnocytophaga canimorsus Cc5]
gi|163754820|ref|ZP_02161941.1|, hypothetical protein KAOT1_02362 [Kordia algicida OT-1]
gi|295136244|ref|YP_003586920.1|, hypothetical protein ZPR_4422 [Zunongwangia profunda SM-A87]
gi|150025575|ref|YP_001296401.1|, CRISPR-associated endonuclease Csn1 family protein [Flavobacterium psychrophilum JIP02/86]
gi|212694363|ref|ZP_03302491.1|, hypothetical protein BACDOR_03889 [Bacteroides dorei DSM 17855]
gi|301311869|ref|ZP_07217791.1|, conserved hypothetical protein [Bacteroides sp. 20_3]
gi|224026357|ref|ZP_03644723.1|, hypothetical protein BACCOPRO_03113 [Bacteroides coprophilus DSM 18228]
gi|261414553|ref|YP_003248236.1|, CRISPR-associated protein, Csn1 family [Fibrobacter succinogenes subsp. succinogenes S85]
gi|260592128|ref|ZP_05857586.1|, CRISPR-associated protein, Csn1 family [Prevotella veroralis F0319]
gi|288802595|ref|ZP_06408034.1|, conserved hypothetical protein [Prevotella melaninogenica D18]
gi|315507525|ref|ZP_07882520.1|, csn1 family CRISPR-associated protein [Prevotella buccae ATCC 33574]
gi|282858617|ref|ZP_06267779.1|, CRISPR-associated protein, Csn1 family [Prevotella bivia JCVIHMP010]
gi|300771242|ref|ZP_07081118.1|, csn1 family CRISPR-associated protein [Sphingobacterium spiritivorum ATCC 33861]

FIG. 5B

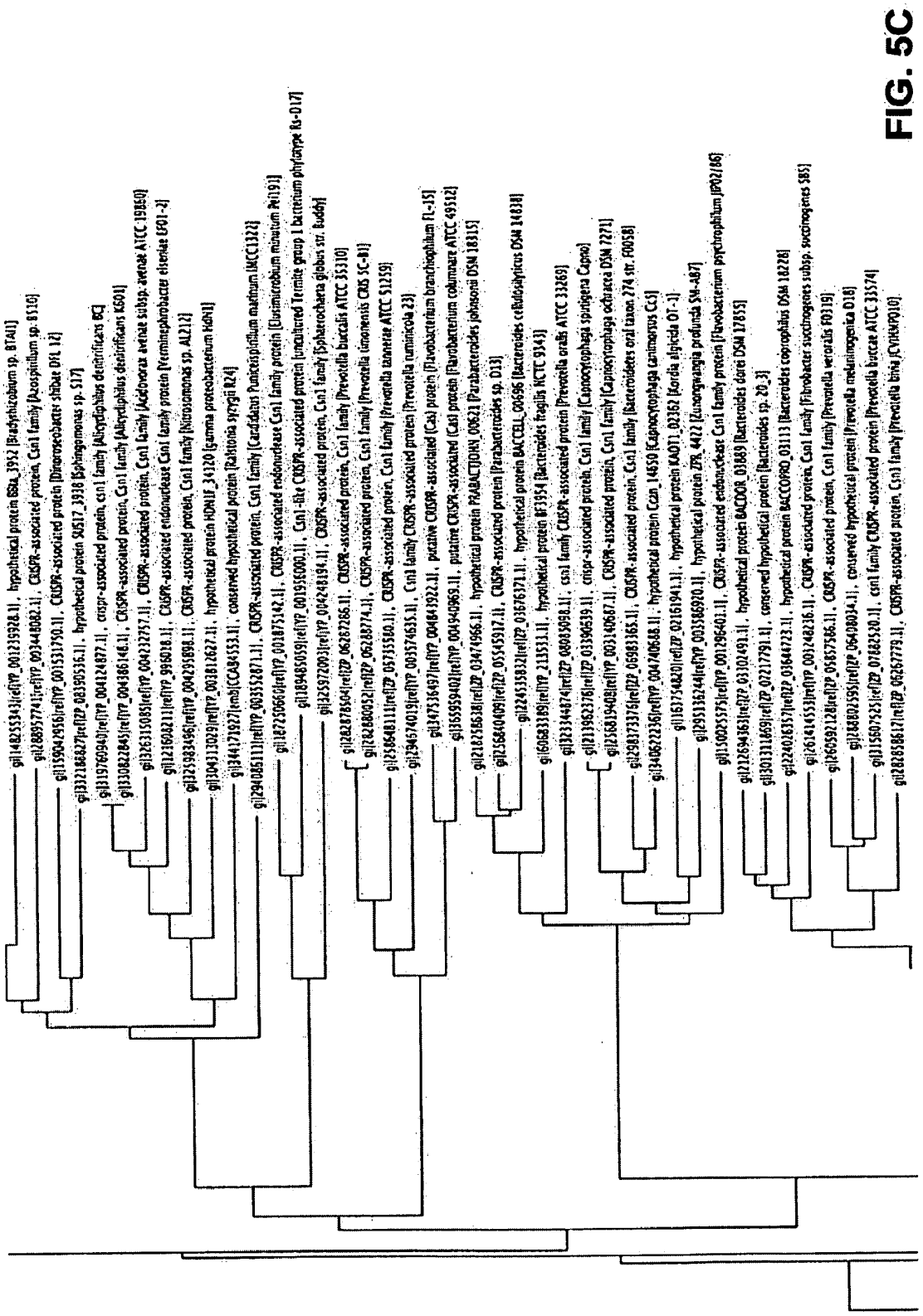


FIG. 5C

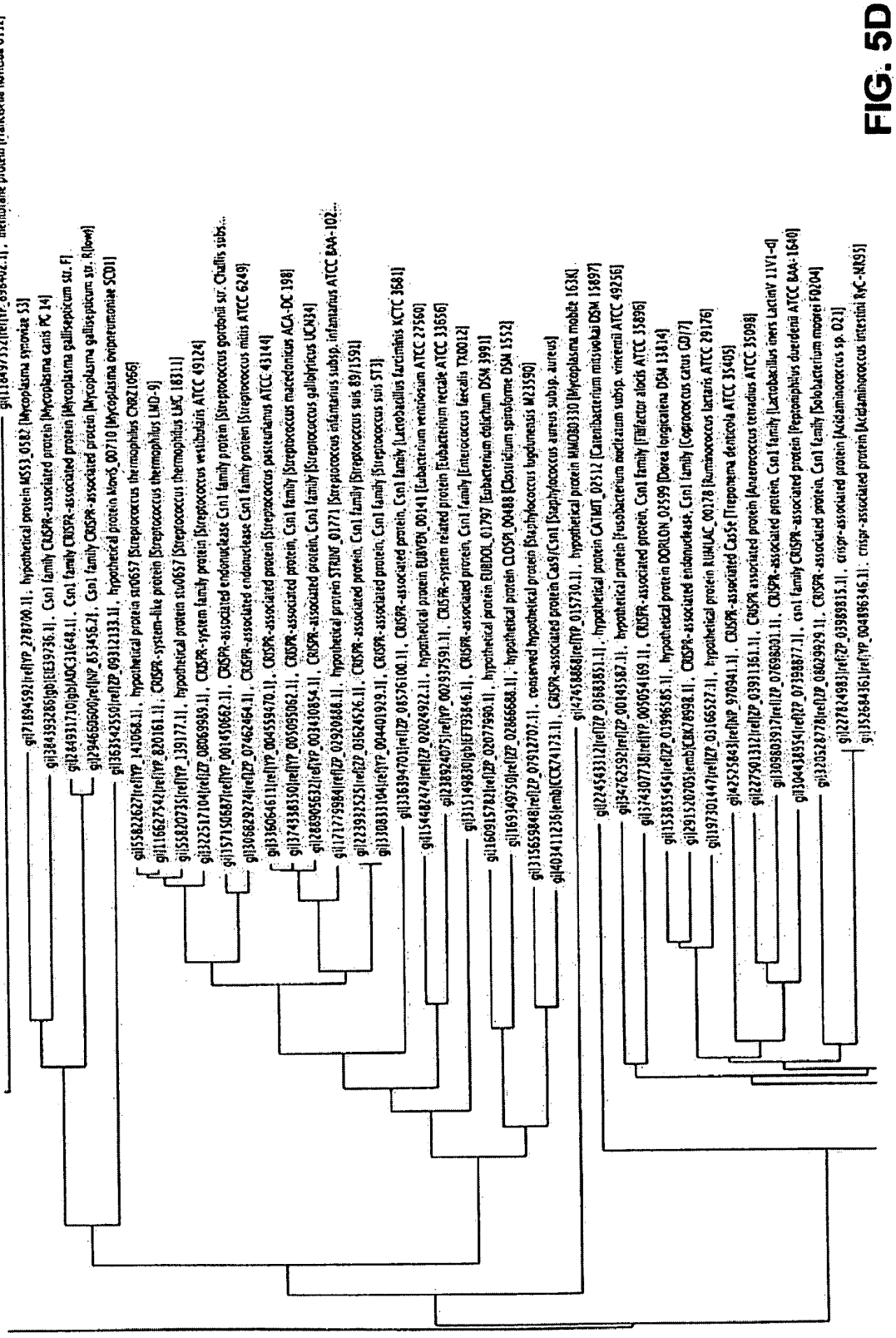

gi|118497352|ref|YP_898402.1|, membrane protein [Francisella novicida U112]
gi|71894592|ref|YP_278700.1|, hypothetical protein MS53_0582 [Mycoplasma synoviae 53]
gi|384393286|gb|EIE39736.1|, Csn1 family CRISPR-associated protein [Mycoplasma canis PG 14]
gi|284931710|gb|ADC31648.1|, Csn1 family CRISPR-associated protein [Mycoplasma gallisepticum str. F]
gi|294668690|ref|YP_853456.2|, Csn1 family CRISPR-associated protein [Mycoplasma gallisepticum str. R(low)]
gi|363542550|ref|ZP_09312133.1|, hypothetical protein MvanS_00710 [Mycoplasma ovipneumoniae SC01]
gi|55822627|ref|YP_141068.1|, hypothetical protein str0657 [Streptococcus thermophilus CNRZ1066]
gi|111662742|ref|YP_820161.1|, CRISPR-system-like protein [Streptococcus thermophilus LMD-9]
gi|55820735|ref|YP_139177.1|, hypothetical protein str0657 [Streptococcus thermophilus LMG 18311]
gi|322517104|ref|ZP_08069989.1|, CRISPR-system family protein [Streptococcus vestibularis ATCC 49124]
gi|315150687|ref|YP_001450662.1|, CRISPR-associated endonuclease Csn1 family protein [Streptococcus gordonii str. Challis subs...
gi|306829274|ref|ZP_07462464.1|, CRISPR-associated endonuclease Csn1 family protein [Streptococcus mitis ATCC 6249]
gi|336064611|ref|YP_004559470.1|, CRISPR-associated protein [Streptococcus pasteurianus ATCC 43144]
gi|374338350|ref|YP_005093062.1|, CRISPR-associated protein, Csn1 family [Streptococcus macedonicus ACA-DC 198]
gi|288905632|ref|YP_003430854.1|, CRISPR-associated protein, Csn1 family [Streptococcus gallolyticus UCN34]
gi|171779984|ref|ZP_02920468.1|, hypothetical protein STRINF_01771 [Streptococcus infantarius subsp. infantarius ATCC BAA-102...
gi|223933525|ref|ZP_03624526.1|, CRISPR-associated protein, Csn1 family [Streptococcus suis 89/1591]
gi|330831104|ref|YP_004401929.1|, CRISPR-associated protein, Csn1 family [Streptococcus suis ST3]
gi|313639470|ref|ZP_05376100.11, CRISPR-associated protein, Csn1 family [Lactobacillus farciminis KCTC 3681]
gi|154482474|ref|ZP_02024922.1|, hypothetical protein EUBVEN_00141 [Eubacterium ventriosum ATCC 27560]
gi|238924075|ref|YP_002937591.1|, CRISPR-system related protein [Eubacterium rectale ATCC 33656]
gi|315149838|gb|EFT93846.1|, CRISPR-associated protein, Csn1 family [Enterococcus faecalis TX0012]
gi|160915782|ref|ZP_02077990.1|, hypothetical protein EUBDOL_01797 [Eubacterium dolichum DSM 3991]
gi|169349750|ref|ZP_02866688.1|, hypothetical protein CLOSPI_00488 [Clostridium spiroforme DSM 1552]
gi|315659848|ref|ZP_07912707.1|, conserved hypothetical protein [Staphylococcus lugdunensis M23590]
gi|403411236|emb|CCK74173.1|, CRISPR-associated protein Cas9/Csn1 [Staphylococcus aureus subsp. aureus]
gi|474518868|ref|YP_015710.1|, hypothetical protein MMOB0330 [Mycoplasma mobile 163K]
gi|224543312|ref|ZP_03683851.1|, hypothetical protein CATMIT_02512 [Catenibacterium mitsuokai DSM 15897]
gi|347629592|ref|ZP_00149387.1|, hypothetical protein Fisuba_0178 [Fusobacterium mediatum subsp. vincentii ATCC 49256]
gi|187430773|ref|YP_005054169.1|, CRISPR-associated protein, Csn1 family [Filifactor alocis ATCC 35896]
gi|153855454|ref|ZP_01996185.1|, hypothetical protein DORLON_02599 [Dorea longicatena DSM 13814]
gi|291520705|emb|CBK78998.1|, CRISPR-associated endonuclease, Csn1 family [Coprococcus catus GD/7]
gi|197301447|ref|ZP_03166527.1|, hypothetical protein RUMLAC_00178 [Ruminococcus lactaris ATCC 29176]
gi|42525843|ref|NP_970941.1|, CRISPR-associated Cas5e [Treponema denticola ATCC 35405]
gi|227501312|ref|ZP_03931161.1|, CRISPR-associated protein [Anaerococcus tetradius ATCC 35098]
gi|305680917|ref|ZP_07598201.1|, CRISPR-associated protein, Csn1 family [Lactobacillus iners LactinV 11V1-d]
gi|304438354|ref|ZP_07198877.1|, csn1 family CRISPR-associated protein [Peptoniphilus duerdenii ATCC BAA-1640]
gi|325127877|ref|ZP_08029929.1|, CRISPR-associated protein, Csn1 family [Sebaldella termitidis ATCC 33386]
gi|227814983|ref|ZP_03983815.1|, crispr-associated protein [Acidaminococcus sp. D21]
gi|352684316|ref|YP_004896346.1|, crispr-associated protein [Acidaminococcus intestini RyC-MR95]

FIG. 5D

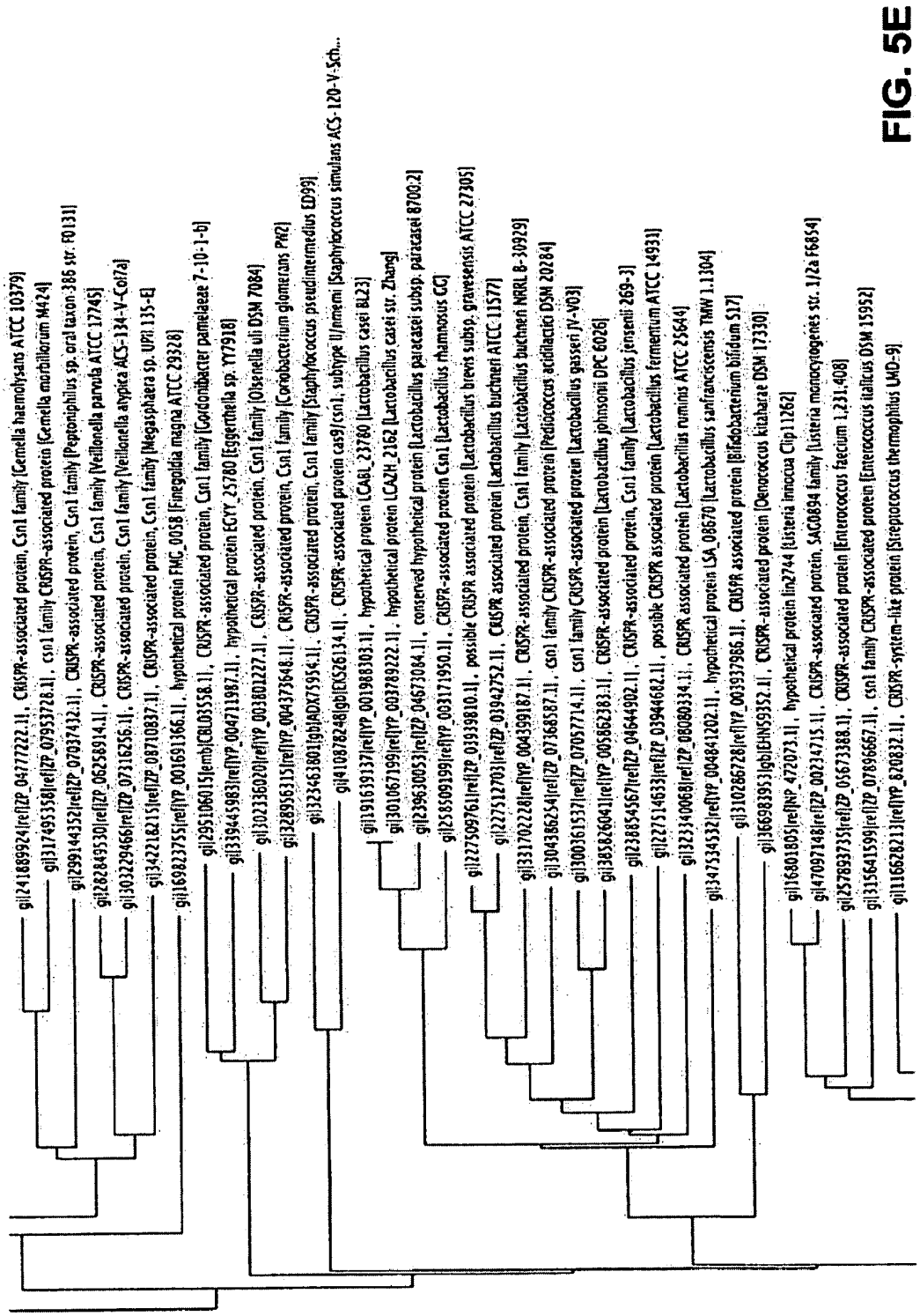

gi|241689924|ref|ZP_04777222.1|, CRISPR-associated protein, Csn1 family [Gemella haemolysans ATCC 10379]
gi|317495358|ref|ZP_07953728.1|, csn1 family CRISPR-associated protein [Gemella morbillorum M424]
gi|299144352|ref|ZP_07037432.1|, CRISPR-associated protein, Csn1 family [Peptoniphilus sp. oral taxon 386 str. F0131]
gi|262845930|ref|ZP_06258914.1|, CRISPR-associated protein, Csn1 family [Veillonella parvula ATCC 17745]
gi|303229466|ref|ZP_07316256.1|, CRISPR-associated protein, Csn1 family [Veillonella atypica ACS-134-V-Col7a]
gi|342218215|ref|ZP_08710837.1|, CRISPR-associated protein, Csn1 family [Megasphaera sp. UPII 135-E]
gi|169823755|ref|YP_001691366.1|, hypothetical protein FMG_0058 [Finegoldia magna ATCC 29328]
gi|295106915|emb|CBL03558.1|, CRISPR-associated protein, Csn1 family [Gordonibacter pamelaeae 7-10-1-b]
gi|339445983|ref|YP_004711987.1|, hypothetical protein EGYY_25780 [Eggerthella sp. YY7918]
gi|302336020|ref|YP_003802227.1|, CRISPR-associated protein, Csn1 family [Olsenella uli DSM 7084]
gi|328956315|ref|YP_004373648.1|, CRISPR-associated protein, Csn1 family [Coriobacterium glomerans PW2]
gi|323463801|gb|ADX75954.1|, CRISPR-associated protein, Csn1 family [Staphylococcus pseudintermedius ED99]
gi|410878248|gb|EKS26134.1|, CRISPR-associated protein cas9/csn1, subtype II/nmeni [Staphylococcus simulans ACS-120-V-Sch...
gi|319163137|ref|YP_001988383.1|, hypothetical protein LCABL_23780 [Lactobacillus casei BL23]
gi|301067199|ref|YP_003789222.1|, hypothetical protein LCAZH_2162 [Lactobacillus casei str. Zhang]
gi|239630053|ref|ZP_04673084.1|, conserved hypothetical protein [Lactobacillus paracasei subsp. paracasei 8700:2]
gi|258509199|ref|YP_003171950.1|, CRISPR-associated protein Csn1 [Lactobacillus rhamnosus GG]
gi|227509761|ref|ZP_03939810.1|, possible CRISPR associated protein [Lactobacillus brevis subsp. gravesensis ATCC 27305]
gi|227512703|ref|ZP_03942752.1|, CRISPR associated protein [Lactobacillus buchneri ATCC 11577]
gi|331702228|ref|YP_004399187.1|, CRISPR-associated protein, Csn1 family [Lactobacillus buchneri NRRL B-30929]
gi|304386254|ref|ZP_07368587.1|, csn1 family CRISPR-associated protein [Pediococcus acidilactici DSM 20284]
gi|300361537|ref|ZP_07057714.1|, csn1 family CRISPR-associated protein [Lactobacillus gasseri JV-V03]
gi|385826041|ref|YP_005862383.1|, CRISPR-associated protein [Lactobacillus johnsonii DPC 6026]
gi|238854567|ref|ZP_04644902.1|, CRISPR-associated protein, Csn1 family [Lactobacillus jensenii 269-3]
gi|227514633|ref|ZP_03944682.1|, possible CRISPR associated protein [Lactobacillus fermentum ATCC 14931]
gi|323340068|ref|ZP_08080334.1|, CRISPR associated protein [Lactobacillus ruminis ATCC 25644]
gi|347534532|ref|YP_004841202.1|, hypothetical protein LSA_08670 [Lactobacillus sanfranciscensis TMW 1.1304]
gi|310286728|ref|YP_003937986.1|, CRISPR associated protein [Bifidobacterium bifidum S17]
gi|366983953|gb|EHN59352.1|, CRISPR-associated protein [Oenococcus kitaharae DSM 17330]
gi|116801805|ref|NP_472073.1|, hypothetical protein lin2744 [Listeria innocua Clip11262]
gi|470971481|ref|ZP_00234715.1|, CRISPR-associated protein, SAG0894 family [Listeria monocytogenes str. 1/2a F6854]
gi|257893735|ref|ZP_05673388.1|, CRISPR-associated protein [Enterococcus faecium 1,231.408]
gi|315641599|ref|ZP_07896667.1|, csn1 family CRISPR-associated protein [Enterococcus italicus DSM 15952]
gi|116628213|ref|YP_820832.1|, CRISPR-system-like protein [Streptococcus thermophilus LMD-9]

FIG. 5E

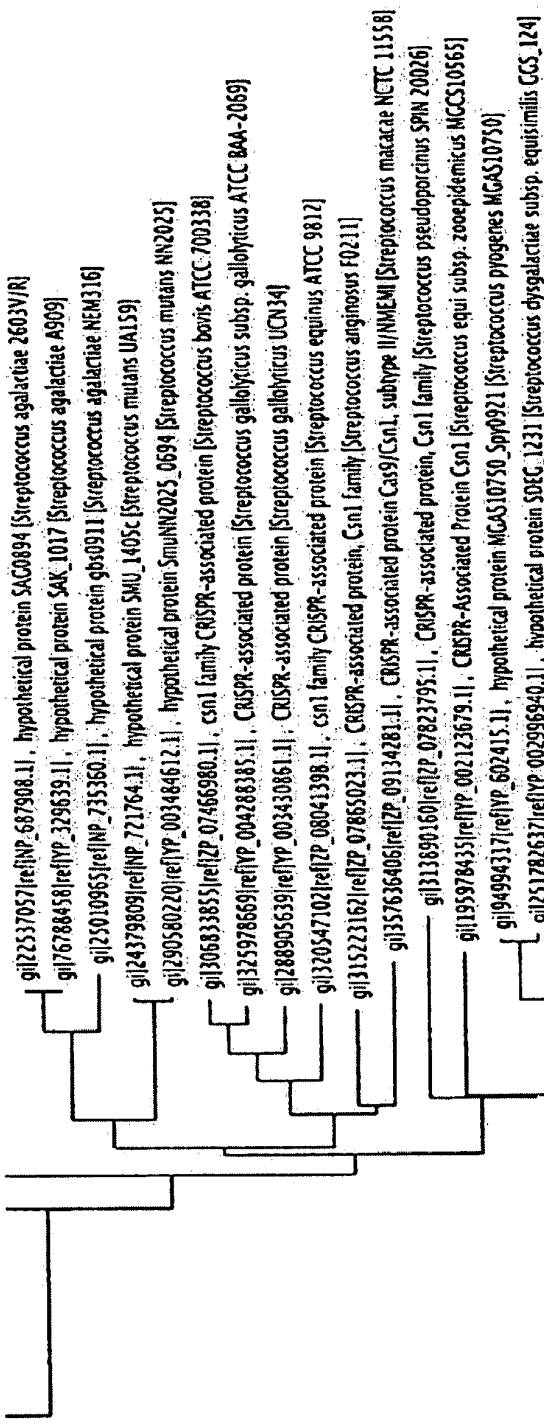

gi|22537057|ref|NP_687908.1|, hypothetical protein SAG0894 [Streptococcus agalactiae 2603V/R]
gi|76788458|ref|YP_329639.1|, hypothetical protein SAK_1017 [Streptococcus agalactiae A909]
gi|25010965|ref|NP_735360.1|, hypothetical protein gbs0911 [Streptococcus agalactiae NEM316]
gi|24379809|ref|NP_721764.1|, hypothetical protein SMU_1405c [Streptococcus mutans UA159]
gi|29058022O|ref|YP_003484612.1|, hypothetical protein SmuNN2025_0694 [Streptococcus mutans NN2025]
gi|30683385|ref|ZP_07466980.1|, csn1 family CRISPR-associated protein [Streptococcus bovis ATCC 700338]
gi|32597869|ref|YP_004288385.1|, CRISPR-associated protein [Streptococcus gallolyticus subsp. gallolyticus ATCC BAA-2069]
gi|28890563|ref|YP_003430861.1|, CRISPR-associated protein [Streptococcus gallolyticus UCN34]
gi|32054702|ref|ZP_08041398.1|, csn1 family CRISPR-associated protein [Streptococcus equinus ATCC 9812]
gi|31522316|ref|ZP_07865023.1|, CRISPR-associated protein, Csn1 family [Streptococcus anginosus F0211]
gi|35763640|ref|ZP_09134281.1|, CRISPR-associated protein Cas9/Csn1, subtype II/NMEMI [Streptococcus mutacae NCTC 11558]
gi|31189016O|ref|ZP_07823795.1|, CRISPR-associated protein, Csn1 family [Streptococcus pseudoporcinus SPIN 20026]
gi|19597843S|ref|YP_002123679.1|, CRISPR-Associated Protein Csn1 [Streptococcus equi subsp. zooepidemicus MGCS10565]
gi|94994317|ref|YP_602415.1|, hypothetical protein MGAS10750_Spy0921 [Streptococcus pyogenes MGAS10750]
gi|25178263 7|ref|YP_002996940.1|, hypothetical protein SDEG_1231 [Streptococcus dysgalactiae subsp. equisimilis GGS_124]
gi|21910213|ref|NP_664481.1|, hypothetical protein SpyM3_0677 [Streptococcus pyogenes MGAS315]
gi|28896088|ref|NP_802438.1|, hypothetical protein SPs1176 [Streptococcus pyogenes SSI-1]
gi|20955935S6|ref|YP_002285828.1|, hypothetical protein Spy49_0823 [Streptococcus pyogenes NZ131]
gi|15675041|ref|NP_269215.1|, hypothetical protein SPy_1046 [Streptococcus pyogenes M1 GAS]
gi|71910582|ref|YP_282132.1|, hypothetical protein MS005_Spy_0768 [Streptococcus pyogenes MGAS5005]
gi|94990395|ref|YP_598495.1|, putative cytoplasmic protein [Streptococcus pyogenes MGAS10270]
gi|94988516|ref|YP_596617.1|, cytoplasmic protein [Streptococcus pyogenes MGAS9429]
gi|94992340|ref|YP_600439.1|, putative cytoplasmic protein [Streptococcus pyogenes MGAS2096]
gi|71903413|ref|YP_280216.1|, cytoplasmic protein [Streptococcus pyogenes MGAS6180]

FIG. 5F

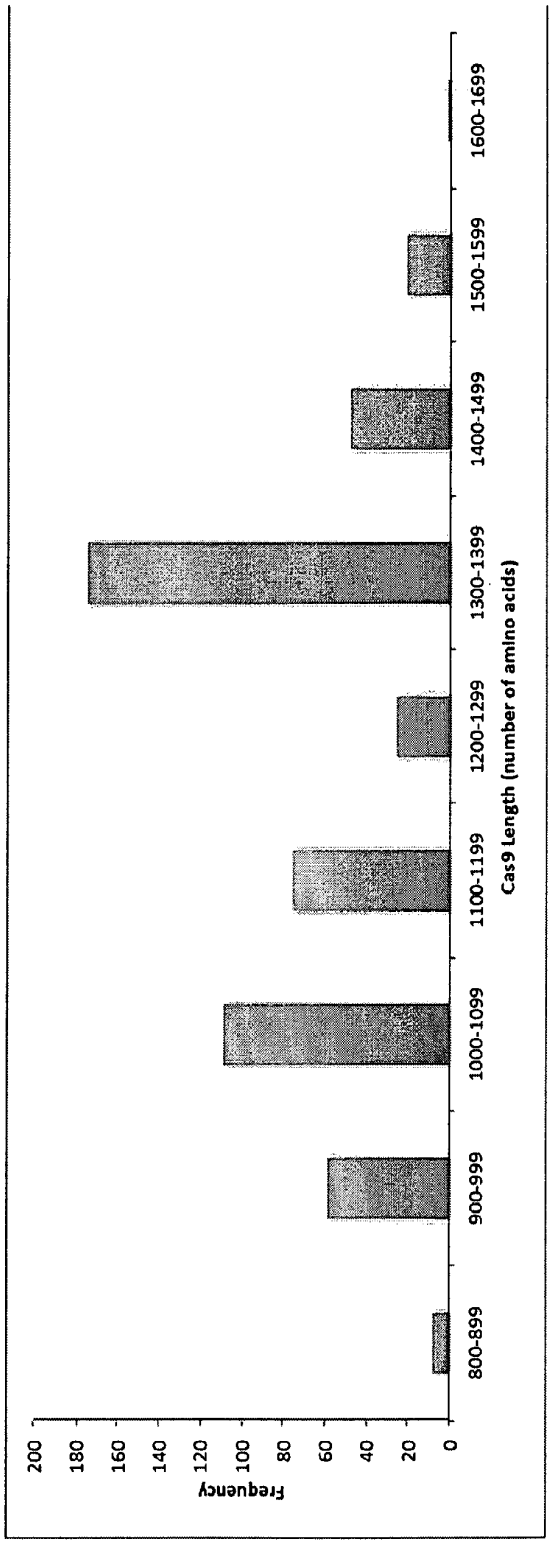

Frequency
0
20
40
60
80
100
120
140
160
180
200
800-899
900-999
1000-1099
1100-1199
1200-1299
1300-1399
1400-1499
1500-1599
1600-1699
Cas9 Length (number of amino acids)

FIG. 6

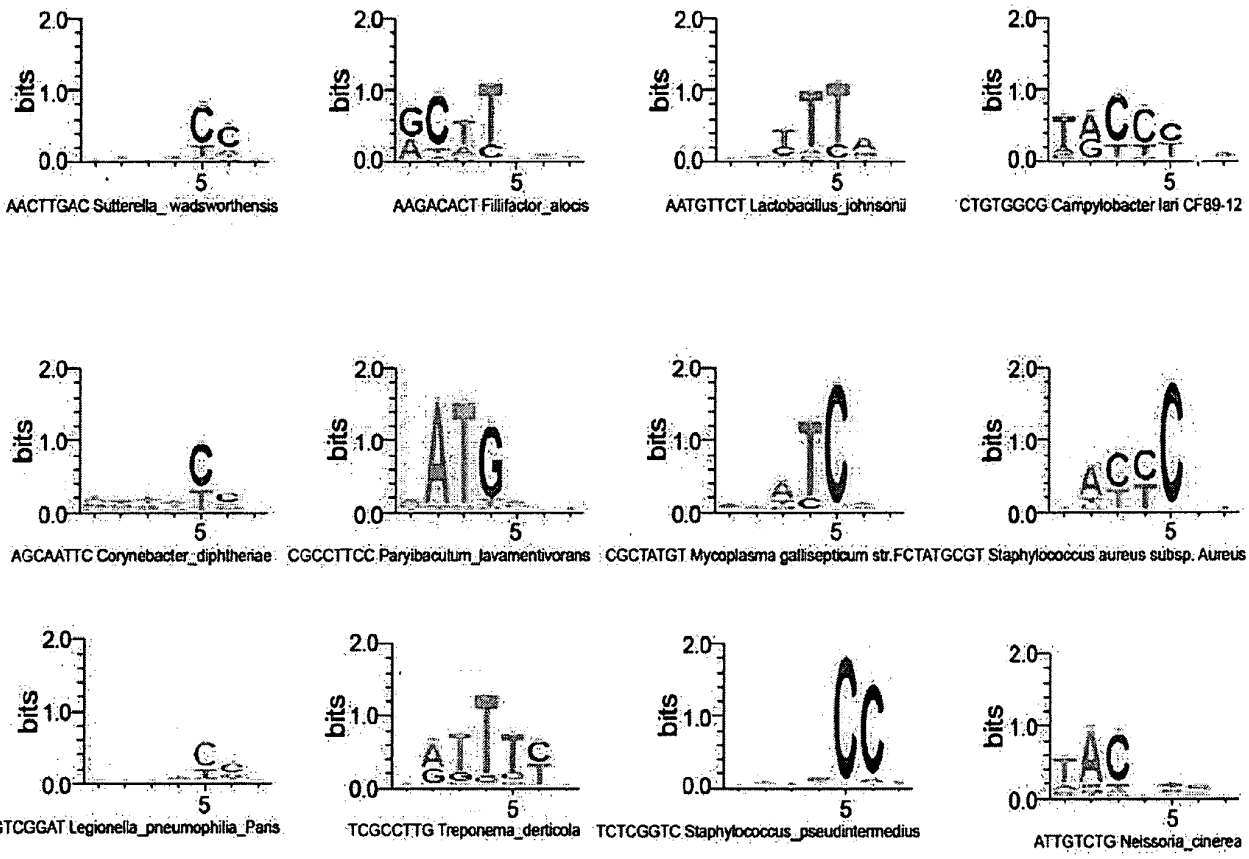
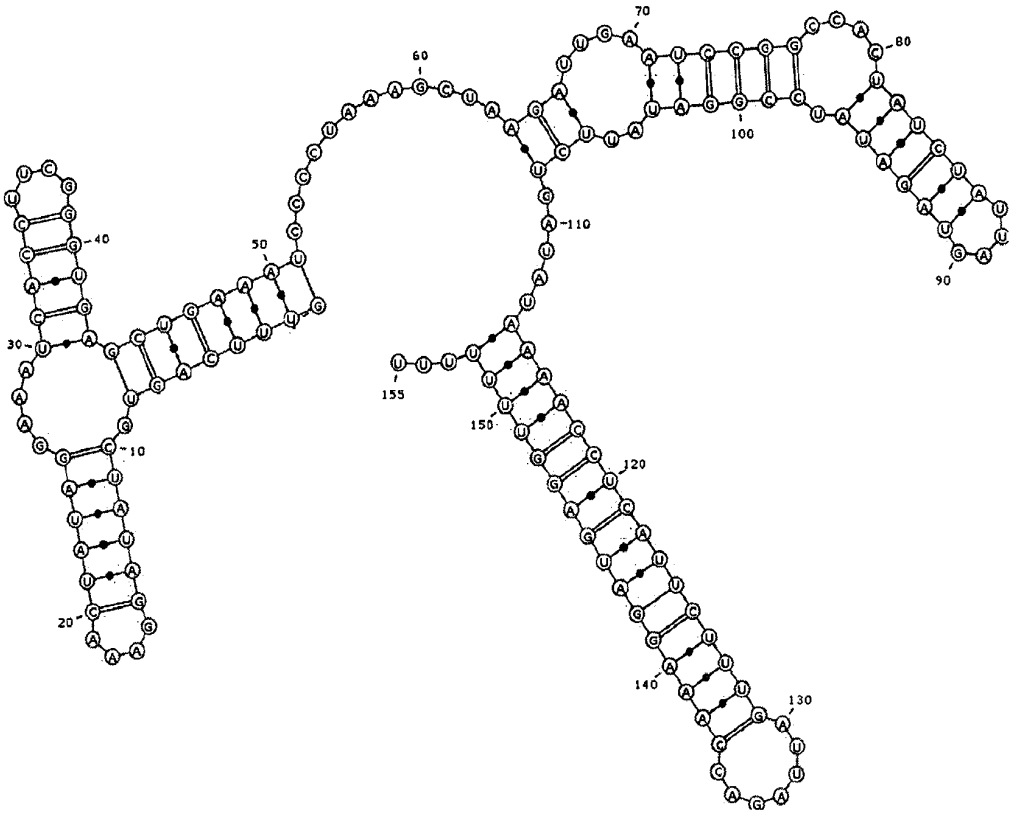
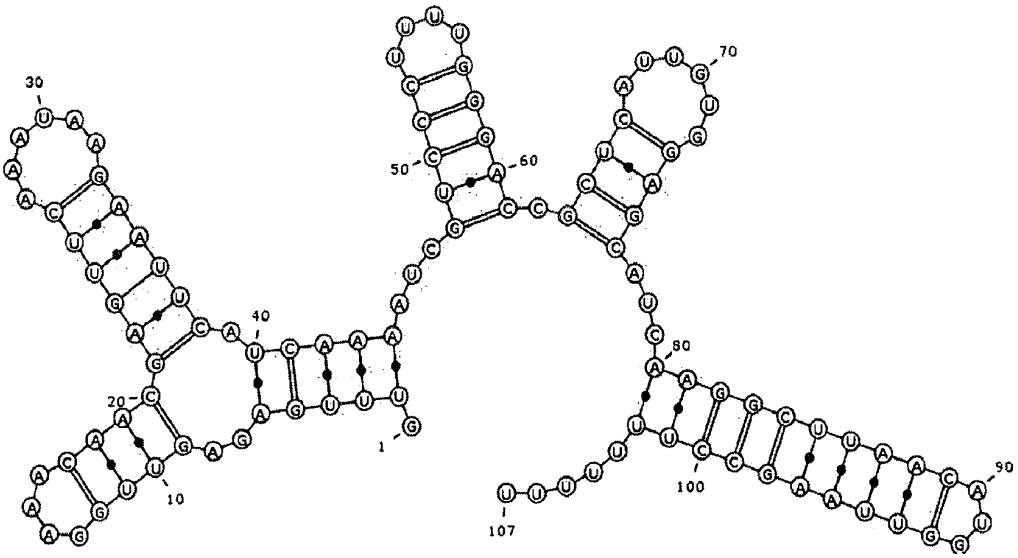

bits
2.0
1.0
0.0
5
AACTTGAC Sutterella_wadsworthensis
AAGACACT Filifactor_alocis
AATGTTCT Lactobacillus_johnsonii
CTGTGGCG Campylobacter lari CF89-12
AGCAATTC Corynebacter_diphtheriae
CGCCTTCC Paryibaculum_lavamentivorans
CGCTATGT Mycoplasma gallisepticum str.FCTATGCGT Staphylococcus aureus subsp. Aureus
TGTCGGAT Legionella_pneumophilia_Paris
TCGCCTTG Treponema_denticola
TCTCGGTC Staphylococcus_pseudintermedius
ATTGTCTG Neissoria_cinerea

FIG. 7

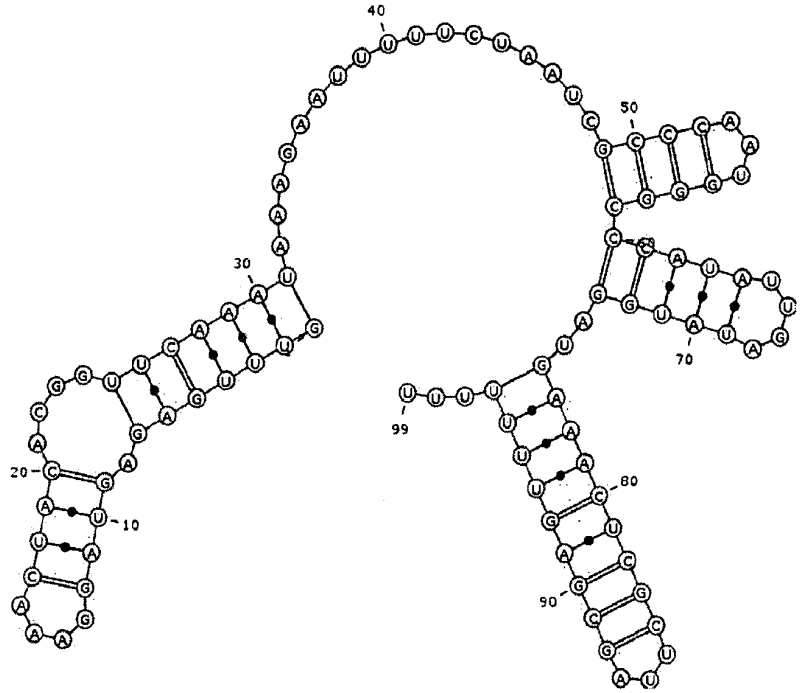


Chimeric guide RNA for Cas9 homolog 2

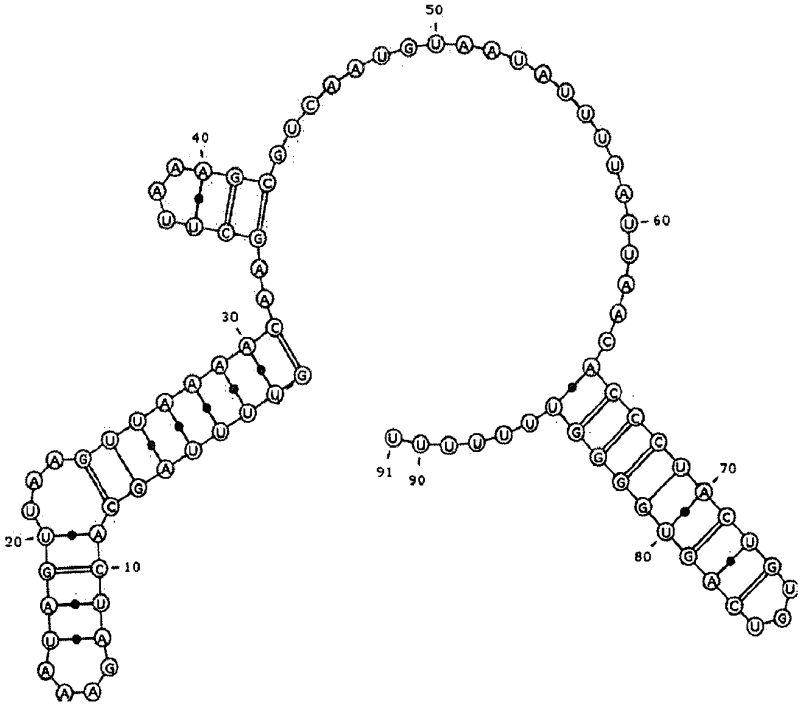


Chimeric guide RNA for Cas9 homolog 4

FIG. 8A

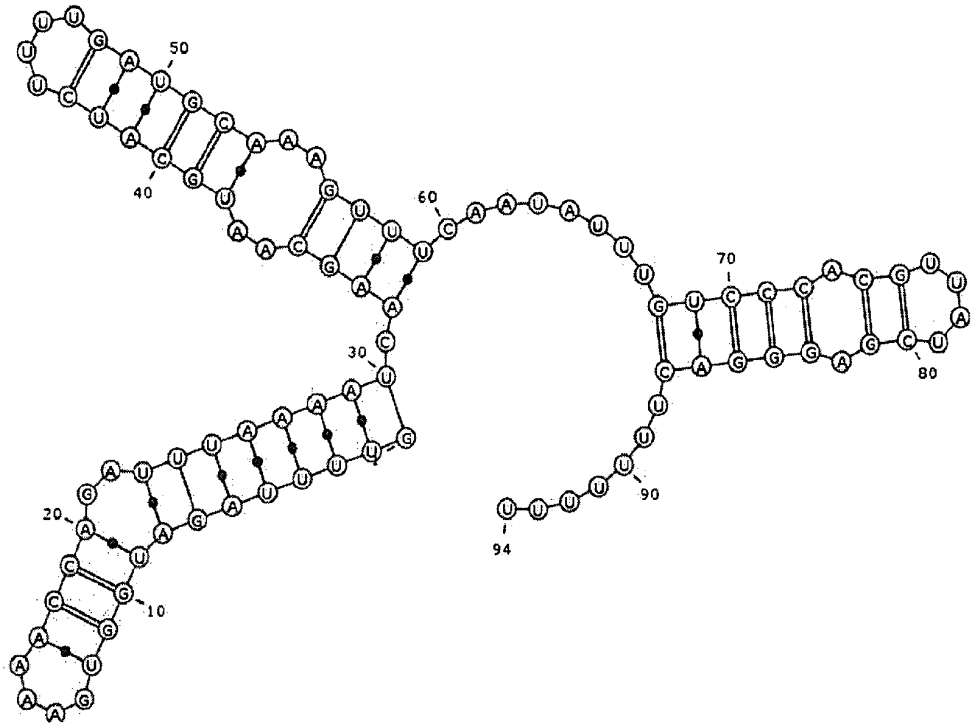


Chimeric guide RNA for Cas9 homolog 5

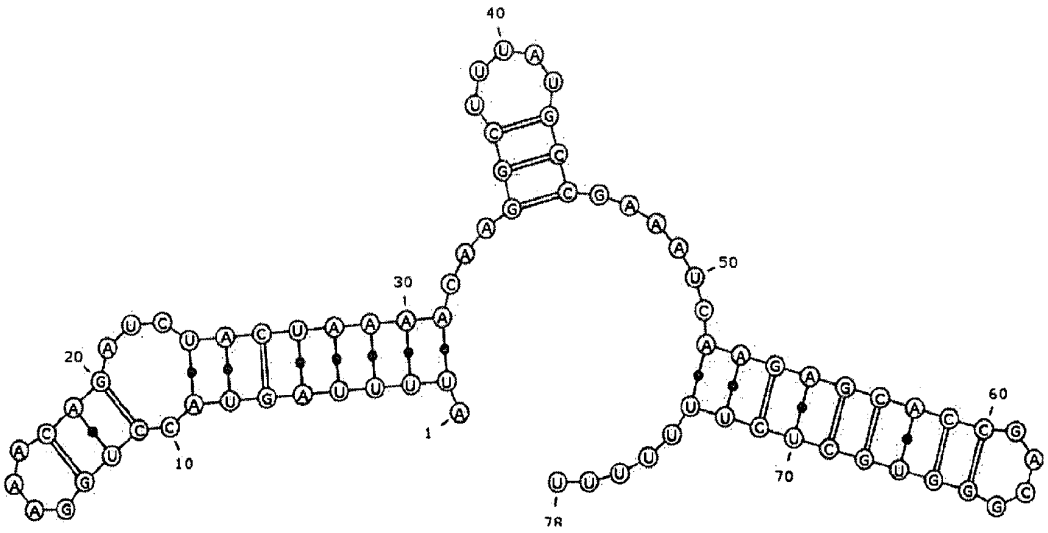


Chimeric guide RNA for Cas9 homolog 6

FIG. 8B

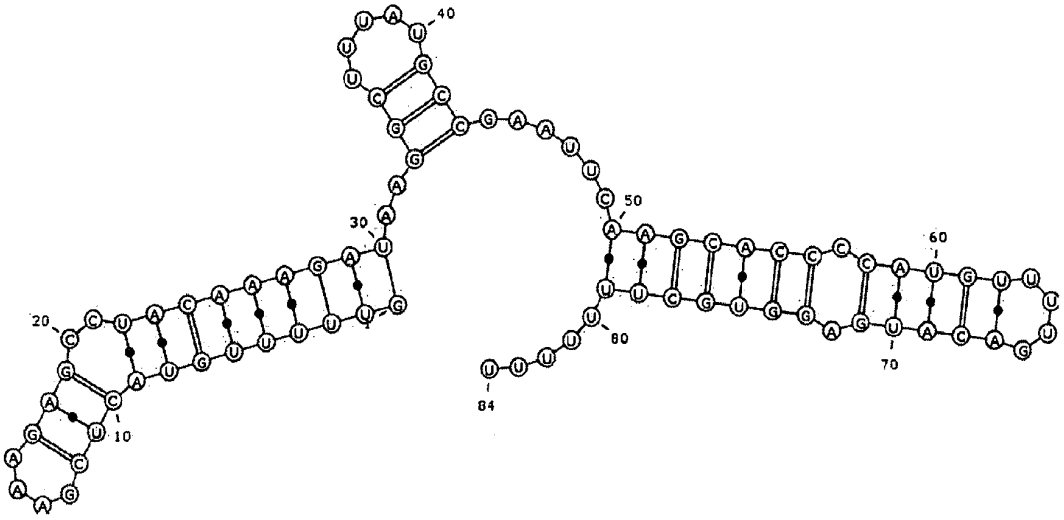


Chimeric guide RNA for Cas9 homolog 7

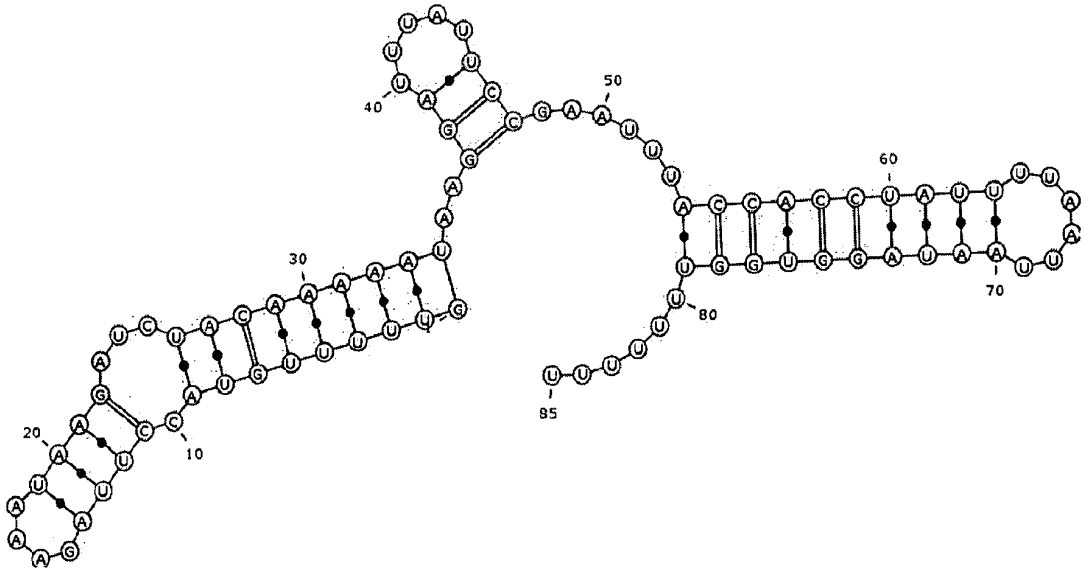


Chimeric guide RNA for Cas9 homolog 8

FIG. 8C

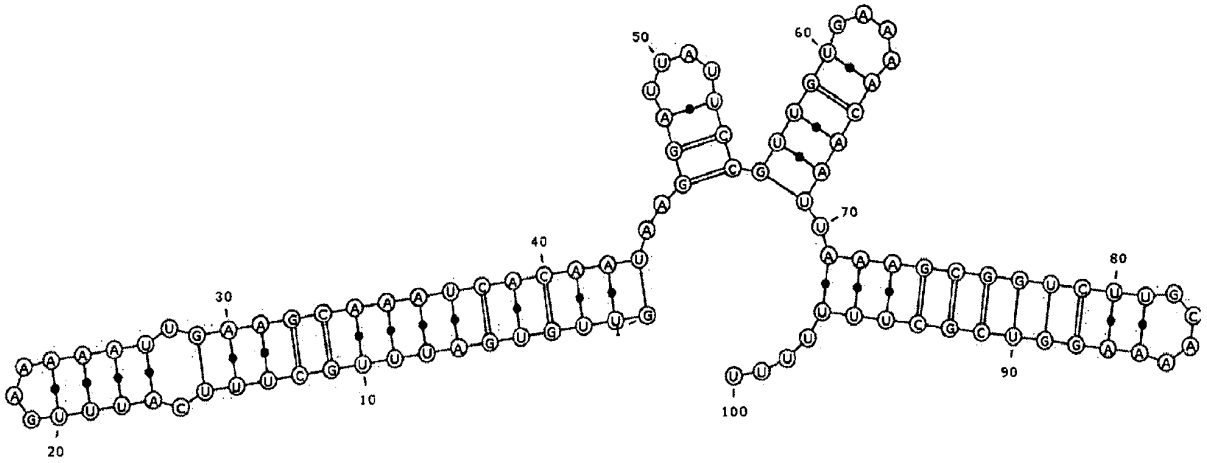


Chimeric guide RNA for Cas9 homolog 9

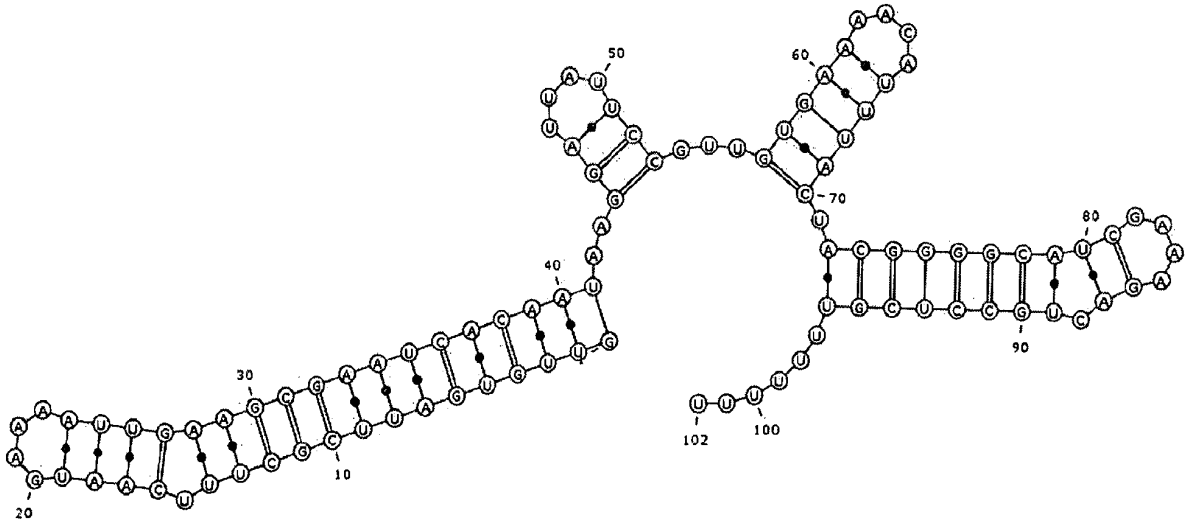


Chimeric guide RNA for Cas9 homolog 10

FIG. 8D

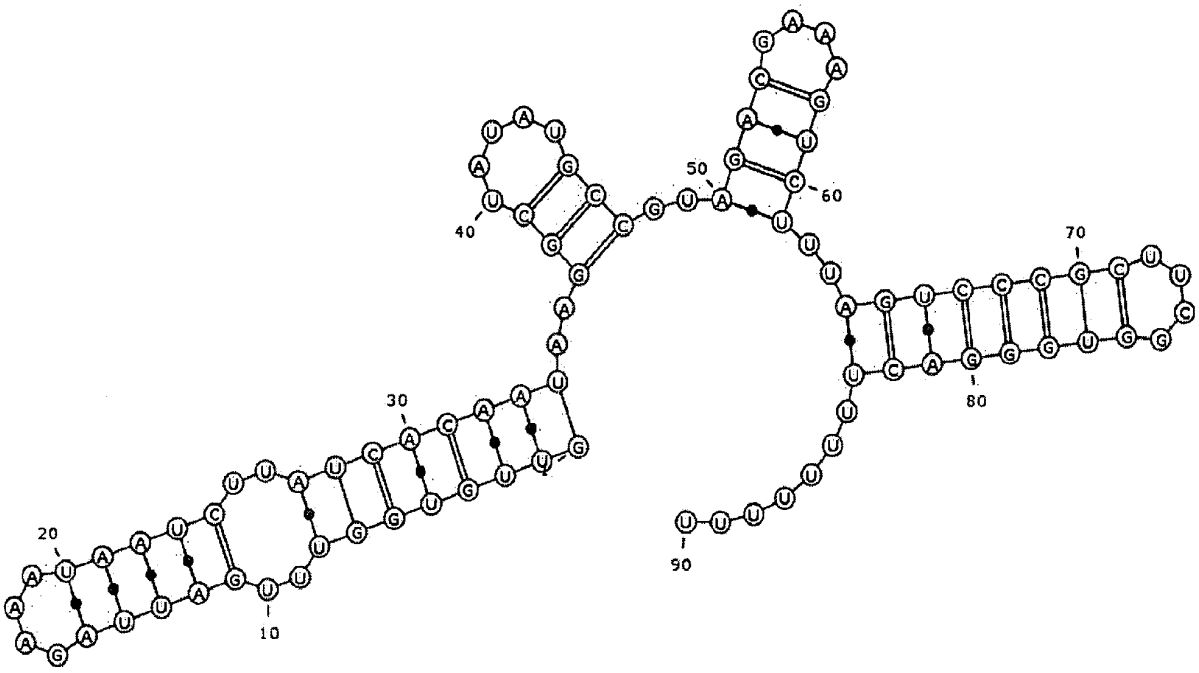


Chimeric guide RNA for Cas9 homolog 12



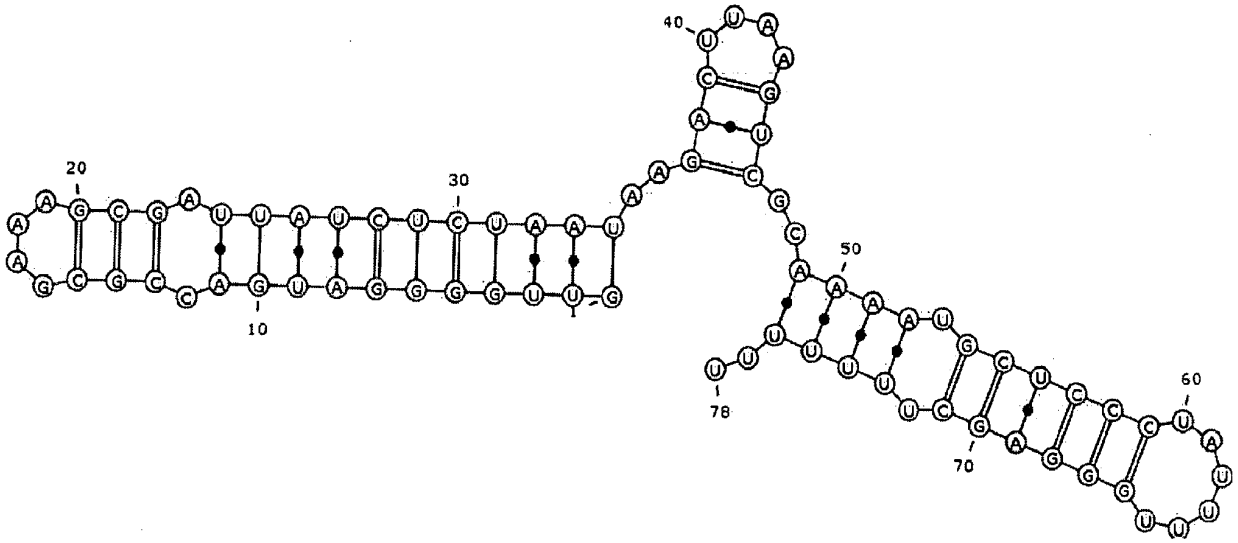
Chimeric guide RNA for Cas9 homolog 13

FIG. 8E

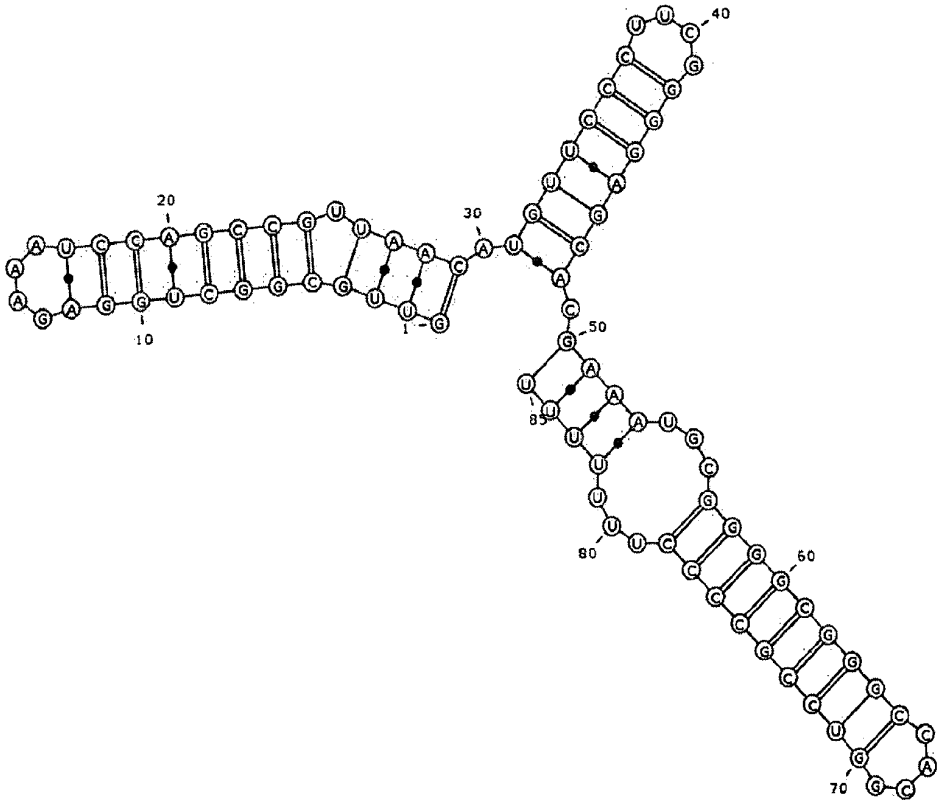


Chimeric guide RNA for Cas9 homolog 14

FIG. 8F

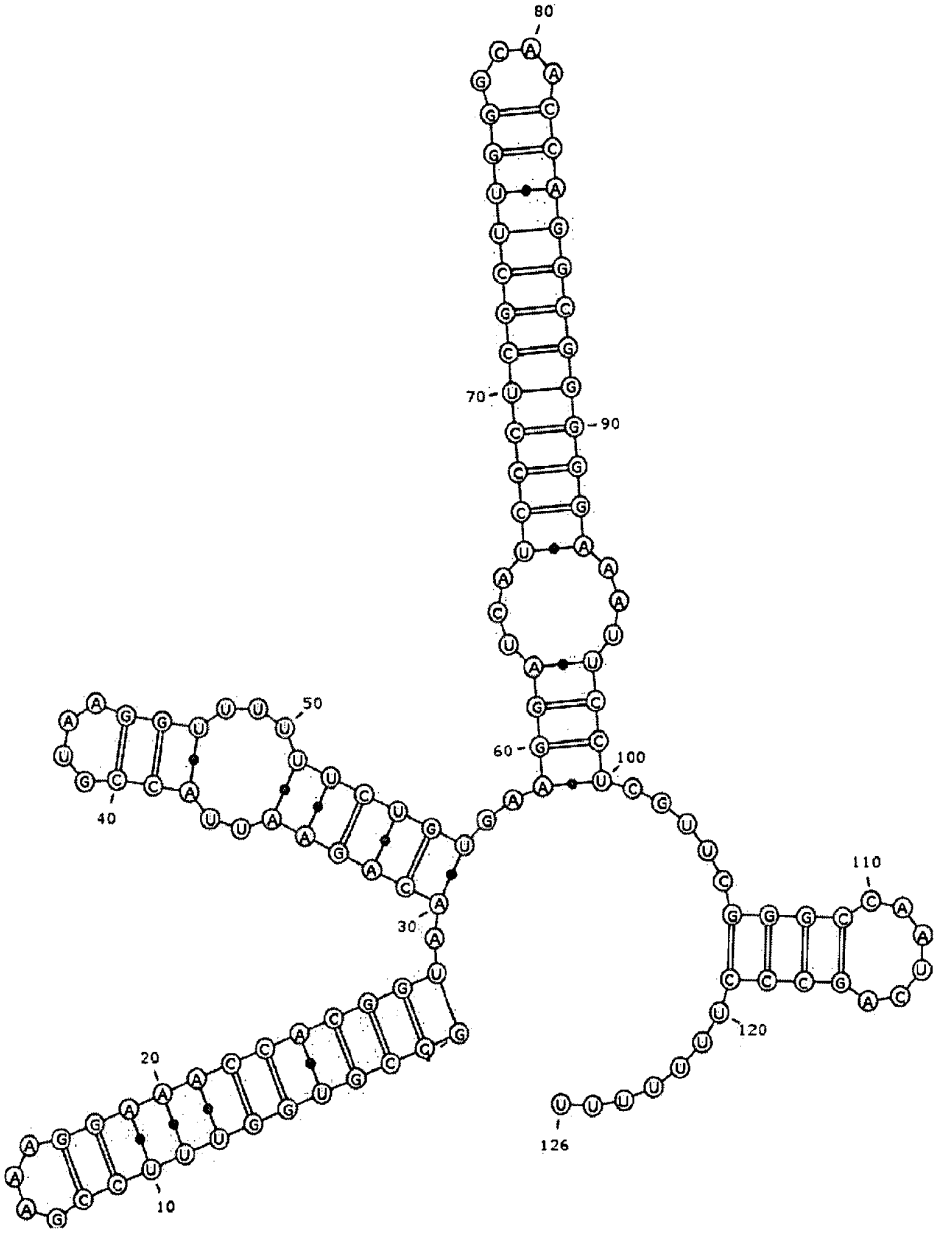


Chimeric guide RNA for Cas9 homolog 15



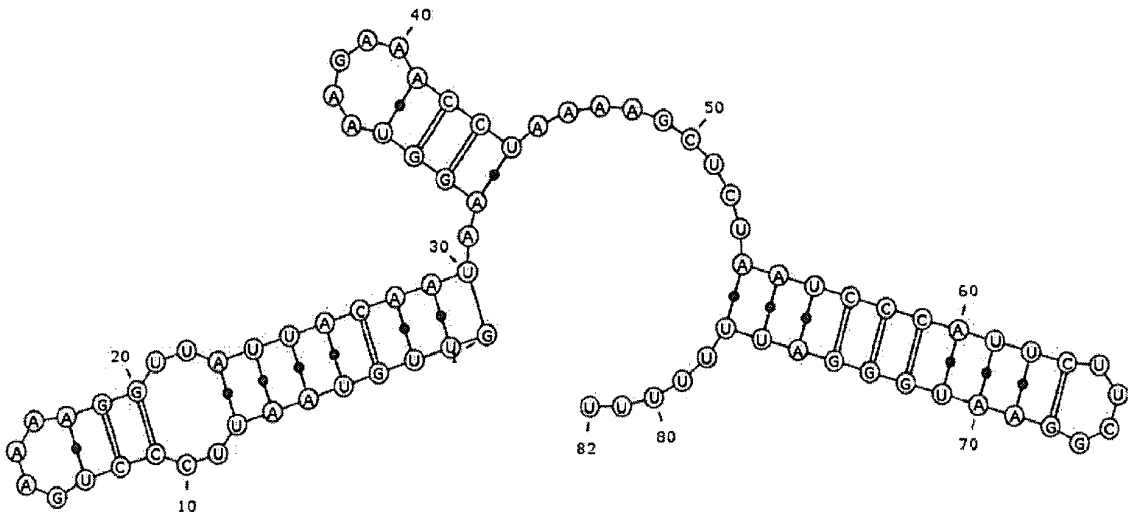
Chimeric guide RNA for Cas9 homolog 16

FIG. 8G



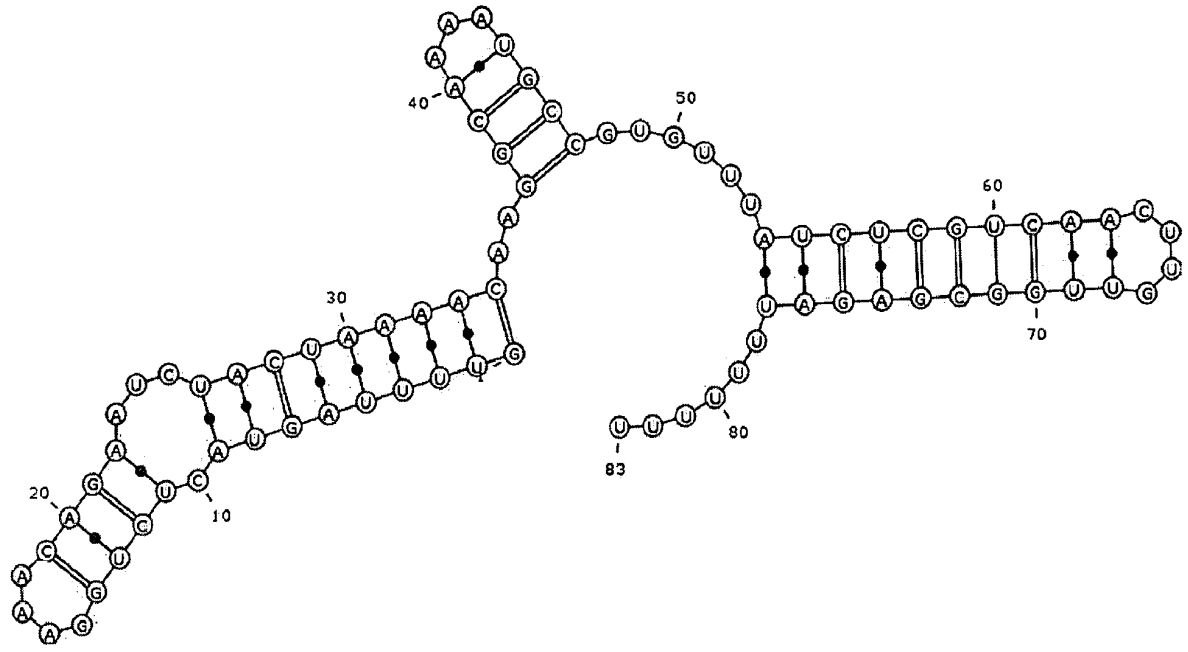
Chimeric guide RNA for Cas9 homolog 17

FIG. 8H

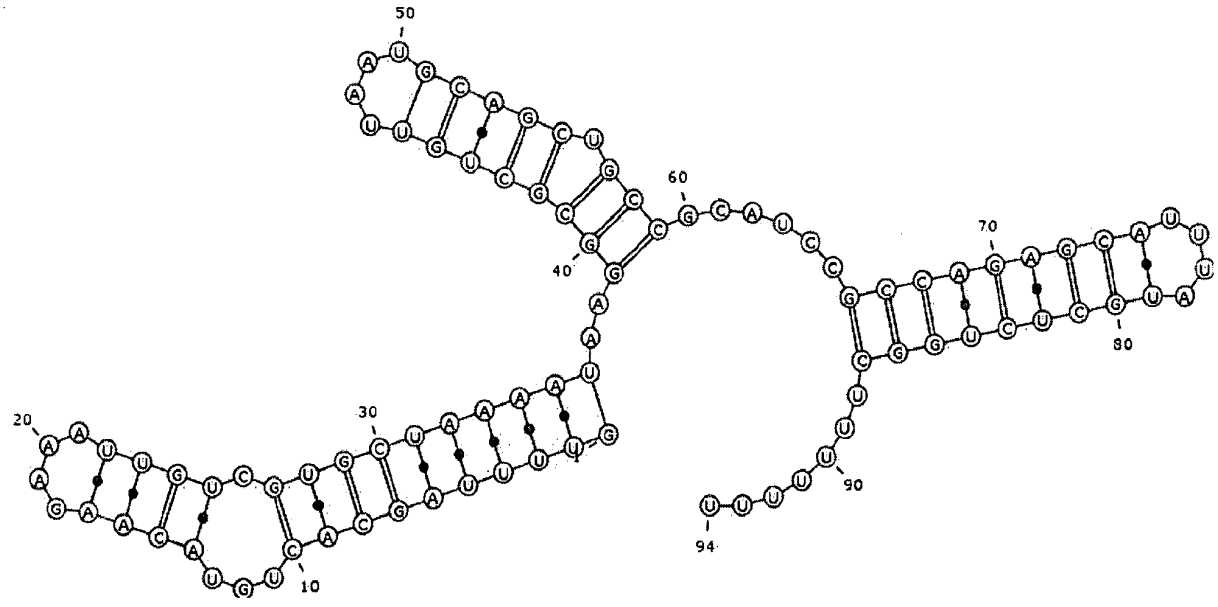


Chimeric guide RNA for Cas9 homolog 19

FIG. 8I



Chimeric guide RNA for Cas9 homolog 21



Chimeric guide RNA for Cas9 homolog 23

FIG. 8J

FIG. 9A

human codon optimized Cas9 sequences

#2

ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGACTCAG
AGCGAGCGACGATTTTCTTGCAGCATTGGCATTGACATGGGGGCTAAGTACACTGGGGTGTTCTACGCACTGTTCGACCGGGAG
GAACTGCCCACAAACCTGAACAGCAAGGCCATGACCCTGGTCATGCCTGAGACAGGGCCAAGATACGTGCAGGCACAGAGAACT
GCCGTCAGACACAGGCTGCGCGGACAGAAGAGATATACCCTGGCTAGGAAACTGGCATTTCTGGTGGTCGACGATATGATCAAG
AAACAGGAAAAGAGGCTGACTGATGAGGAATGGAAACGAGGACGGGAGGCCCTGTCCGGCCTGCTGAAGCGGAGAGGGTACTCT
CGGCCCAACGCTGACGGCGAAGATCTGACCCCTCTGGAGAATGTGAGAGCAGACGTGTTCGCCGCTCATCCTGCCTTCAGCACA
TATTTTTCCGAAGTGCGCTCTCTGGCTGAGCAGTGGGAGGAGTTCACCGCAAACATCAGCAATGTCGAGAAGTTTCTGGGCGAC
CCAAACATCCCCGCCGATAAAGAGTTCATTGAATTTGCCGTGGCTGAAGGGCTGATTGACAAGACCGAGAAGAAAGCCTACCAG
TCAGCTCTGAGCACCCTGAGGGCAAACGCCAATGTGCTGACAGGACTGCGGCAGATGGGCCACAAGCCTAGATCAGAATATTTT
AAAGCAATCGAGGCCGACCTGAAGAAAGATAGCCGCCTGGCCAAGATTAACGAAGCATTCGGAGGAGCAGAGCGCCTGGCTCGA
CTGCTGGGAAACCTGTCCAATCTGCAGCTGCGGGCAGAAAGATGGTACTTCAATGCCCCCGACATCATGAAGGATAGGGGCTGG
GAGCCTGATCGCTTCAAGAAAACACTGGTGCGGGCTTTTAAGTTCTTTCACCCAGCAAAGGACCAGAACAAACAGCATCTGGAA
CTGATCAAACAGATTGAGAACAGCGAAGATATCATTGAGACTCTGTGCACCCTGGACCCAAACAGAACCATCCCCCCTTACGAG
GATCAGAACAATAGGCGCCCACCCCTGGACCAGACTCTGCTGCTGAGTCCCGAAAAGCTGACCCGGCAGTATGGCGAGATCTGG
AAAACATGGAGCGCCAGACTGACCTCCGCTGAACCCACACTGGCACCTGCAGCCGAGATTCTGGAAAGATCTACCGACAGGAAG
AGTCGCGTGGCAGTCAACGGACACGAGCCACTGCCTACACTGGCTTACCAGCTGAGTTATGCACTGCAGAGAGCCTTCGACAGG
TCAAAAGCCCTGGATCCATATGCTCTGAGGGCACTGGCTGCAGGCTCAAAAAGCAATAAGCTGACATCCGCCCGCACTGCTCTG
GAGAACTGCATCGGAGGCCAGAATGTGAAAACCTTCCTGGACTGTGCCCGACGGTACTATCGGGAAGCAGACGATGCCAAAGTC
GGGCTGTGGTTCGACAACGCCGATGGACTGCTGGAGAGATCTGACCTGCATCCTCCAATGAAGAAAAAGATCCTGCCCCTGCTG
GTGGCCAATATTCTGCAGACAGATGAAACCACAGGCCAGAAGTTTCTGGACGAGATCTGGCGAAAACAGATTAAGGGGCGGGAA
ACTGTGGCTAGCCGATGTGCACGGATCGAGACAGTGCGGAAATCCTTCGGGGGAGGCTTTAACATTGCCTACAATACCGCTCAG
TATAGGGAGGTGAACAAGCTGCCCCGCAATGCCCAGGATAAAGAACTGCTGACAATCAGAGATAGGGTGGCTGAGACTGCAGAC
TTCATTGCCGCTAACCTGGGGCTGTCTGACGAGCAGAAAAGAAAGTTCGCCAATCCTTTTAGTCTGGCTCAGTTCTACACCCTG
ATCGAGACAGAAGTGTCCGGATTTTCTGCAACTACCCTGGCCGTCCACCTGGAGAACGCCTGGAGGATGACAATCAAGGATGCT
GTGATTAATGGGGAAACTGTCAGAGCAGCACAGTGCAGCAGGCTGCCTGCAGAGACAGCTCGCCCATTCGATGGACTGGTGAGA
AGGCTGGTCGACAGACAGGCTTGGGAGATCGCAAAGAGGGTGTCAACTGACATTCAGAGCAAAGTCGATTTCTCCAACGGCATC
GTGGACGTCAGCATTTTTGTGGAGGAAAATAAGTTCGAGTTTTCCGCATCTGTGGCCGATCTGAAAAAGAACAAACGGGTCAAA
GACAAGATGCTGTCCGAGGCCGAAAAGCTGGAAACCAGATGGCTGATCAAAAATGAGCGGATCAAGAAGGCCAGCCGGGGAACT
TGTCCCTACACCGGCGATAGGCTGGCTGAGGGGGGAGAAATCGACCACATTCTGCCCCGAAGCCTGATCAAGGATGCCCGGGGA
ATTGTGTTTAACGCTGAGCCTAATCTGATCTATGCAAGCTCCCGCGGCAACCAGCTGAAAAAGAATCAGCGATACAGTCTGTCA
GATCTGAAGGCCAACTATCGGAATGAGATCTTCAAAACTAGCAACATCGCTGCAATTACCGCCGAGATTGAGGACGTGGTCACT
AAGCTGCAGCAGACCCATAGACTGAAATTCTTTGATCTGCTGAATGAGCACGAACAGGACTGCGTGCGGCACGCCCTGTTCCTG
GACGATGGCAGCGAAGCTCGCGACGCAGTGCTGGAGCTGCTGGCAACACAGCGCCGAACTCGCGTCAACGGGACACAGATCTGG
ATGATTAAGAACCTGGCCAACAAGATCCGAGAGGAACTGCAGAATTGGTGTAAGACAACTAACAATAGACTGCACTTTCAGGCC
GCTGCAACTAACGTGTCCGATGCAAAGAATCTGAGGCTGAAACTGGCCCAGAACCAGCCCGACTTCGAGAAGCCAGATATCCAG
CCCATTGCCAGCCATTCCATCGACGCCCTGTGCTCTTTCGCTGTGGGGAGTGCTGACGCAGAACGCGATCAGAATGGATTTGAC
TACCTGGATGGCAAGACCGTGCTGGGACTGTATCCACAGAGCTGTGAGGTCATTCACCTGCAGGCCAAGCCCCAGGAGGAAAAA
AGTCATTTCGATTCAGTGGCTATCTTTAAGGAAGGCATCTACGCAGAGCAGTTCCTGCCTATCTTTACCCTGAACGAAAAGATC
TGGATTGGATATGAGACACTGAATGCCAAAGGCGAAAGATGCGGGGCTATTGAGGTGAGTGGCAAACAGCCAAAGGAGCTGCTG
GAAATGCTGGCCCCCTTCTTTAACAAGCCTGTGGGCGACCTGTCAGCCCACGCTACTTACCGGATCCTGAAAAAGCCTGCATAT
GAGTTTCTGGCAAAGGCAGCTCTGCAGCCACTGAGCGCAGAGGAAAAAAGACTGGCAGCCCTGCTGGATGCTCTGCGCTACTGT
ACCAGTCGAAAGTCACTGATGAGCCTGTTCATGGCTGCAAACGGGAAATCCCTGAAAAAGCGGGAGGACGTGCTGAAACCCAAG
CTGTTCCAGCTGAAGGTCGAGCTGAAAGGCGAAAAGAGCTTCAAGCTGAACGGGAGCCTGACCCTGCCTGTGAAGCAGGACTGG
CTGAGAATCTGCGATAGCCCAGAACTGGCAGACGCCTTTGGCAAACCCTGTTCCGCCGATGAGCTGACATCTAAGCTGGCTCGC
ATTTGGAAACGACCTGTGATGCGGGATCTGGCTCATGCACCAGTCCGGAGAGAGTTCAGCCTGCCCGCAATCGACAACCCAAGT
GGAGGGTTCAGGATTAGGCGCACCAACCTGTTTGGCAATGAGCTGTACCAGGTGCACGCCATCAACGCTAAAAAGTATCGCGGC
TTCGCCTCCGCTGGGTCTAATGTCGACTGGTCCAAGGGGATCCTGTTTAACGAGCTGCAGCATGAAAATCTGACCGAGTGCGGA
GGCAGGTTCATTACAAGCGCCGATGTGACTCCTATGTCCGAATGGCGCAAGGTGGTCGCAGAGGACAACCTGTCTATCTGGATT
GCTCCAGGGACAGAAGGACGACGGTACGTGAGGGTCGAGACAACATTCATCCAGGCCAGTCACTGGTTTGAACAGTCAGTGGAG
AATTGGGCCATTACTAGTCCTCTGTCACTGCCAGCTTCCTTCAAGGTGGACAAACCAGCTGAGTTTCAGAAGGCAGTCGGAACC
GAGCTGTCAGAACTGCTGGGCCAGCCCAGGAGCGAAATCTTCATTGAGAACGTGGGCAATGCCAAGCATATCCGCTTTTGGTAC
ATTGTGGTGAGCAGCAACAAAAAGATGAACGAGTCTTACAACAATGTGTCTAAGAGTTAAGAATTC

#4

ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGAAAAAA
GAAATCAAAGACTACTTCCTGGGGCTGGATGTGGGGACTGGGAGCGTGGGGTGGGCTGTGACCGATACTGACTACAAACTGCTG
AAGGCTAACCGAAAAGACCTGTGGGGCATGAGATGCTTCGAGACAGCCGAAACTGCTGAGGTGCGGAGACTGCACAGGGGAGCC
AGGCGCCGAATCGAGCGGAGAAAGAAACGCATTAAGCTGCTGCAGGAGCTGTTCTCTCAGGAAATCGCCAAAACCGATGAGGGC
TTCTTTCAGAGAATGAAGGAAAGCCCCTTTTACGCTGAGGACAAAACAATCCTGCAGGAAAACACTCTGTTCAATGACAAGGAT
TTTGCTGATAAGACTTACCACAAAGCATATCCTACCATTAATCATCTGATCAAGGCTTGGATTGAGAACAAGGTGAAACCAGAC
CCCCGACTGCTGTACCTGGCATGTCACAACATCATTAAGAAAAGGGGACATTTCCTGTTTGAAGGCGACTTCGATTCAGAGAAT
CAGTTTGATACCAGCATCCAGGCACTGTTCGAGTATCTGCGCGAGGACATGGAGGTGGACATCGATGCCGACAGCCAGAAGGTC

FIG. 9B

AAAGAGATTCTGAAGGATAGCTCCCTGAAGAACTCTGAAAAACAGAGTCGGCTGAATAAGATCCTGGGGCTGAAGCCTTCCGAC
AAACAGAAGAAAGCCATCACAAACCTGATTTCTGGAAACAAGATCAATTTCGCCGATCTGTACGACAATCCAGATCTGAAGGAC
GCTGAGAAAAACTCAATCAGCTTCTCCAAGGACGATTTTGATGCACTGAGTGACGATCTGGCCTCAATTCTGGGCGACAGCTTT
GAACTGCTGCTGAAGGCCAAAGCTGTCTATAACTGCTCTGTGCTGAGTAAGGTCATCGGGGACGAGCAGTACCTGAGCTTCGCC
AAGGTGAAAATCTACGAAAAGCACAAAACCGATCTGACAAAGCTGAAAAACGTGATCAAGAAACATTTCCCCAAGGACTACAAG
AAGGTCTTTGGATACAACAAGAACGAGAAAAACAACAACAATTACTCCGGCTATGTGGGAGTCTGTAAGACCAAGAGTAAGAAA
CTGATCATTAACAACTCAGTCAACCAGGAAGATTTCTACAAGTTTCTGAAAACTATCCTGTCAGCCAAGAGCGAGATCAAGGAA
GTGAATGACATCCTGACCGAGATTGAAACTGGCACCTTTCTGCCAAAGCAGATCTCTAAAAGTAACGCAGAGATTCCCTATCAG
CTGAGGAAAATGGAGCTGGAAAAGATCCTGTCCAATGCCGAAAAGCACTTCTCTTTTCTGAAGCAGAAAGACGAAAAAGGACTG
TCACATAGCGAGAAGATCATTATGCTGCTGACCTTCAAGATCCCTTACTATATTGGCCCAATCAACGATAATCACAAGAAATTC
TTTCCCGACACGATGCTGGGTGGTCAAGAAAGAGAAATCCCCTTCTGGCAAGACCACACCATGGAACTTCTTTGATCATATCGAC
AAGGAAAAAACAGCAGAGGCCTTCATTACTTCTAGGACCAATTTTTGCACTTACCTGGTGGGAGAGAGCGTCCTGCCTAAGTCT
AGTCTGCTGTACTCCGAATATACCGTGCTGAACGAGATCAACAATCTGCAGATCATTATCGATGGCAAGAATATTTGTGACATC
AAGCTGAAACAGAAGATCTACGAGGACCTGTTCAAGAAGTACAAGAAAATTACCCAGAAGCAGATCAGCACCTTCATCAAGCAC
GAAGGCATCTGCAACAAAACCGATGAAGTGATCATCCTGGGGATTGACAAGGAATGTACATCAAGCCTGAAAAGCTACATCGAG
CTGAAAAACATTTTCGGCAAGCAGGTGGATGAGATCTCCACTAAGAATATGCTGGAGGAAATTATCAGATGGGCTACCATCTAC
GACGAGGGGGAAGGAAAGACCATCCTGAAAACAAAGATCAAGGCAGAATACGGAAAGTATTGCTCCGACGAGCAGATTAAGAAA
ATCCTGAACCTGAAGTTCTCCGGCTGGGGGCGACTGTCTCGGAAATTTCTGGAGACAGTGACTAGTGAAATGCCCGGCTTCTCA
GAACCTGTCAATATTATCACCGCCATGAGGGAGACACAGAACAATCTGATGGAGCTGCTGTCCTCTGAGTTCACCTTCACCGAG
AACATTAAGAAAATCAATTCTGGATTCGAAGATGCCGAGAAGCAGTTTAGTTACGACGGCCTGGTGAAACCACTGTTTCTGAGT
CCCTCAGTCAAGAAAATGCTGTGGCAGACCCTGAAGCTGGTGAAAGAGATTAGCCATATCACACAGGCCCCCCCTAAGAAAATT
TTCATCGAAATGGCAAAGGGGGCCGAGCTGGAACCTGCTCGGACTAAGACCAGACTGAAAATCCTGCAGGATCTGTATAACAAT
TGTAAGAACGATGCTGACGCCTTCAGCTCAGAGATCAAAGACCTGAGCGGAAAGATTGAGAACGAAGATAATCTGAGGCTGCGC
TCCGACAAGCTGTACCTGTACTATACTCAGCTGGGGAAATGCATGTATTGTGGAAAGCCAATTGAGATCGGCCACGTGTTCGAT
ACCTCAAACTACGATATTGACCATATCTATCCCCAGAGCAAGATCAAAGACGATAGCATTTCCAATCGGGTGCTGGTCTGCAGC
TCCTGTAACAAGAACAAGGAGGACAAGTACCCACTGAAATCAGAGATCCAGAGCAAGCAGCGCGGCTTCTGGAACTTTCTGCAG
CGAAACAATTTCATTTCTCTGGAGAAGCTGAATAGACTGACAAGGGCCACTCCAATCAGTGACGATGAGACAGCCAAGTTTATT
GCTAGGCAGCTGGTGGAAACTCGCCAGGCTACCAAGGTGGCCGCTAAAGTCCTGGAAAAGATGTTCCCCGAGACAAAAATCGTG
TACAGCAAGGCCGAGACTGTCTCCATGTTCCGGAACAAGTTTGATATCGTGAAGTGCAGAGAAATTAACGATTTTCACCATGCT
CACGACGCATACCTGAATATCGTGGTCGGCAACGTGTATAATACCAAGTTCACAAACAATCCTTGGAACTTTATCAAGGAGAAA
AGAGATAATCCAAAGATTGCTGACACCTACAACTACTATAAGGTGTTCGATTATGACGTCAAAAGGAACAATATCACAGCATGG
GAGAAGGGGAAAACTATTATCACCGTGAAAGACATGCTGAAGAGAAACACACCAATCTACACTAGGCAGGCAGCCTGTAAGAAA
GGGGAGCTGTTCAATCAGACCATTATGAAGAAAGGACTGGGCCAGCACCCCCTGAAGAAAGAAGGACCTTTTTCCAATATCTCT
AAATACGGCGGGTATAACAAGGTGAGCGCTGCATACTATACACTGATTGAGTATGAGGAAAAGGGCAACAAAATCCGCTCCCTG
GAAACTATTCCCCTGTACCTGGTGAAAGATATCCAGAAGGATCAGGACGTCCTGAAGTCTTATCTGACAGACCTGCTGGGGAAG
AAAGAGTTCAAGATCCTGGTGCCCAAGATCAAGATCAACAGCCTGCTGAAGATCAATGGGTTTCCTTGCCATATTACAGGAAAA
ACTAACGATAGTTTCCTGCTGCGCCCTGCCGTGCAGTTTTGCTGTTCAAACAATGAGGTCCTGTACTTCAAGAAAATTATCCGG
TTTTCCGAAATCCGCTCTCAGCGAGAGAAGATCGGGAAAACAATTAGCCCATACGAGGACCTGAGCTTCCGGTCATATATCAAG
GAGAACCTGTGGAAGAAAACTAAGAACGATGAAATCGGAGAGAAGGAATTTTACGACCTGCTGCAGAAGAAAAACCTGGAGATC
TATGATATGCTGCTGACTAAGCACAAAGACACCATCTACAAGAAACGCCCTAATTCTGCCACCATTGATATCCTGGTGAAGGGG
AAAGAGAAGTTCAAAAGCCTGATTATCGAAAACCAGTTTGAAGTGATCCTGGAGATCCTGAAGCTGTTTTCTGCAACACGGAAT
GTCAGTGACCTGCAGCATATCGGAGGCAGCAAGTACTCCGGCGTGGCCAAAATCGGGAACAAGATCTCTAGTCTGGATAACTGT
ATCCTGATCTATCAGTCCATCACCGGCATCTTCGAGAAACGGATCGACCTGCTGAAGGTGTAAGAATTC

#5

ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGACCAAG
GAGTATTACCTGGGGCTGGATGTGGGGACCAATTCCGTGGGATGGGCAGTGACCGATTCTCAGTACAACCTGTGCAAGTTTAAG
AAAAAGGATATGTGGGGCATCCGGCTGTTCGAAAGCGCCAACACAGCAAAGGACCGGAGACTGCAGAGAGGGAATAGGCGCCGA
CTGGAGCGGAAAAAGCAGAGAATTGATCTGCTGCAGGAAATCTTCTCCCCAGAGATCTGCAAGATTGACCCCACTTTCTTTATC
CGACTGAACGAATCCCGGCTGCACCTGGAGGACAAGTCTAACGATTTCAAATACCCACTGTTTATTGAGAAGGACTATTCTGAT
ATCGAGTACTATAAAGAGTTCCCCACCATTTTTCACCTGAGGAAGCATCTGATCGAGAGTGAGGAAAAACAGGATATCCGGCTG
ATCTACCTGGCCCTGCACAACATCATTAAGACCCGAGGACATTTTCTGATTGACGGCGATCTGCAGAGCGCCAAGCAGCTGAGG
CCCATCCTGGATACATTCCTGCTGTCCCTGCAGGAGGAACAGAACCTGTCAGTGAGCCTGTCCGAAAATCAGAAGGACGAGTAT
GAGGAAATTCTGAAAAACCGCAGCATCGCCAAGTCCGAAAAAGTGAAAAAGCTGAAGAATCTGTTTGAGATCTCAGACGAGCTG
GAAAAAGAGGAGAAGAAGGCCCAGAGCGCCGTGATCGAGAACTTCTGCAAGTTTATCGTGGGAAATAAGGGCGATGTCTGTAAA
TTCCTGCGGGTGTCTAAGGAGGAACTGGAGATTGACTCTTTCAGTTTTTCAGAGGGCAAGTACGAGGACGACATCGTGAAAAAC
CTGGAGGAAAAAGTGCCTGAAAAGGTCTACCTGTTTGAGCAGATGAAGGCAATGTATGATTGGAATATTCTGGTCGACATCCTG
GAAACCGAGGAATACATCAGCTTCGCCAAAGTGAAGCAGTATGAGAAACACAAGACTAACCTGCGGCTGCTGAGAGACATCATT
CTGAAATACTGCACCAAGGATGAGTATAATCGGATGTTTAACGACGAGAAGGAAGCTGGCAGCTACACCGCATATGTGGGGAAA
CTGAAAAAGAACAACAAGAAGTACTGGATCGAGAAAAAGAGAAATCCCGAGGAGTTCTACAAATCCCTGGGCAAGCTGCTGGAT
AAAATTGAGCCTCTGAAGGAGGACCTGGAAGTGCTGACTATGATGATCGAGGAGTGTAAGAACCACACCCTGCTGCCAATTCAG
AAAAATAAGGACAACGGCGTGATCCCCCACCAGGTGCATGAGGTCGAACTGAAAAAGATCCTGGAAAATGCCAAAAAGTACTAT
TCCTTCCTGACCGAGACAGACAAGGATGGGTACTCAGTGGTCCAGAAAATCGAGAGCATTTTCAGGTTTCGCATCCCCTACTAT

FIG. 9C

GTGGGGCCTCTGAGTACCCGGCACCAGGAAAAGGGATCAAACGTGTGGATGGTCAGAAAACCTGGCAGGGAGGATCGCATCTAC
CCATGGAATATGGAGGAAATCATTGACTTTGAGAAGAGCAACGAAAATTTCATTACACGGATGACTAACAAATGTACATATCTG
ATCGGGGAAGATGTCCTGCCCAAGCATTCTCTGCTGTACAGTAAATATATGGTGCTGAATGAGCTGAACAATGTGAAGGTCAGA
GGAAAAAAGCTGCCTACATCTCTGAAACAGAAGGTGTTCGAGGACCTGTTTGAAAACAAATCCAAAGTGACTGGAAAGAATCTG
CTGGAGTACCTGCAGATCCAGGACAAAGATATCCAGATTGACGATCTGTCTGGCTTCGACAAGGACTTCAAGACCAGCCTGAAG
AGCTATCTGGACTTCAAAAAGCAGATTTTTGGGGAGGAAATCGAGAAGGAAAGCATTCAGAACATGATCGAAGATATCATTAAG
TGGATCACCATCTACGGCAATGACAAGGAGATGCTGAAACGAGTGATTCGGGCTAATTATAGCAACCAGCTGACAGAGGAACAG
ATGAAAAAGATCACTGGATTTCAGTACAGTGGCTGGGGGAACTTCTCAAAGATGTTTCTGAAAGGGATCAGCGGATCCGACGTG
AGCACCGGCGAAACATTCGACATCATTACCGCAATGTGGGAGACAGACAACAATCTGATGCAGATCCTGTCAAAAAAGTTCACC
TTTATGGACAACGTCGAGGACTTCAACAGCGGCAAGGTCGGGAAAATCGACAAGATTACTTACGATAGCACCGTGAAGGAAATG
TTCCTGTCCCCTGAGAACAAAAGGGCCGTCTGGCAGACCATTCAGGTGGCTGAGGAGATCAAGAAAGTGATGGGCTGCGAGCCA
AAAAAGATCTTTATTGAAATGGCACGGGGCGGGGAGAAGGTGAAAAAGAGGACAAAATCTCGCAAGGCCCAGCTGCTGGAGCTG
TACGCCGCTTGCGAGGAAGATTGTAGAGAACTGATCAAGGAGATTGAGGACCGGGACGAGAGGGACTTCAATAGCATGAAGCTG
TTTCTGTACTATACCCAGTTCGGGAAATGTATGTATTCCGGCGACGACATCGATATTAACGAGCTGATTCGCGGCAATTCTAAG
TGGGACCGAGATCACATCTACCCCCAGAGCAAAATTAAGGACGATTCCATCGATAACCTGGTGCTGGTCAATAAGACATATAAT
GCCAAAAAGTCCAATGAGCTGCTGTCTGAGGACATCCAGAAAAAGATGCATTCATTCTGGCTGAGCCTGCTGAACAAAAAGCTG
ATCACTAAAAGCAAGTACGACCGCCTGACTCGAAAGGGCGACTTTACCGATGAGGAACTGAGTGGGTTCATCGCTAGACAGCTG
GTGGAAACAAGGCAGTCAACTAAGGCAATCGCCGATATCTTCAAGCAGATCTACAGCTCCGAGGTGGTCTATGTGAAGAGCAGC
CTGGTGAGCGACTTCAGGAAAAAGCCACTGAACTACCTGAAGTCTCGGAGAGTCAATGATTACCACCATGCAAAAGACGCCTAT
CTGAACATTGTGGTCGGGAACGTGTACAACAAAAAGTTTACCAGTAATCCCATCCAGTGGATGAAAAAGAATCGCGATACAAAC
TATAGCCTGAACAAGGTGTTCGAACACGACGTGGTCATTAACGGAGAAGTGATCTGGGAAAAGTGCACATACCATGAGGACACT
AATACCTATGATGGAGGCACTCTGGACCGAATCCGGAAGATTGTGGAACGCGATAACATTCTGTACACCGAGTACGCTTATTGT
GAGAAGGGCGAACTGTTTAATGCAACCATCCAGAACAAAAATGGAAACTCCACAGTCTCTCTGAAAAAGGGCCTGGACGTGAAA
AAGTACGGGGGATACTTCAGCGCCAACACAAGTTACTTCTCACTGATCGAGTTTGAGGACAAGAAGGGGGATAGAGCAAGGCAC
ATCATTGGAGTGCCTATCTATATTGCAAACATGCTGGAGCATTCTCCAAGTGCCTTCCTGGAGTACTGCGAACAGAAGGGGTAT
CAGAATGTGCGGATTCTGGTCGAGAAAATCAAAAAGAACAGCCTGCTGATCATTAATGGATACCCTCTGCGCATTCGAGGCGAG
AACGAAGTGGATACTTCCTTTAAGAGGGCCATCCAGCTGAAGCTGGACCAGAAAAACTATGAGCTGGTCCGCAATATCGAGAAG
TTCCTGGAAAAATACGTGGAGAAAAAGGGAAACTATCCAATTGACGAGAATAGAGATCACATCACACATGAAAAGATGAACCAG
CTGTACGAGGTGCTGCTGTCCAAAATGAAAAAGTTCAACAAGAAGGGCATGGCCGACCCCTCTGATAGGATCGAAAAGAGTAAG
CCTAAATTCATCAAGCTGGAGGACCTGATCGATAAGATTAATGTGATCAACAAAATGCTGAACCTGCTGCGCTGTGACAATGAT
ACTAAGGCCGACCTGTCTCTGATTGAGCTGCCCAAAAACGCTGGGAGTTTCGTGGTCAAAAAGAATACCATCGGAAAGTCAAAA
ATCATCCTGGTGAATCAGAGCGTGACTGGACTGTACGAGAATAGACGGGAACTGTAAGAATTC

#6
ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGGGGAGG
AAACCTTACATTCTGTCTCTGGATATTGGAACTGGGTCCGTCGGCTACGCTTGCATGGATAAAGGATTCAACGTGCTGAAGTAC
CACGACAAAGATGCCCTGGGAGTGTATCTGTTCGACGGCGCTCTGACTGCACAGGAGCGGAGACAGTTTAGGACCTCCAGGCGC
CGAAAGAACCGGAGAATCAAACGCCTGGGCCTGCTGCAGGAACTGCTGGCACCCCTGGTGCAGAACCCTAATTTCTACCAGTTT
CAGCGGCAGTTCGCCTGGAAGAACGACAATATGGATTTTAAGAACAAGAGCCTGTCTGAGGTGCTGAGCTTCCTGGGATATGAA
TCCAAGAAATACCCTACCATCTACCACCTGCAGGAGGCTCTGCTGCTGAAAGACGAGAAGTTTGATCCAGAACTGATCTACATG
GCACTGTATCATCTGGTGAAATACAGAGGCCACTTTCTGTTCGATCATCTGAAGATCGAGAACCTGACTAACAATGACAATATG
CACGATTTCGTGGAGCTGATTGAAACCTATGAGAACCTGAACAATATCAAGCTGAATCTGGACTACGAGAAAACCAAAGTGATC
TATGAGATTCTGAAAGACAACGAAATGACTAAGAATGATAGAGCCAAAAGGGTCAAGAACATGGAGAAGAAACTGGAACAGTTC
TCTATCATGCTGCTGGGGCTGAAGTTCAATGAGGGAAAACTGTTTAACCACGCCGATAATGCTGAGGAACTGAAGGGGGCTAAC
CAGAGCCATACATTTGCAGACAACTACGAGGAAAATCTGACTCCCTTCCTGACCGTGGAACAGTCAGAGTTTATTGAAAGGGCC
AACAAAATCTATCTGAGCCTGACTCTGCAGGATATCCTGAAGGGCAAGAAATCAATGGCTATGAGCAAAGTGGCCGCTTACGAC
AAGTTCAGAAATGAGCTGAAACAGGTGAAGGACATTGTCTATAAGGCTGATTCTACCAGGACACAGTTCAAGAAAATCTTTGTG
AGCTCCAAGAAAAGTCTGAAGCAGTACGACGCAACTCCCAACGATCAGACCTTCTCTAGTCTGTGCCTGTTTGACCAGTACCTG
ATTCGCCCAAAGAAACAGTATAGCCTGCTGATCAAGGAGCTGAAGAAAATCATTCCCCAGGACTCCGAACTGTACTTTGAGGCA
GAAAATGATACCCTGCTGAAGGTGCTGAACACCACAGACAATGCTAGCATCCCTATGCAGATTAACCTGTACGAGGCAGAAACC
ATCCTGCGAAATCAGCAGAAATATCACGCCGAGATCACAGATGAGATGATTGAAAAGGTGCTGTCTCTGATCCAGTTCCGCATT
CCATACTATGTGGGGCCCCTGGTCAACGACCATACAGCCAGTAAGTTTGGATGGATGGAGCGCAAAAGTAACGAATCAATCAAG
CCTTGGAATTTCGACGAGGTGGTCGATCGAAGTAAATCAGCCACTCAGTTTATTAGGCGCATGACCAACAAGTGTTCCTACCTG
ATCAATGAGGATGTGCTGCCAAAAAACTCTCTGCTGTATCAGGAGATGGAAGTCCTGAACGAACTGAATGCCACACAGATCAGG
CTGCAGACTGACCCAAAAAACCGCAAGTACCGAATGATGCCCCAGATTAAGCTGTTCGCTGTGGAGCACATCTTTAAGAAATAT
AAAACCGTCAGCCATTCCAAGTTCCTGGAAATTATGCTGAACAGCAATCACAGGGAGAACTTTATGAATCATGGAGAAAAGCTG
AGTATCTTCGGCACACAGGACGATAAGAAATTTGCATCAAAGCTGTCAAGCTACCAGGACATGACTAAAATCTTCGGGGATATT
GAGGGAAAGCGCGCCCAGATTGAGGAAATCATTCAGTGGATCACCATTTTTGAGGACAAGAAAATCCTGGTGCAGAAGCTGAAA
GAGTGCTATCCTGAACTGACATCCAAGCAGATCAACCAGCTGAAGAAACTGAATTACTCTGGCTGGGGGAGGCTGAGTGAGAAG
CTGCTGACTCACGCCTATCAGGGCCATAGCATCATTGAACTGCTGCGCCACTCCGATGAGAATTTCATGGAAATTCTGACCAAC
GACGTGTACGGGTTCCAGAATTTTATCAAAGAGGAAAACCAGGTCCAGAGCAATAAGATCCAGCATCAGGATATTGCCAACCTG
ACTACCTCTCCCGCTCTGAAGAAAGGCATCTGGAGTACAATTAAGCTGGTGCGGGAGCTGACTTCCATTTTCGGGGAGCCTGAA
AAGATCATTATGGAGTTTGCTACCGAGGACCAGCAGAAAGGCAAGAAACAGAAATCAAGAAAGCAGCTGTGGGACGATAACATC

FIG. 9D

AAGAAAAATAAGCTGAAAAGCGTGGACGAGTACAAATATATCATTGATGTCGCCAATAAGCTGAACAATGAGCAGCTGCAGCAG
GAAAAACTGTGGCTGTACCTGAGCCAGAACGGCAAGTGTATGTATAGCGGGCAGTCCATCGACCTGGATGCCCTGCTGTCCCCC
AATGCTACCAAGCACTACGAGGTGGATCATATTTTCCCTCGGAGCTTCATCAAGGACGATAGCATTGACAACAAGGTGCTGGTC
ATCAAGAAAATGAATCAGACAAAGGGCGATCAGGTGCCCCTGCAGTTCATTCAGCAGCCTTACGAGAGAATCGCATATTGGAAG
AGCCTGAACAAAGCCGGGCTGATCTCTGATAGTAAACTGCACAAGCTGATGAAACCAGAGTTCACAGCTATGGACAAGGAAGGC
TTCATCCAGCGGCAGCTGGTGGAGACTAGACAGATCAGCGTGCATGTCCGGGATTTTCTGAAAGAGGAATACCCTAATACCAAA
GTGATCCCAATGAAGGCCAAAATGGTGAGCGAGTTCCGGAAGAAATTTGACATCCCAAAGATTAGACAGATGAACGACGCACAC
CATGCCATCGATGCTTACCTGAATGGCGTGGTCTATCACGGGGCACAGCTGGCCTACCCCAACGTGGACCTGTTTGATTTCAAT
TTTAAGTGGGAGAAAGTCCGAGAAAAGTGGAAAGCCCTGGGAGAGTTCAACACAAAGCAGAAATCTCGGGAACTGTTCTTTTTC
AAGAAACTGGAGAAGATGGAAGTGTCCCAGGGCGAGCGGCTGATCTCTAAGATCAAGCTGGACATGAACCACTTCAAGATCAAC
TACTCCAGAAAGCTGGCCAACATCCCTCAGCAGTTTTATAATCAGACCGCAGTGTCTCCAAAGACAGCCGAGCTGAAATACGAA
TCTAACAAGAGTAATGAGGTGGTCTATAAGGGACTGACACCATACCAGACTTATGTGGTCGCCATCAAGAGCGTGAACAAGAAA
GGCAAGGAGAAAATGGAATACCAGATGATCGACCACTACGTGTTCGATTTTTATAAATTCCAGAACGGCAATGAGAAGGAACTG
GCTCTGTACCTGGCACAGAGGGAGAACAAGGACGAAGTGCTGGATGCTCAGATTGTCTATAGTCTGAATAAGGGGGATCTGCTG
TACATCAACAATCATCCCTGCTATTTCGTGTCACGCAAAGAGGTCATCAACGCAAAGCAGTTTGAGCTGACCGTGGAACAGCAG
CTGTCTCTGTACAACGTGATGAACAACAAGGAGACAAATGTCGAAAAGCTGCTGATCGAGTATGACTTCATTGCCGAGAAAGTG
ATCAACGAATACCACCATTATCTGAATAGCAAGCTGAAAGAAAAGCGAGTCCGGACCTTTTTCTCAGAGAGCAACCAGACACAC
GAGGACTTCATCAAGGCCCTGGACGAGCTGTTTAAGGTGGTCACCGCATCCGCCACAAGGTCTGATAAAATCGGGAGTCGCAAG
AACAGCATGACTCATCGAGCCTTCCTGGGAAAAGGCAAGGACGTGAAGATTGCTTACACCTCCATCTCTGGACTGAAAACAACT
AAACCTAAGAGTCTGTTTAAGCTGGCCGAGTCAAGAAACGAACTGTAAGAATTC

#7
ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGACAAAA
ATCAAAGACGACTACATCGTGGGACTGGACATCGGCACAGACTCCTGCGGGTGGGTGGCTATGAACAGCAATAATGACATTCTG
AAACTGCAGGGCAAGACCGCAATCGGGTCACGCCTGTTCGAGGGAGGGAAGAGCGCAGCTGAACGGAGACTGTTTCGCACCACA
CACAGGCGCATCAAACGACGGAGATGGCGACTGAAGCTGCTGGAGGAGTTCTTCGACCCCTACATGGCAGAGGTGGATCCTTAT
TTCTTTGCCCGGCTGAAGGAATCTGGCCTGAGTCCACTGGACAAAAGAAAGACCGTGAGCTCCATTGTGTTCCCCACATCCGCC
GAGGATAAGAAGTTCTACGACGATTACCCTACAATCTACCATCTGAGGTATAAACTGATGACTGAGGACGAAAAGTTCGATCTG
CGCGAAGTGTACCTGGCTATCCACCATATCATTAAGTACCGAGGAAACTTCCTGTATAATACCAGTGTGAAAGACTTCAAGGCA
TCAAAGATCGATGTCAAATCTAGTATCGAGAAGCTGAACGAGCTGTATGAAAATCTGGGCCTGGACCTGAACGTGGAGTTCAAC
ATTAGCAATACTGCCGAGATCGAAAAGGTGCTGAAAGACAAGCAGATCTTCAAGCGGGATAAAGTCAAGAAAATTGCCGAGCTG
TTTGCTATCAAAACCGACAACAAGGAACAGAGCAAGAGAATCAAAGATATTTCCAAACAGGTGGCCAATGCTGTCCTGGGGTAC
AAGACCAGGTTCGACACAATCGCTCTGAAAGAGATTTCCAAGGACGAACTGTCTGATTGGAACTTCAAACTGTCAGACATCGAT
GCAGACAGCAAGTTTGAGGCCCTGATGGGAAACCTGGATGAGAATGAACAGGCCATCCTGCTGACTATTAAGGAGCTGTTTAAC
GAAGTGACCCTGAATGGAATTGTCGAGGACGGCAACACCCTGAGCGAATCCATGATCAACAAGTACAATGATCACCGGGACGAT
CTGAAGCTGCTGAAAGAAGTGATCGAAAATCATATTGACAGAAAGAAAGCCAAGGAGCTGGCACTGGCCTACGATCTGTATGTC
AACAATAGGCACGGACAGCTGCTGCAGGCTAAGAAAAAGCTGGGCAAAATCAAGCCCCGCTCTAAGGAGGACTTCTACAAAGTG
GTCAACAAGAATCTGGACGATTCACGGGCAAGCAAGGAGATCAAAAAGAAAATTGAACTGGACAGCTTTATGCCTAAGCAGAGA
ACCAACGCCAATGGCGTGATCCCATACCAGCTGCAGCAGCTGGAGCTGGATAAGATCATCGAAAACCAGTCTAAGTACTATCCA
TTCCTGAAGGAGATTAATCCCGTGTCAAGCCACCTGAAAGAGGCCCCCTATAAGCTGGACGAACTGATCCGATTTCGGGTGCCT
TACTATGTCGGCCCCCTGATTTCTCCTAACGAGAGTACCAAGGATATCCAGACAAAGAAAAACCAGAATTTCGCCTGGATGATT
CGCAAAGAGGAAGGGCGAATCACACCTTGGAACTTTGACCAGAAGGTGGATCGAATTGAGAGCGCCAATAAGTTCATCAAACGG
ATGACTACCAAGGACACTTACCTGTTTGGGGAGGATGTGCTGCCAGCTAACAGCCTGCTGTATCAGAAGTTCACCGTCCTGAAC
GAACTGAACAACATCCGGATTAATGGAAAAAGAATCTCCGTGGACCTGAAGCAGGAGATCTACGAAAACCTGTTTAAGAAACAC
ACAACTGTGACCGTCAAGAAACTGGAGAATTATCTGAAGGAAAACCATAATCTGGTGAAAGTCGAGATCAAGGGGCTGGCCGAT
GAAAAGAAATTCAACAGCGGACTGACCACATACAATAGATTCAAGAACCTGAACATCTTTGACAACCAGATTGACGATCTGAAG
TACAGGAACGATTTCGAGAAGATCATCGAATGGTCTACAATTTTTGAGGACAAGAGTATCTACAAAGAAAAGCTGAGGAGCATC
GATTGGCTGAACGAGAAGCAGATTAACGCTCTGTCTAATATCAGACTGCAGGGGTGGGGAAGGCTGAGTAAGAAACTGCTGGCA
CAGCTGCACGACCATAATGGCCAGACCATCATTGAGCAGCTGTGGGATTCCCAGAACAATTTCATGCAGATTGTGACACAGGCC
GACTTTAAAGATGCTATCGCAAAGGCCAACCAGAATCTGCTGGTGGCTACCTCAGTCGAGGACATTCTGAACAATGCATACACA
AGCCCCGCAAACAAGAAAGCCATCAGACAGGTCATCAAGGTGGTCGACGATATCGTGAAGGCAGCCTCCGGAAAGGTCCCAAAA
CAGATCGCCATTGAGTTCACTAGGGATGCTGACGAAAATCCCAAGAGAAGTCAGACCAGGGGCTCAAAGCTGCAGAAAGTGTAC
AAGGACCTGAGCACTGAGCTGGCCTCCAAGACCATTGCTGAGGAACTGAACGAAGCAATCAAAGACAAGAAACTGGTGCAGGAT
AAGTACTATCTGTACTTTATGCAGCTGGGGCGGGACGCCTATACAGGAGAGCCTATCAATATCGATGAAATCCAGAAGTACGAT
ATCGACCACATTCTGCCACAGTCTTTCATCAAGGACGATGCCCTGGACAACAGGGTGCTGGTGAGCCGGGCTGTGAACAATGGC
AAATCTGATAATGTGCCTGTCAAGCTGTTTGGCAACGAGATGGCTGCAAATCTGGGGATGACTATCAGGAAAATGTGGGAGGAA
TGGAAGAACATCGGCCTGATTAGCAAAACAAAGTACAACAATCTGCTGACTGATCCCGACCACATTAACAAGTATAAGAGTGCC
GGGTTCATCAGGCGCCAGCTGGTGGAGACATCACAGATCATCAAGCTGGTGAGCACTATCCTGCAGAGTCGCTACCCTAACACT
GAAATCATTACCGTGAAGGCTAAGTACAATCATTATCTGCGGGAGAAATTTGACCTGTATAAGAGCAGAGAAGTCAACGACTAC
CACCATGCTATTGATGCATATCTGTCCGCCATCTGCGGAAATCTGCTGTACCAGAACTATCCAAATCTGCGGCCCTTCTTTGTG
TACGGCCAGTATAAGAAATTCTCCTCTGATCCTGACAAAGAGAAGGCCATTTTTAACAAAACCCGCAAGTTCTCCTTTATCTCT
CAGCTGCTGAAAAACAAGAGTGAGAACAGCAAGGAAATCGCTAAGAAACTGAAACGGGCATACCAGTTCAAGTATATGCTGGTG
TCTCGAGAGACTGAAACCCGGGACCAGGAGATGTTCAAAATGACCGTGTACCCCCGGTTCAGCCACGATACAGTCAAGGCTCCT

FIG. 9E

AGGAACCTGATTCCAAAGAAAATGGGCATGTCCCCTGACATCTACGGAGGCTATACAAACAATTCTGACGCATACATGGTCATC
GTCCGCATTGATAAGAAAAAGGGAACTGAGTATAAGATCCTGGGCATTCCAACCCGGGAACTGGTGAATCTGAAAAAGGCCGAG
AAGGAGGACCATTACAAAAGCTATCTGAAGGAGATCCTGACACCAAGGATTCTGTACAACAAAAATGGGAAGCGCGATAAAAAG
ATCACTTCCTTCGAAATTGTGAAATCTAAGATCCCCTATAAGCAGGTCATCCAGGATGGGGACAAAAAGTTTATGCTGGGAAGT
TCAACATACGTGTATAACGCAAAGCAGCTGACACTGAGCACTGAGTCCATGAAAGCCATCACTAACAATTTCGATAAGGACAGC
GATGAGAACGACGCTCTGATTAAGGCATACGATGAAATCCTGGACAAAGTGGATAAGTATCTGCCACTGTTCGACATCAACAAG
TTCCGGGAGAAGCTGCACAGTGGGCGAGAAAAGTTCATCAAGCTGAGCCTGGAGGACAAAAAGGATACCATCCTGAAAGTGCTG
GAAGGACTGCATGATAACGCTGTCATGACAAAGATCCCTACTATTGGCCTGTCCACACCACTGGGGTTCATGCAGTTTCCCAAC
GGCGTGATTCTGAGCGAGAATGCCAAACTGATCTACCAGTCCCCCACCGGGCTGTTCAAAAAGTCAGTGAAGATCAGCGACCTG
TAAGAATTC

#8
ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGGGCTAC
ACTGTGGGACTGGATATTGGGGTGGCTTCCGTCGGGGTGGCTGTGCTGGATGAGAATGACAACATCGTGGAGGCTGTGTCAAAC
ATCTTTGATGAAGCCGACACAAGCAACAATAAGGTGCGGAGAACTCTGAGGGAGGGCAGGCGCACAAAGCGGCGGCAGAAAACC
CGCATTGAGGACTTCAAGCAGCTGTGGGAGACTTCAGGCTACATCATTCCTCACAAGCTGCATCTGAATATCATTGAGCTGCGC
AACAAAGGGCTGACCGAACTGCTGAGCCTGGATGAGCTGTATTGCGTGCTGCTGTCCATGCTGAAGCACCGGGGGATCTCCTAC
CTGGAGGACGCCGACGATGGCGAGAAGGGGAATGCCTATAAGAAAGGACTGGCTTTTAACGAAAAACAGCTGAAGGAGAAAATG
CCATGTGAGATCCAGCTGGAACGCATGAAGAAATACGGGAAGTACCATGGAGAGTTCATCATCGAAATTAATGATGAGAAGGAA
TACCAGAGCAACGTGTTCACCACAAAGGCTTATAAGAAGGAGCTGGAAAAGATCTTCGAGACACAGCGGTGCAACGGCAACAAG
ATCAACACAAAGTTCATTAAGAAATACATGGAGATCTACGAACGAAAGCGGGAATACTATATCGGACCAGGCAATGAGAAAAGC
AGAACAGACTACGGCATCTATACTACCAGGACTGATGAGGAAGGGAATTTCATCGACGAGAAGAACATTTTTGGCAAACTGATC
GGGAAGTGTAGTGTGTACCCCGAGGAATATAGAGCAAGCTCCGCCTCATACACCGCCCAGGAGTTCAATCTGCTGAACGATCTG
AACAATCTGAAAATCAACAATGAGAAGCTGACAGAATTTCAGAAGAAAGAGATTGTCGAAATCATTAAGGACGCTTCTAGTGTG
AACATGAGGAAAATCATTAAGAAAGTCATCGATGAGGACATTGAACAGTACAGCGGAGCACGAATCGATAAGAAAGGCAAGGAA
ATCTACCACACCTTCGAGATCTATCGGAAGCTGAAGAAAGAGCTGAAAACAATCAATGTGGATATCGACTCTTTTACTAGAGAG
GAACTGGATAAGACCATGGACATCCTGACCCTGAACACAGAGAGGGAAAGTATTGTGAAGGCCTTCGACGAACAGAAATTTGTC
TACGAGGAAAATCTGATCAAGAAACTGATTGAGTTTCGGAAGAACAATCAGAGACTGTTCAGCGGCTGGCATAGTTTTTCATAC
AAGGCTATGCTGCAGCTGATCCCAGTGATGTACAAGGAGCCCAAAGAACAGATGCAGCTGCTGACCGAAATGAACGTGTTCAAA
AGTAAGAAAGAGAAGTACGTCAACTACAAGTACATCCCAGAGAACGAAGTGGTCAAGGAGATCTATAACCCCGTGGTCGTGAAG
AGCATTAGAACAACTGTGAAAATTCTGAATGCACTGATCAAGAAATACGGGTATCCTGAATCCGTCGTGATCGAGATGCCAAGG
GATAAGAACTCTGACGATGAGAAGGAAAAGATCGACATGAACCAGAAGAAAAACCAGGAGGAATACGAGAAAATCCTGAACAAG
ATCTACGATGAGAAGGGAATCGAAATTACCAACAAGGACTACAAGAAACAGAAGAAACTGGTGCTGAAGCTGAAACTGTGGAAC
GAGCAGGAAGGACTGTGCCTGTATTCCGGCAAGAAAATCGCTATTGAGGATCTGCTGAATCACCCCGAGTTCTTTGAAATTGAC
CATATCATTCCTAAGAGCATCTCCCTGGACGATTCTCGCAGTAACAAGGTCCTGGTGTACAAAACAGAAAATTCTATCAAGGAG
AACGATACCCCCTACCACTATCTGACACGGATTAACGGAAAGTGGGGCTTTGACGAATATAAAGCTAATGTGCTGGAGCTGAGA
AGGCGCGGCAAGATCGACGATAAGAAAGTGAACAATCTGCTGTGCATGGAGGATATCACTAAGATTGACGTCGTGAAAGGGTTC
ATTAACCGCAATCTGAACGACACCAGATACGCATCCAGGGTGGTGCTGAACGAAATGCAGTCCTTCTTTGAGTCTCGAAAGTAC
TGTAATACTAAGGTCAAAGTGATCCGAGGCTCTCTGACCTATCAGATGCGGCAGGATCTGCACCTGAAGAAAAACAGAGAGGAA
TCATACAGCCACCATGCTGTGGACGCAATGCTGATCGCATTCTCCCAGAAGGGGTACGAGGCCTATAGGAAGATCCAGAAAGAT
TGCTACGACTTTGAGACAGGCGAAATTCTGGACAAGGAAAAATGGAATAAGTACATTGACGATGACGAGTTTGATGACATCCTG
TATAAAGAGAGGATGAACGAAATCCGCAAGAAAATCATTGAGGCCGAGGAAAAGGTGAAGTACAACTACAAGATCGATAAGAAG
TGCAATCGCGGGCTGTGTAACCAGACTATCTACGGGACCCGAGAAAAGGACGGAAAAATCCACAAGATTTCAAGCTACAACATC
TATGATGACAAGGAGTGTAATTCCCTGAAGAAAATGATTAACAGTGGGAAAGGATCAGATCTGCTGATGTACAACAATGATCCT
AAGACATATCGCGACATGCTGAAAATCCTGGAAACTTACTCCTCTGAGAAGAATCCATTCGTGGCATATAACAAAGAGACAGGA
GACTACTTTCGGAAATATTCTAAGAATCACAACGGACCCAAGGTCGAGAAGGTGAAATACTATAGCGGCCAGATCAACTCCTGC
ATCGATATTTCTCACAAGTACGGCCATGCCAAAAATAGTAAGAAAGTCGTGCTGGTGTCACTGAACCCTTATAGAACCGACGTC
TACTATGATAATGACACAGGCAAGTACTATCTGGTCGGGGTGAAGTACAATCATATCAAATGTGTCGGAAACAAGTACGTGATT
GATAGCGAGACATATAACGAACTGCTGAGGAAGGAGGGCGTGCTGAACAGCGATGAGAACCTGGAGGACCTGAACAGCAAAAAC
ATCACTTACAAGTTCAGTCTGTACAAGAACGACATCATCCAGTACGAGAAGGGCGGGGAATACTATACAGAGCGCTTTCTGAGC
CGAATCAAAGAACAGAAGAACCTGATTGAGACTAAACCCATCAATAAGCCTAACTTCCAGCGGAAGAATAAGAAAGGCGAGTGG
GAAAATACCAGAAACCAGATCGCCCTGGCTAAGACTAAATACGTGGGGAAGCTGGTCACCGATGTGCTGGGAAACTGTTACATC
GTGAACATGGAGAAGTTCTCCCTGGTCGTGGACAAATAAGAATTC

#9
ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGACTAAC
GGCAAGATTCTGGGGCTGGACATTGGCATCGCAAGCGTGGGGGTGGGGATTATTGAGGCAAAAACTGGAAAGGTGGTGCATGCC
AATTCCCGGCTGTTCTCTGCCGCTAACGCTGAGAACAATGCAGAACGGAGAGGGTTTAGGGGATCTAGGCGCCTGAATCGACGG
AAGAAACACCGCGTGAAGCGAGTCCGGGATCTGTTCGAGAAATACGGAATCGTCACCGACTTTCGCAACCTGAATCTGAACCCT
TATGAGCTGCGAGTGAAGGGCCTGACCGAACAGCTGAAAAACGAGGAACTGTTCGCAGCCCTGAGAACAATCTCTAAGAGAAGG
GGGATTAGTTACCTGGACGATGCCGAGGACGATAGTACCGGATCAACAGACTATGCTAAGAGCATCGATGAGAATCGCCGACTG
CTGAAAAACAAGACACCAGGCCAGATTCAGCTGGAGAGGCTGGAAAAGTACGGCCAGCTGCGCGGGAATTTCACCGTCTATGAC
GAGAACGGGGAAGCCCATCGCCTGATCAATGTGTTTAGTACATCAGATTACGAGAAAGAAGCACGGAAGATCCTGGAGACACAG

FIG. 9F

GCCGACTACAACAAGAAAATCACAGCTGAGTTCATTGACGATTATGTGGAAATCCTGACCCAGAAACGAAAGTACTATCACGGC
CCCGGGAACGAAAAGAGCCGGACTGACTACGGACGGTTCCGGACCGATGGGACCACACTGGAGAATATTTTCGGAATCCTGATT
GGCAAGTGCAACTTTTACCCTGATGAATATCGAGCAAGCAAGGCCAGCTACACCGCACAGGAGTATAATTTCCTGAACGACCTG
AACAATCTGAAGGTGAGCACCGAAACAGGGAAGCTGTCAACAGAGCAGAAAGAAAGCCTGGTGGAGTTTGCCAAGAATACTGCT
ACCCTGGGACCCGCTAAACTGCTGAAGGAGATCGCAAAAATTCTGGACTGTAAGGTGGATGAGATCAAAGGATACAGAGAGGAC
GATAAAGGCAAGCCAGATCTGCATACCTTCGAGCCCTATAGGAAACTGAAGTTTAATCTGGAAAGCATCAACATTGACGATCTG
TCCCGCGAAGTGATCGACAAGCTGGCTGATATTCTGACTCTGAACACCGAGAGAGAAGGAATCGAGGACGCAATTAAGAGGAAT
CTGCCAAACCAGTTCACAGAGGAACAGATCAGCGAGATCATCAAGGTGCGGAAGAGCCAGTCCACTGCTTTCAATAAGGGCTGG
CACTCTTTTAGTGCAAAACTGATGAACGAGCTGATCCCCGAACTGTACGCCACCTCCGACGAGCAGATGACAATTCTGACTCGG
CTGGAAAAATTCAAGGTCAATAAGAAAAGCTCCAAAAACACAAAGACTATCGACGAGAAGGAAGTCACTGATGAGATCTACAAT
CCTGTGGTCGCCAAGAGCGTGAGACAGACCATCAAAATCATTAACGCTGCAGTCAAGAAATATGGCGACTTCGATAAGATCGTG
ATTGAAATGCCACGGGATAAAAATGCTGACGATGAGAAGAAGTTCATCGACAAGAGAAATAAGGAGAACAAGAAGGAAAAGGAC
GATGCCCTGAAAAGGGCCGCTTACCTGTATAATTCTAGTGACAAGCTGCCCGATGAGGTGTTCCACGGCAACAAGCAGCTGGAA
ACCAAAATCCGACTGTGGTATCAGCAGGGGGAGCGGTGCCTGTATAGTGGAAAGCCCATCTCAATTCAGGAGCTGGTGCATAAC
TCTAACAATTTCGAAATCGATCACATTCTGCCTCTGTCACTGAGCTTTGACGATAGTCTGGCCAATAAGGTGCTGGTCTACGCT
TGGACAAACCAGGAGAAAGGCCAGAAAACCCCTTATCAGGTCATCGACTCCATGGATGCAGCCTGGTCTTTCAGGGAGATGAAG
GACTACGTGCTGAAACAGAAGGGACTGGGCAAGAAAAAGCGCGACTATCTGCTGACTACCGAGAACATCGATAAGATTGAAGTG
AAGAAGAAGTTCATCGAGAGGAATCTGGTGGATACTCGCTACGCATCTCGAGTGGTCCTGAACTCTCTGCAGAGTGCCCTGAGA
GAGCTGGGGAAAGACACTAAGGTGTCTGTGGTCAGGGGACAGTTCACCAGTCAGCTGCGGAGAAAATGGAAGATCGATAAGAGC
CGCGAGACATACCACCATCACGCAGTGGACGCCCTGATCATTGCTGCATCAAGCCAGCTGAAACTGTGGGAGAAGCAGGACAAT
CCCATGTTTGTGGATTATGGCAAGAACCAGGTGGTCGACAAACAGACTGGGGAGATCCTGTCCGTGTCTGACGATGAGTACAAG
GAACTGGTGTTCCAGCCCCCTTATCAGGGCTTTGTGAATACCATCTCCTCTAAAGGGTTCGAGGACGAAATTCTGTTTAGCTAC
CAGGTGGATTCCAAATATAACCGGAAGGTCAGTGACGCAACCATCTACTCAACAAGAAAAGCCAAGATTGGCAAGGATAAGAAA
GAGGAAACCTACGTGCTGGGAAAAATCAAGGACATCTACTCCCAGAATGGCTTCGATACCTTCATCAAGAAGTACAACAAAGAT
AAGACTCAGTTCCTGATGTATCAGAAGGACTCTCTGACATGGGAGAACGTGATCGAAGTCATTCTGAGGGACTACCCAACAACT
AAGAAAAGCGAGGACGGCAAAAATGATGTGAAGTGCAACCCCTTTGAGGAATACAGGCGCGAGAATGGGCTGATCTGTAAGTAT
TCCAAGAAAGGGAAAGGAACTCCCATCAAGAGCCTGAAGTACTATGACAAGAAACTGGGGAACTGCATCGATATTACCCCAGAG
GAATCACGCAATAAGGTCATCCTGCAGAGCATTAACCCTTGGCGAGCCGACGTGTACTTCAATCCAGAGACACTGAAGTACGAA
CTGATGGGCCTGAAATATTCAGATCTGAGCTTTGAAAAGGGCACTGGGAACTACCATATCAGCCAGGAGAAATATGACGCTATC
AAAGAGAAGGAAGGAATTGGCAAGAAATCCGAGTTCAAGTTTACACTGTACCGCAACGACCTGATCCTGATCAAGGATATCGCC
AGTGGCGAGCAGGAAATCTACAGATTCCTGTCAAGAACTATGCCCAATGTGAACCACTACGTCGAGCTGAAGCCTTACGACAAG
GAAAAGTTCGATAACGTGCAGGAGCTGGTCGAAGCACTGGGAGAGGCAGATAAAGTGGGACGATGTATCAAAGGACTGAATAAG
CCAAACATCAGCATCTACAAGGTGAGAACCGACGTCCTGGGAAACAAATATTTCGTGAAGAAAAAGGGCGACAAACCCAAGCTG
GATTTTAAGAACAACAAGAAGTAAGAATTC

#10

ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGGCCGAC
CGAATCTCTCTGGGGCTGGACATTGGGGTGGCAAGCGTGGGGTTCTCAGTGCTGGACATTGATAAGGGAAAAGTCATTGAGCTG
GGCGCCAGGCTGTTCTCTGCTACTGTGGCCGCTGGCAACCAGGATCGAAGAGACATGCGAGGAGCCAGGCGCCTGCTGAACCGG
AACAAGCAGCGACGGCAGGATACCGGAAAGCTGTTCAAGAAATTTGGCCTGATCGACGATTTTGATAAGGGCAGCTTCTACGAC
AACTTTAATCAGAACCTGAATCCTTATGAGCTGAGAGTGAAAGGCCTGACAGAACAGCTGACTAAGGAGGAACTGGCCGAGTCT
CTGTACCAGATCGTGAAACATAGGGGGATTAGTTATGCACTGAAGGACGCCGATGTGGACGAAGGCGGGACAGACTACTCAGTC
AGCCTGAAAATCAACAGCCAGGAGCTGGCAGAAAAGACTCCAGCCCAGATTCAGCTGCAGAGACTGAATGATTATGGAAAGGTG
AGGGGCCAGGTGGTCATCGGCGACGATCCAGACAACCAGAAGGTGCTGCTGAATGTGTTCCCCACATCAGCTTACGAGAAAGAA
GCAAAGCAGATCATTGCCACTCAGCAGCAGTTCTATCCTGAGAGCCTGACCGACAAGTTCACCGAGGAATACTGCCAGATCCTG
ACTCGCAAGCGAGATTATTTTGTGGGGCCAGGAAACGAGAAAAGCCGGACCGACTACGGGATCTACAAGACTGATGGAAGAACC
CTGGACAATCTGTTCGAGGAACTGATCGGCCACGATAAGATCTACCCCGAGGAACTGCGGGCATCTGCAGCCAGTTATACCGCC
CAGCTGTTTAACGTGCTGAATGACCTGAACAATCTGAGAATCCTGAACTACGAGGATGGGAAACTGACAAAGGAGGACAAGGAA
AAGATCATCGCTGAAATTAAGAACAACACCACAACTATCAACATGCTGAATGTGATTAAGAAAGTCGCCGGGTGTTCCAAGGAC
GATATCAAAGGATTCCGAGTGAATGAGAAGGATAAACCCGAAATCAGCTCCATGCCTGTGTACCGCAAAATTCATAAGGACCTG
CTGAAGGCCGGCGTGGATATCTCAGACTGGCCCGTCGAGTTCATCGACGAACTGAGCTTTATTCTGACACTGAACACTGAGAAT
GGGGAAATTCGCAAACAGCTGAACAATCGACTGGCCCCTAAGTTCGATTTTCTGAACGCTGACCTGATCCAGCTGATCATTGAT
AATAAGGACTCCTTTGAGATTAAGACTAACAACAAGTGGCACAGATTCAGCGTGAAAACCATGAACAAACTGATCCCAGAGATG
ATGGAAAGACCCGTGGAGCAGATGACCCTGCTGAATGAAATGGGACTGGTCAAGAAAGATAAGAAACGCTTTGAGAACAATAAG
TACCTGCCTTACAAGGAAATCGCAAAGGACATTTTCAACCCAGTGGCCTCCAAATCTGTCCGCGAGGCCCTGAAGATCGTGAAT
GCTGTCCTGAAGAAATACGGCCACATTGATTATCTGGTGGTCGAAATGCCTCGGGATAAAAACCTGAAGGAGGAACAGGACAAT
ATCAAGGAGTTCCAGAACAAAAATAAGAAAGCTAAGGACGCTGCATTCGAAGCATTTGTGAAATCAGTCGGGAGCGAGCAGAGA
GTGAAGGAAGCCCTGTCTAAAAACCGGAAGCTGCAGATGAAGATGAGACTGTGGTATCAGCAGCAGGAGATCGATCCATATAAT
GGAAAGACAATCGATGCCACTGACCTGATTAACAATCCTGATAAGTTCGAGATTGACCATATCATTCCACAGAGTATCTCATTT
TACGACAGTATTAACAATAAGACCCTGTGCTTCGCCTCAATGAACCAGGTGAAAGGACAGAAAACCCCCTACGAGTTTATGCTG
GAAGGCCACGGGCAGTCCTATGACAAGTTCAAAGCTACAGTGATGGCAAACAAGAATTTTGGCAAGGCTAAAAGGGCAAACTAC
CTGTTCGAGGAAAATGTGAGCGATATCGAGACTCGGAAGAGATTCCTGTCCCGCAACCTGGTGGACACCCGATATTCTAGTCGG
GTGGTGCTGAACAGCCTGCAGGATTTCTTTCGGGAGAAATCTGCCGACACCAAGGTGACAGTCATTCGCGGCAAGTTTACCTCC

FIG. 9G

AACATGCGAAAACATTGGCACATCGATAAGACTAGGGAGACATTCCACCATCACGCCATTGACGCTTCTATCATTGCCGCTACA
CCATTTCTGCGCATGTGGAAGAAAGGAGGCACTATCTTCCCCGTGAAGGTCGGAGAAGAAAGTATCGATATTGAGACAGGCGAA
ATTCTGGACGATAAGAATTTTGACAAAGCAATGTACGAGGAACCCTATAGTGGCTTCGTGTCAGAGATCATGAACGCCGACGAT
CGGATCAAGTTCAGCCACCAGGTGGATAAGAAAATGAATAGGAAGGTGAGCGACGCCACCATCTACAGTACTCGCACCGGGAAA
CTGGCTAAGGATAAGAAAGACGCTGAGTACATCGTGGCAAAGGTCAAAGATATCTACAGCGTGGACGGATTCAAGAAGTTCAAG
AAAGTCTACGATAAGGACAAAACCAAGTTTCTGCTGTACAAATATGATCCTAGGACATTCTCAAAGCTGGAGCGCATCATTAGC
GATTGCCCAGACAAAGTGGAAAAGGTCCAGACAAACGGCAAAGTGAAGGCTGTCGATATCAGTCCATTCGAGATGTACAGAAGG
GACCATGGGATGATCAAGAAATACTCAAAGAAAGGAAACGGCCCCGCCATCAAGCAGCTGAAGTACCTGGATAAGAAACTGGGC
AGCCACATCGACATTACCCCCGCAAACGCCAATGGAAAACACGTGATCCTGCAGAGTCTGAAGCCTTGGAGAACCGACGTCTAT
CTGAACCACGAGACAGGCGAGTACGAAATCATGGGGATTAAGTATAGCGATCTGAAGTTCAACAAGAATGAGGGGTACGGAATC
AAGAAAGACAAGTATCTGGAAATTAAGAAAGTGGAGGAAGTCTCCGAGAAGTCTGAGTTCATGTTTAGCCTGTACAGGAAGGAT
CGCGTGAAAGTCCAGGACATGAAAACCGGCGAGTCCGTGGAACTGCTGTTCTGGAGCAGGAACTTTTCCAATAAGAAATACGCT
GAGCTGAAGCCCATCTCCCAGGCAGAAAACGACAAGAAACTGCCTGTGTATGGCAAAGGGAGACTGATCAAGAGGCTGATTCCC
AAAAACTGTAAGATCTGGAAAGTGAATACCACAATTCTGGGCGATCCCTACTATCTGGAGAAAGAAAGCGACTCCCCTAAGGAT
ATCCTGGACTAAGAATTC

#12
ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGAAGAAA
ACACTGGGACTGGACCTGGGGACTAACAGCATTGGGTGGGCCGTCATCAACTCAAACATCGACTCAGAAGGAAAAGAAAAGCTG
GTGGGGATCAGCTCCTGCGGAAGCCGGATCATTCCTATGGACGCTACCACACTGGGAGATTTCGGAAAGGGCAACACTAAAAGT
CCAGTGGCAGAGCGAACCCGACTGCGGGGCATTCGGAGACTGCTGGAAAGGTCACTGCTGAGGCGCGAGCGACTGCACCGAGTG
CTGTCAGTCATGGGGTTCCTGCCAGAGCATTACGCTAGCCAGCTGGACCGCTATGGAAAGTTTCTGCCAGAAACCGAGCCCAAG
CTGGCATGGTACAAAGACGATAGCGGCAGGTATCAGTTCCTGTTTCAGAAGTCCTTCCACGAGATGCTGGAGGACTTTCGACAG
CATCAGCCAGAACTGGTGGCAGGAGAGAAGAAAATCCCTTACGATTGGACAATCTACTATCTGCGGAAGAAAGCACTGTCTCAG
GAGATCACTAAGGAGGAACTGGCCTGGATTCTGCTGAATTTCAACCAGAAACGCGGCTACTATCAGCTGCGAGGGGAGGAAGAG
CAGGAAGAGAACAATAAGAGCGTGGAGTATCACGCCCTGAAAGTGGTGAGCGTCGAGGACTCTGGAGAAAGAAAGGGCAAAGAT
ATCTGGTACAATGTGACTCTGGAGAACGGATGGGTCTATCGACGGGCTAGCAACATTCCCCTGGACTGGACCGGCAAGGTGAAA
GAGTTCGTGGTCACTACCGAGCTGGACGATGCCGGGAACCCAAAGAAAGATAAGGAAGGAAATGTGAAAAGATCCTTTAGGGCA
CCAAAGGAGGACGATTGGGGGCTGCTGAAAACTAGGACCCAGGCTCAGATCGACGAATCCGGCAAGACCGTGGGAACTTACATC
TACGAGTCTCTGCTGTGTATGCCCAACCAGAAGATCCGGGGAAAACTGGTGAGGACCATTGAACGCAAGTACTATAAAGACGAG
CTGAGACAGATCCTGGTGAAGCAGAGCGAGTTCCACGCCGCTCTGCAGGATCATAATCTGCTGCTGTCCTGTATCGAAGAGCTG
TACCCTAACAATGAGGCCCACAGAAGGCTGCTGAGCGCCAGCAGCTTCATCTACTTCCTGATCGAGGACATTCTGTTTTATCAG
CGCCCACTGAAGTCCCAGAAAGGGCTGATCGATAACTGCCCCTACGAGTCTCACATCTACAAGGATAAGAAAGATGGAAGTCTG
CACCATGTGCCTCTGAAGTGTGTCAGCAAATCCCATCCACTGTTCCAGGAATTTCGCCTGTGGCAGTTCCTGTCTAACCTGCGA
ATCTACCAGAGAGAGAGGATGATGGACGGCAGTCTGAAACTGGACGTGGATGTCACCCGGGAGTTCCTGCCCTCAGAAGAGGAC
TACGTGAAGCTGTTTGATTGGCTGAATGAGAGAAAGGAAATCTCTCAGAAATTCCTGCTGGCTTATAAACCTTTTGGGCTGAAG
AAAAACGAACAGGCAAATTACAGATGGAACTATGTGGAGGACAAGAGCTACCCCTGCAACGAGACACGGGCAGAAATCAAGAGC
AGACTGTCCAAAGCCGGAGTGCCTGAAGAGTTTCTGACTGAAGAGAAGGAAGAGGCCCTGTGGCACATCCTGTATTCTATTAGT
GATAAGAAAGAGCTGACTAAGGCTCTGGGCACCTTCGCAGCCAAAAACTGTCTGAATGAGTCTTTCGTGGAAGTCTTTGCCAAG
ATCCCCCCTTTTGAGTCAAACTACGCTGCATATAGCCTGAAGGCTATTAGGAAACTGCTGGCACTGATGCGCATGGGGAAGTAC
TGGAATGAACAGGCCATCGACAGGCAGACTCGCGATCGAATCGAGAAAATTCTGACCGGAGAGTATGACGAAACAATCCGGAGC
AGAGTGAGGGAGAAGGCAATGCTGCTGACCGATATTAGCAGCTTCCGGGGCCTGCCTCTGTGGCTGGCCTGTTACATCGTGTAT
GACCGCCACTCAGAGAGCAAGGAACTGGTCAAATGGGAGACACCAGCCGACATCGATCATTTCCTGTCCAAGTTTAAACAGAAC
AGCCTGCGGAACCCCATCGTGGAACAGGTCATTACAGAGTCCCTGCGCACTGTGCGAGACATTTGGAAGCAGGAGGGAAAAATC
GATGAGATTCACGTGGAACTGGGCCGGGAGATGAAGAACCCTGCAAAAGAGCGCGCCCGAATTACAGCTCAGGTGCAGGAAAAT
GAGAACACTAATCTGAGAATCAAGGCTCTGCTGGCAGAGTTCATGAACCCCGAATTTGAGATTGAAAATGTGCATCCATACTCA
CCCGGCCAGCAGGAAATCCTGCGGATCTACGAGGACGGCGTGCTGAGCGGGATCGCTGAGAAGGATCTGCCTGACGATATCACA
GCAATTCTGAAGAAATTCCGAGAAAACGACGTGAAGAAACGGCCAACAACTAGCGAAGTCCTGCGGTACAAACTGTGGCTGGAG
CAGCGGTACAGATCCCCATATACCGGAAGAGTGATCCCCCTGGGCAAGCTGTTCACACCTGCTTACGAGATCGAACACGTGATT
CCCCAGAGCCGGTATTTTGACGATTCCATCTCTAACAAAGTGATCTGCGAAAGTGCCGTCAATAAGCTGAAAGATAACTGTCTG
GGCTATGAGTTCATCAAGAAACATTCCGGGGAGATGGTGGAACTGGGGAATGGAGAGACAGTGCCCGTGTTCAGCGTGGAAGAG
TACGAACGGTTCGTCAAGGAGTCTTACTTCGGCAACAGTAAGAAAATGAAGAAACTGCTGCTGGAGGACATCCCAGATAGCTTC
ATTGAGAGACAGCTGAATGACAGTCGATACATCTCACGGGTGGTCACATCTCTGCTGAGTAACCTGGTGTGCGAAGAGGGAGAG
CAGGATGGCCTGTCCAAGAATGTGATCGTCTGTACCGGCGGGATTACAGACAGGCTGAAGAAAGATTGGGGAGTGCAGGAAGTC
TGGAACCGCATCATTCTGCCTCGGTTCCTGAGACTGAATGAGATCACCGGACGGACAGACTTTACAAGTACTTCAGTGAACGGC
CACCTGCTGCCTGCCCTGCCACTGTACCTGCAGAAGGGCTTTAATAAGAAAAGAATTGACCATAGGCACCATGCCATGGATGCT
ATCGTGATTGCCTGCGCTAACCGGAATATCGTCAACTACCTGAACAATTCCTCTGCTAGAAAGAACAGCGAAATTAGCCGATAT
GACCTGCAGCGGCTGCTGTGTGAGAAGGTGAAAACCGATGCCAACGGCAATTACAAGTGGATCCTGAGGAAACCATGGGAGACA
TTCACCCAGGATGTGTATGCCGCTCTGACAAACATCGTGGTCAGCTTTAAGCAGAATCTGCGCGTGATTAATCGAACCACAAAC
TACTATCAGCACTACAACGAGGCAGGGGAAAAGCGCATGATCCCTCAGACCAAAGGCGACAGATGGGCCATTAGAAAGCCAATG
CATAAAGATACTGTGTATGGCGAAGTCAATCTGAGAAAGGAGAAAACCCTGCCACTGAGGGACGTGGTCAAGAACCCCAGTATC
GTGGTCGATAAATCACTGAAGAACAAGCTGTACGAGCTGCTGAAGAGCCAGTATGACCTGAAGGCAATCGCCAAATACTTCGAG
ACACACCAGGACGTGTGGGCAGATGTCAACCTGAAGAAAATTAAGGTGTACTACTTCACTAAGGAGACAAACGAAAGATTCTTT

FIG. 9H

GCCACTAGGAAGAGCCTGGACCCATCCTTTGATCAGAAGAAAATCGAAGAGGAAGTGACTGATACCGGCGTCCAGAAGATTCTG
CTGCACCATCTGCAGCAGAACAATAACGACCCTGATATGGCCTTTTCCCCAGACGGCATCGATAGGATGAACCAGAATATGACC
ATTCTGAATGACGGGAAGTGGCACCAGCCCATCTACAAAGTGCGCACATATGAAAAGGCAGATAAATTCGCCGTCGGCGAGTCT
GGGAACAAGGCCAAGAAATTTGTGGAAGCAGCCAAAGGCACCAATCTGTTCTTTGCCGTGTACGAGAGTGTCCAGGAGGACGAA
GCTATCGGGAAGCAGGTCTGCAAACGGACATTCGCCACTATCCCCCTGAACGAAGTGATCAAGAGAAAGAAACAGGGCCTGCCC
GCTGCACCTGAGGACCTGAACGGGAATCTGCCCAAGTTTGTGCTGAGCCCTAACGATCTGGTCTACCTGCCCACCGAGGAAGAG
AGGAATAGTTCACGCATCATTCAGCCTCTGGACAGGGAGCGCATCTATAAGATGGTGAGCTCCTCTGGGAGTCAGTGCTTCTTT
ATCAAAGTGTTCGTCGCCAATTCAATTTGGGATAAGAACGAATACAGCAGCCTGAACAAGATGGAGAGGGCTATTACAAACGAA
ATGATCAAGGAGATTTGTGTGCCTATCAAAATTGACCGCCTGGGCAATGTCAGCCTGATCCAGATTTAAGAATTC

#13
ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGACAAAG
ACTATTCTGGGACTGGATCTGGGGACTAACAGCATTGGATGGGCACTGATTAACCAGGACTTCGATAACAAAAAGGGGGAGATC
CTGGGCATGGGCAGCCGGATCATTCCTATGACACAGGATATTAAGGACGAGTTCGGAAAAGGCAACTCTATCAGTCAGACAGCT
GAGCGAACTCGACTGCGGGGAGTCCGGAGACTGATCCACAGAACTCTGCTGAGGCGCGAGCGGCTGCATAGAGTGCTGAATATC
ATTGGCTTCCTGCCAGAACACTATGCCAACCAGATTGACTTTGAGAAGAGGTTCGGCAAGTTCAAGCCTGAATCTGAGCCAAAA
ATCAGTTTCGATGGAAATGACGTGTTCCTGTTTGAGAAGAGCTACCAGGAAATGCTGGTCGACTTTAAAATCCACCAGCCACAG
CTGGTGAGCAACGGCAAGAAAATTCCCCATGACTGGACAATCTACTATCTGAGAAAGAAAGCTCTGTCTAAGAAAATTGAGAAG
GAGGAACTGGCATGGATCCTGCTGAACTTCAATCAGAAACGGGGATACTATCAGCTGAGAGGCGAGGAAGAGGAAGAGACACCA
AATAAGCTGGTGGAGTTTCATAGCCTGAAAATTGTGGAGGTCAACGCCGATGAACCCCAGAAAGGAAAGCCTGAGATCTGGTAC
TCACTGGTGCTGGAAAACGGCTGGGTCTATCGACGGGCTAGCAAGACTCCCCTGTTCGATTGGAAAGACAAGATTAGGGAGTTT
ATCGTGACCACAGAAATCAACAATGATGGAACTGTCAAAACCGACAAGGAGGGCACCGAAAAGAGGAGCTTCCGCGCACCAAAA
CCCGAGGACTGGACTCTGCAGAAGAAAAAGACCGAGTTCGAGCTGTCAAAGAGCGGGAATGAAGTGGGAGCCTTCATCTACGAA
AGTATCCTGCAGAAACCCAACCAGAAGATTAGAGGCAAACTGATTTCAACCATCGAGAGGAAATTTTATAAGGAAGAGCTGAAA
ACCATCCTGAAAACACAGCTGTTCTTTCACAAAGAACTGAAGGATGAGAAACTGTACAATGCCTGCATCGAAGAGCTGTATAAG
AACAATGAAGCTCACCGGAGCCTGCTGTCCAACAAGGGGTTCGAGCATCTGTTTATTAACGACATCCTGTTCTACCAGCGACCT
CTGCGGTCTAAAAAGAGTCAGATCTCAAACTGCCCACTGGAGAAGCGCACATATAAAAAGGAGGGGATTGAAATCACTGAGGGC
ATCAAAGTGATCTCCAAATCTAATCCAATCTACCAGGAGTTCCGGCTGTGGCAGTGGATTAGCAACCTGTCCCTGTATTGTATC
GAACCCACCGAGACAAATGTGACTTCAACCTTTCTGAACAGCATTGAAGATTACGAGAATCTGTTCGAATTTCTGAACAATCGC
AAGGAAATCGAGCAGAAGCACCTGCTGAAATATCTGCTGGAGAACCAGGGGTTTAAAGGAAAGCTGCTGACAAACGAACTGGAG
AAGTTCCGCTGGAATTTTGTCGCTGACAAAAAGTACCCCTGTAATGAGACAGGCAGCCTGCTGCATACTCGGCTGAGCAAAGTG
AAGAACATTTCCCCTGATTTCCTGACCAAGGAAATCGAGCACCAGCTGTGGCATATCATCTACAGCGTGACCGACAAGATTGAA
TATGAGCAGGCCCTGAAAACCTTTGCTCGGAAAAACAATCTGGATGTGGACTCCTTCTTTGAGCACTTCAAAAAGATCCCCCCT
TTTGAGTCTACCTACGGAGCATATAGTGAAAAGGCCATCAAGAAGCTGCTGCCACTGATCAGACTGGGCAAATACTGGAACTGG
GAGGCCATTGATAGTATCTCAAAGGACAGGATTAGTAAAATCCTGTCAGGGGAATACGATGAGAACATTAAGAACAGAGTGAGG
GAGAAAGCAGTCCACCTGACCTCCGAAAACAATTTCCAGGGACTGCAGGAGTGGCTGGCCAAGTACATCGTCTATGATCGCCAT
TCTGAGGGCAATGACCTGGGGAAGTGGACTAGCGTGTCCGACCTGGAGACATACCTGAAGGAGTTCAAGCAGCATAGCCTGCGG
AACCCTATTGTGGAGCAGGTCATCACAGAAACTCTGAGAGTGGTCAAGGATATTTGGATCAAGCACGGGAAAGGAACCGAAAAT
TTCTTTGACGAAATCCATGTGGAGCTGGGCCGGGAAATGAAGAACAATTCCGAGGATCGCAAACGACTGACCAACACAATTACT
GAAAACGAGAATACAAACCTGAGAATCAAGGCCCTGCTGATGGAAATGATGAATGATAACGACGTGGAGAACGTCAGGCCTTAC
TCTCCAAGTCAGCAGGAGATTCTGAAGATCTATGAGGACGGAGCTCTGAATAGCAACATCGAGCTGGACGATGAAATTGTGAAG
ATCTCCAAAAAGGCAGAGCCCACCAAATCTGAACTGCAGCGCTACAAGCTGTGGCTGGAGCAGAAATACCGATCCCCTTATACT
GGCCAGGTCATCCCACTGAACAAGCTGTTCACCTCTGAATATGAGATCGAACACGTGGTCCCTCAGAGTCGCTTCTTTGACGAT
AGCTTCAGCAACAAAGTGATCTGCGAGTCAGCCGTCAACAAGCGGAAGGATAACCAGCTGGGGCTGCAGTTCATCAAGAACCAT
AGCGGAGAAAAAGTGGAGCTGGGCTTCGGGAAGGTGGTCCAGGTCTTTACAGAAGAGCAGTACCTGGATTTTGTGAAGGAGCAC
TATAGCAAAAATCGCTCCAAGCATAACAAACTGCTGCTGGAAGAGATTCCCGAGAAGATGATCGAAAGGCAGCTGAATGACACT
CGCTACATCAGTAAGTTCGTGAGCTCCATTCTGTCCAACATCGTCAGATCTGAGAAGGACGATGACGGCCTGAATAGCAAAAAC
ATTCTGCCTGGAAATGGCAAGATCACTACCGAACTGAAAAGGGACTGGGGGCTGAATGATGTGTGGAACGACCTGATTCTGCCA
AGATTCGAGAGGATGAATACCATCACAAACAGCGATCTGTTTACAACTTACAACGACAAGTATCAGAAACACCTGCCCACCGTG
CCTTTCGAGTACTCCAAGGGCTTTCAGAAAAAGCGCATTGATCACCGACACCATGCTATGGACGCACTGGTCATCGCATGTGCC
ACACGGGATCATCTGAATCTGATGAACAATCAGTCTGCCAAGAGTGAACTGAAACGATACGACCTGCAGAACAAGCTGCGGAAA
AAGGAGCCTTACTTCAACCAGAAGGAGAACAAACAGAAGGAAGCCTTCAAGGATTTTATCAAACCATGGGGCACTTTCACCGAG
GACAGCAAGAATGCTCTGGAAAAAATCATTATCTCCTTTAAGCAGAACCTGAGAGTGATCAACAAGGCAACAAACTCATACGAG
AGCTATAAGGATGAGGACGGGAATCTGAGGATCGGCAAGGATGGGAAACCAGAGAAGGGCCTGATCAAGCAGAAGGGGCTGAAC
TGGGCAATCAGAAAGCCCCTGCACAAAGACACCGTGTCAGGCCAGATTAACCTGAGCAGGATCAAGCTGCCCAATGGGAAAATC
CTGACAGCCACTCGCAAGAATCTGGATACCAGTTTCGACCTGAAAACAATTGAGAACTCAATCACCGACACAGGCATTCAGAAA
ATCCTGAAGAATTACCTGCTGTCCAAGGAATCTCCAGAGCTGGCCTTCTCTCCCGAGGGCCTGGAAGAGATGAACAATGAGATC
GAAAAGTACAACAATGGGAAATTTCACCATCCTATTAACAAGGCTAGGATCTATGAACTGGGGAGCAAATTCAATGTCGGACAG
ACAGGCAACAGAAGGGATAAGTTTGTGGAGGCCGCTAAAGGAACTAATCTGTTCTTTGCTATCTACCAGGACGAGAATAAGAAC
CGCTCTTATGAAACTATTCCCCTGAACGAAGTGATCGAGCACCAGAAGTGGCGAGCAAGTCTGCCTAAGGAAGAGCAGGAGAAA
ATTCCACTGGTGCCCGTCAACAAGCTGAAGGGGACCTTCATCTTTTCCCTGTCTCCCAACGATCTGGTGTACGTCCCTTCCAAT
GACGAGCTGGAACGAAGTGCTTCAATTGACTTCTCTAAGCTGAAAAAGGAACAGATCAACCGGCTGTATAAAATGGTGTCTAGT
TCAGGATCCCAGTGCTTCTTTGTGAAGAGTGAGGTCGCAACCTCAGTGGTCAACAAAATGGAATACAGCAGCCTGAACAAGATG

FIG. 9I

GAGAAGTCTATCGATAACCTGATGGTGAAGGAGATTTGTATCAAACTGAAGATTGACAGGCTGGGCAACATCAGCAAGGCCTAA
GAATTC

#14
ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGACCAAA
ATCCTGGGGCTGGACCTGGGGACTAATAGCATCGGGTGGGCAATCCGCGACACAGAAAATGAGGGCATCAATCAGATCCTGGAC
AAGGGCGCCCGCATTTTCAGCGAGGGAGTGAAGTCCGAGAATGGCAAAGAAATCAGTAGGGCAGCTGAACGAACCGCTTACCGG
AGCGCAAGAAAGATCAAATATCGGAGAAAACTGCGGAAGTACGAGACACTGAAGGTGCTGTCAATTAACGGAATGTGCCCCCTG
AGCGTGGAGGAAGTCGAACAGTGGAAGAAATCTGGCTTCAAGGAATATCCCCTGAATCCTGAGTTTCTGGATTGGCTGCGGACC
AACGAGGACAAGAACATCAATCCTTACCTGTTCAGAGATAAAGCTAGCAAGCAGAAAGTGACACTGTTTGAACTGGGAAGAGCT
CTGTATCACATCGCACAGAGGCGCGGCTTCCTGAGTAACAGGCTGGACCAGTCAGCCGAGGGCGTCTTTGAGGAACATAATCCT
CAGATCCAGAACCTGATTGAGGACCTGGACAGCTCCAACACAATTCTGAATGAGCTGAAGGAATACTATATCAATCTGGGGATC
ATTGACGAGACTGAAAAGTCCGGCTTCAAGAAAGACCTGGATGAGGGGGAAAAGAAACTGAAGTCACTGTACAACAGCCTGGTG
GCCATCACAAAGAAAAACGCTAACGATATCGAAACTTGCAAGCAGGAGCTGATCGCTAGACTGAATAAGAAAGAGGACCTGGGC
AAGGTCAAAGGGAAGATCAAAGATATTTCTCAGGCAATGCTGGACAGGAACTTTAAGACTCTGGGACAGTATTTCTTTTCACTG
TACAACAAGGAGCGAATCCGGAACCAGTATACCAGCCGCGAGGAACACTACCTGGAGGAGTTCATCATTATCTGTCAGACCCAG
GGGATCGAGGGAATTAACACAAATGAGAAGCTGCCTGAGAAGAAGTTCACCGGCCTGGCAAAAGATCTGTATCGAGCCATTTTC
TTTCAGCGGCCACTGAAATCCCAGAAGGGCCTGATCGGGAAGTGCTCTTTCGAGAAGAACAAGTCTCGCTGTGCCGTGAGTCAC
CCAGACTATGAGGAATTTCGAATGTGGAGCTACCTGAACACAATCAAAATTGGCACTCAGTCCGAGAAAACCCTGCGCTTCCTG
ACACTGGAGGAAAAGCTGAAACTGGTGCCCAAGTTCTACCGAAAGAGTGATTTCAACTTCGAGGTGCTGGCTAAGGAGCTGGTC
GAAAAAGGAGCATCATTCGGCTACTACAAGTCTAGTAAGAAAAATGAGTTCTTTTACTGGTTCAACTATAAGCCCACCGATTCA
GTGAGCGCCTGCGTGGTGAGCGCTTCTCTGGAGAACGCAATCGGCAAGGACTGGAAGATCAAGACTTTCGAATATCAGACTAGA
AACACCGAGAAGAATGAAGTGACCAAATCCGTCGATTACAAGGACCTGTGGCACCTGCTGTCCGTGGCAACATCTGATACTTAC
CTGTATGACTTTGCCATCGAAAAGCTGAAACTGGAGCCTAAAAACGCAAAGGCCTTCAGCAAGACAAAACTGAAGAAAGACTTT
GCCAGTCTGTCACTGGCAGCCATCGCTAAAATTCTGCCATATCTGAAGCAGGGCCTGCTGTACTCCCACGCCGTGTTCATGGCT
AACATCGAGAATATTGTCGACGCCGATATCTGGAAGGATGAGGAACAGCAGAAGTTCATCCAGTCCAAGATTGTGGAGCTGGTC
GACAATTATATCGTGGAAAAGTCTAAACTGGAGCTGATCAATGGGCTGCTGAAGATCTACAACACCGAGGATAAGGAAGGACGG
AAAGTGTACTATTCAAAGGAGGCTGAAAGCCTGTTCGAGGCAGACCTGAGAAAGAAACTGGTCCCCTTCTACAAGGCTAACATC
ATCATCGAGGAACACGAGCAGGAAATCATTTTCCAGGATCTGTTTCCTATCCTGATGGACCAGCTGAAGAAACAGGAGTTCATC
AAAATTGGCAGACTGGATAAGCAGATTGAAGCCTTCCTGGAGGGGGAAAATGAGGAAGGACAGATCTTTTGTAACCACACAGAT
AAGCTGAAGAAACTGTACCATCCAAGCGACATCGAGGTGTTTAAGAAAAAGACTATCAAAGATGAGTGGGGGAATGAAAAGGTG
GTCCTGGGATCCCCACTGACCACATCTATCAAGAACCCCATGGCAATGAGAGCCCTGCACCAGCTGAGGAAGGTGCTGAATACA
CTGATTGCCAACGACCAGATCGACGAGGATACTCGGATCCATATTGAGATGGCCAGAGAACTGAACGATGCTAATAAGCGAAAA
GGCATCCAGGACTTCCAGAACGAGAACAAGAAGTTTCGGGAGGAAGCCATCAAGGAGATCAAGAAGCTGTACCTGGAGGAATGC
CACAAAGACGTGAACCCCACCGAGGACGATATCCTGAGGTATCAGCTGTGGCTGGAACAGGGAAAGTGCGAGATCTACGAGGAA
GGCAACAATATCAGCATTTGTGATATCATTGGCAGCAATCCCTCCTATGACATCGAGCATACCATTCCTCGGAGCATCTCCCAG
GATAACAGCCAGATGAACAAGACACTGTGTAGCCTGAAGTTCAACAGATCCATCAAAAAGCAGAAGATGCCAGTGGAGCTGTCC
AACTACAATGACATTCTGCCCAGGATCGCACACTGGAAAAAGGAGGCCGAGGAACTGACTAGGCAGATCGAAACCATTTCTCGC
AGCATCAAGAGCGCTGCAACCAAGGTGGCCAAGGATAAGAACATCAGAAAGAGGCATTACCTGACACTGAAGCGAGACTATATT
CTGGGGAAATACGAGCGGTTCACTTGGGAGGAACCTAAAGTGGGCTTTAAAAACTCCCAGATCCCAGACACTGGGATCATTACC
AAGTACGCTCAGGCATATCTGAAATCTTACTTCAAAAGGGTGGAGAGTGTCAAGGGAGGAGCAGTGGCTGAGTTCCGGAAACTG
TGGGGCATCCAGGAATCTTTTATCGATGAGAACTGGTGGAAGCACTATAAGGACAAAGATAGAGACAAACATACCCACCATACA
ATCGACGCAATCACTATTGCCTGTATGCCCAAGGATAAATACGACCTGCTGGCACACGCCTGGAGGCTGGAGGATGAACAGGAC
AAAAAGGCCGCTAAGGTGCTGATCGAGCAGGCCAAACCATGGAAAACCTTCAAGGAGGATATCGAAAAGATTGAGACTGAAATC
CTGGTGAGCCATTTTACCCCCGACAACGTCAAAAAGCAGTCAAAGAGCATCATCAAGAATCGCGGCAAAAAGGTGTACGTCCTG
AAGAACGAGCTGCCTGTGAACTTCAAGAACAAGATCGAAGGGAAGGATTATTTCAAGCTGAAATTTGACAGCAAGATTCTGTAC
AAAATCCCCAAAAAGAAAGAGAAGCAGACCGATACATTCTATGAGGAACTGCCTAAAAACTACCTGAATGGAGTGGAAGGCAAG
GACTACTTCAAAATCAATACTACCGGCAAGACCTTCTACAAAATCCCAATTTTTAACCAGGGCGACACAATCCGGGGGAGCCTG
CACCAGGAGACAACTTACGGAGCCATTAAGCTGCCAGATATCGACATTGAAACAAAGAAACCCCTGCATACTGATAAGGGAGGC
TTCATTCTGAAGAAAGACATCAAAGGCAATGAGATTGTGTTCTTTGTGGTCCGCAAGGAAATCTCTAAAATTAGTGAGAACGAT
GTGCAGAATATCGTCGACAACGTGGTCCGGAAGAAAATTGAGAATGCTATCGCAAACTCCCTGATCACTTTTAAGATCGTGAAG
AAAAAGAAAGTGGCCGTCATCAAAGAGGGGGTCACCATTTGGATGAGGGAGCCCAACATTGAAAAGGGGATCGAGGGAATCCCT
ATTAAGAAAGTGCGCATCATTACCAATTCTGTCAAAAACCCTATCGAGATTAAGGTGCACAGTCCACTGTCCAAGTCTCGCCAC
AAGCATAAACAGAAGGTCTACGGCCAGAACGATGAGAATTATGCCATGGCTCTGTACGAACTGGACGGGAAGCGGGAGTTCGAA
CTGATCAACAATTTTAATCTGGCCAAGCTGCTGAAACAGAGTCAGTCATACTATCCACTGCATAAAGAGAAGGAAATCAAGGGA
AAGAAAATTCTGGTGCCCATCGAGAAGAGAAACAACAAGGATGTCATCCTGAAGAGGGGCCAGCAGGTGGTGTTCTACGACAAA
ACCGTGGAGAACCCCAAGGATATCTCTGAAATCATTGACTTTCGCGAGCGAATCTACATCATTGAAGGGCTGACCATCCAGAGA
CAGAAAGATAAGAAAACATCCAAAGTGAATGAGTACGGAATCATTCAGCTGCGCCACTTCAAGGAAGCTCGAAAAAGTGAGGAA
ATCTCAAAGGATAACTTCAAACCTGACGGCGAGTTCAAGATCAACGAGAATAAGCCAACTAGGAAAATGAACCATAATCAGTTC
ACCGCCTTTGTGGAGGGGATCGACTTCAAGGTCCTGCCTAGCGGAAAGTTTCAGAAAATTTAAGAATTC

#15
ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGAGCAAG

FIG. 9J

AAAGTCAGTAGACGCTATGAAGAACAGGCACAGGAGATTTGTCAGAGGCTGGGGAGTAGACCTTATTCCATCGGGCTGGACCTG
GGAGTGGGATCTATCGGAGTGGCAGTCGCCGCTTACGACCCAATTAAGAAACAGCCCTCCGATCTGGTGTTCGTCAGCTCCAGG
ATTTTTATCCCTTCTACCGGCGCAGCCGAGCGGAGACAGAAGAGAGGACAGAGGAACAGCCTGCGCCACCGAGCAAATCGCCTG
AAATTCCTGTGGAAGCTGCTGGCTGAACGAAACCTGATGCTGTCATATAGCGAGCAGGACGTGCCAGATCCTGCACGGCTGAGA
TTTGAGGACGCTGTGGTCCGAGCAAACCCATACGAGCTGCGGCTGAAGGGCCTGAATGAACAGCTGACCCTGAGCGAGCTGGGG
TATGCACTGTACCACATCGCCAATCATAGGGGATCTAGTTCAGTGCGCACATTCCTGGATGAGGAAAAGAGCTCCGACGATAAG
AAACTGGAGGAACAGCAGGCTATGACAGAACAGCTGGCAAAAGAGAAGGGAATTTCCACTTTCATCGAAGTGCTGACAGCCTTT
AACACTAATGGCCTGATCGGGTACAGGAACTCCGAGTCTGTGAAGAGTAAGGGCGTGCCAGTCCCCACTCGCGACATCATTTCA
AATGAGATTGATGTGCTGCTGCAGACCCAGAAGCAGTTCTATCAGGAAATCCTGTCAGACGAGTACTGCGATCGGATTGTCAGC
GCAATCCTGTTTGAAAACGAGAAGATCGTGCCAGAAGCCGGCTGCTGTCCCTATTTCCCTGACGAGAAGAAACTGCCCAGATGT
CACTTTCTGAATGAGGAAAGGCGCCTGTGGGAAGCCATTAACAATGCTAGGATCAAGATGCCCATGCAGGAGGGCGCTGCAAAA
CGCTACCAGAGTGCTTCATTCAGCGACGAGCAGAGACACATTCTGTTTCATATCGCAAGGAGCGGGACTGATATCACCCCTAAA
CTGGTGCAGAAGGAGTTCCCAGCCCTGAAAACCTCCATCATTGTGCTGCAGGGAAAAGAGAAGGCTATTCAGAAGATCGCAGGC
TTCCGATTTCGACGGCTGGAGGAAAAATCTTTTTGGAAGAGACTGAGTGAGGAACAGAAGGACGATTTCTTTAGCGCCTGGACA
AACACTCCTGACGATAAAAGACTGTCCAAGTACCTGATGAAACACCTGCTGCTGACAGAAAATGAGGTGGTCGACGCCCTGAAA
ACCGTGAGCCTGATTGGAGATTATGGCCCAATCGGCAAGACCGCAACACAGCTGCTGATGAAACATCTGGAGGATGGCCTGACT
TACACCGAAGCCCTGGAGCGGGGAATGGAAACCGGCGAGTTCCAGGAACTGTCAGTGTGGGAGCAGCAGAGCCTGCTGCCCTAC
TATGGGCAGATCCTGACAGGATCTACTCAGGCCCTGATGGGGAAGTATTGGCACTCTGCTTTTAAAGAAAAGAGAGACAGTGAG
GGATTCTTTAAGCCTAACACAAATAGCGATGAGGAAAAATACGGCAGGATTGCCAACCCAGTGGTCCATCAGACTCTGAACGAA
CTGCGCAAGCTGATGAATGAGCTGATTACCATCCTGGGAGCTAAACCTCAGGAGATCACAGTGGAACTGGCACGAGAGCTGAAG
GTCGGAGCTGAGAAAAGAGAGGACATCATTAAGCAGCAGACCAAACAGGAAAAGGAGGCTGTGCTGGCATATAGCAAGTACTGC
GAGCCCAACAATCTGGACAAAAGGTATATTGAAAGGTTCCGCCTGCTGGAGGATCAGGCCTTTGTGTGCCCTTACTGTCTGGAG
CACATTAGCGTCGCAGATATCGCAGCTGGAAGGGCAGACGTGGATCATATCTTCCCACGCGACGATACAGCTGACAACTCCTAC
GGGAATAAGGTGGTCGCACACCGACAGTGTAACGATATCAAGGGAAAGCGGACCCCCTATGCAGCCTTCAGTAATACATCAGCC
TGGGGCCCCATCATGCATTACCTGGACGAAACCCCTGGGATGTGGCGCAAAAGAAGGAAGTTTGAGACAAACGAGGAAGAGTAT
GCTAAGTACCTGCAGTCAAAAGGCTTCGTGAGCAGGTTTGAAAGCGATAACTCCTATATCGCAAAAGCTGCAAAGGAGTACCTG
CGCTGCCTGTTCAATCCAAACAATGTGACTGCCGTCGGGTCCCTGAAGGGAATGGAGACATCTATCCTGCGGAAGGCCTGGAAT
CTGCAGGGAATTGACGATCTGCTGGGCAGCCGGCACTGGAGTAAGGACGCCGATACCAGCCCCACAATGCGCAAAAACCGGGAC
GACAATCGGCACCATGGCCTGGACGCCATCGTGGCTCTGTATTGTTCCAGATCTCTGGTCCAGATGATTAACACCATGTCCGAG
CAGGGCAAGCGAGCAGTGGAAATCGAGGCTATGATTCCTATCCCAGGGTACGCATCCGAACCAAATCTGTCTTTCGAAGCCCAG
CGGGAGCTGTTTAGAAAGAAAATCCTGGAGTTCATGGACCTGCACGCCTTTGTGAGTATGAAAACCGACAACGATGCAAATGGC
GCCCTGCTGAAAGATACTGTGTATTCAATTCTGGGAGCAGACACCCAGGGAGAGGATCTGGTGTTCGTGGTCAAGAAAAAGATT
AAGGACATCGGCGTGAAAATCGGGGATTATGAAGAGGTCGCATCTGCCATTCGAGGCCGGATCACCGACAAACAGCCAAAGTGG
TATCCCATGGAAATGAAAGATAAGATCGAGCAGCTGCAGTCTAAGAACGAAGCCGCTCTGCAGAAATACAAGGAGAGTCTGGTG
CAGGCAGCCGCTGTCCTGGAAGAGAGTAATAGGAAGCTGATTGAGTCAGGCAAAAAGCCCATCCAGCTGAGTGAAAAAACAATT
TCAAAAAAGGCCCTGGAGCTGGTGGGCGGGTACTATTACCTGATTAGCAACAACAAGCGCACAAAGACTTTCGTGGTCAAGGAA
CCTTCAAACGAGGTGAAAGGGTTCGCATTTGACACTGGAAGCAATCTGTGCCTGGACTTTTATCACGATGCCCAGGGAAAGCTG
TGTGGCGAGATCATTAGAAAAATCCAGGCTATGAACCCTTCCTATAAGCCAGCATACATGAAACAGGGGTATTCTCTGTACGTG
AGACTGTACCAGGGCGACGTGTGCGAGCTGAGGGCAAGCGATCTGACTGAAGCAGAGTCCAACCTGGCCAAGACCACACATGTC
CGCCTGCCCAATGCTAAACCTGGGCGAACCTTCGTGATCATTATCACCTTTACAGAGATGGGGTCTGGATATCAGATCTACTTC
AGCAACCTGGCCAAGTCCAAAAAGGGACAGGACACTAGTTTTACCCTGACTACCATCAAGAATTATGATGTGCGGAAAGTCCAG
CTGTCTAGTGCCGGGCTGGTGAGATACGTCAGCCCTCTGCTGGTGGACAAAATCGAGAAGGATGAAGTCGCTCTGTGTGGAGAG
TAAGAATTC

#16

ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGGCAAGA
CCTGCATTTCGGGCACCTCGGAGAGAACACGTCAACGGCTGGACCCCTGACCCACATCGGATTAGCAAACCCTTTTTCATCCTG
GTGAGCTGGCACCTGCTGTCCCGGGTGGTCATTGACAGCTCCTCTGGATGCTTCCCAGGCACCAGCCGGGACCACACCGACAAG
TTTGCCGAGTGGGAATGTGCTGTGCAGCCCTACAGGCTGAGTTTCGACCTGGGGACCAACTCAATCGGATGGGGCCTGCTGAAT
CTGGATCGCCAGGGAAAACCAAGGGAGATCCGAGCACTGGGGTCCCGCATCTTCAGCGACGGACGGGATCCCCAGGACAAGGCT
TCTCTGGCTGTGGCACGGAGACTGGCCAGACAGATGAGGCGCCGACGGGACAGATATCTGACTAGAAGGACCAGGCTGATGGGA
GCTCTGGTGCGCTTCGGCCTGATGCCAGCAGACCCTGCAGCTAGGAAGCGCCTGGAAGTGGCCGTCGATCCATACCTGGCACGA
GAGCGGGCCACAAGAGAAAGGCTGGAGCCCTTCGAAATCGGGAGGGCCCTGTTTCACCTGAACCAGCGCCGAGGATATAAACCC
GTGCGCACCGCCACAAAGCCTGATGAGGAAGCCGGCAAGGTGAAAGAGGCTGTCGAAAGGCTGGAGGCAGCAATCGCTGCAGCC
GGAGCACCTACTCTGGGAGCTTGGTTCGCATGGCGAAAAACACGAGGAGAAACTCTGCGAGCACGACTGGCTGGGAAGGGAAAA
GAGGCTGCATACCCATTCTATCCCGCACGGAGAATGCTGGAGGCCGAATTTGACACTCTGTGGGCAGAGCAGGCCAGGCACCAT
CCAGATCTGCTGACCGCCGAAGCTCGCGAGATCCTGCGGCACAGAATTTTTCATCAGCGGCCCCTGAAGCCACCTCCAGTGGGA
AGATGCACTCTGTACCCTGACGATGGGAGAGCTCCTAGGGCACTGCCAAGCGCTCAGAGGCTGCGCCTGTTCCAGGAGCTGGCC
AGCCTGAGAGTGATCCACCTGGACCTGTCCGAACGCCCTCTGACCCCAGCTGAGCGAGATCGGATTGTGGCATTTGTCCAGGGC
AGACCCCCTAAAGCCGGAAGGAAGCCTGGCAAAGTGCAGAAGAGCGTCCCATTCGAAAAGCTGAGGGGGCTGCTGGAGCTGCCA
CCAGGCACTGGGTTTTCTCTGGAGAGTGACAAACGGCCTGAACTGCTGGGCGACGAGACAGGCGCCAGAATCGCACCAGCATTC
GGACCTGGATGGACAGCTCTGCCTCTGGAGGAACAGGACGCCCTGGTGGAACTGCTGCTGACAGAGGCAGAACCAGAGAGGGCA
ATTGCAGCTCTGACTGCACGATGGGCTCTGGACGAGGCAACTGCAGCAAAGCTGGCTGGAGCAACCCTGCCAGATTTTCACGGA

FIG. 9K

CGATATGGCAGGCGCGCAGTGGCTGAACTGCTGCCTGTCCTGGAACGCGAGACACGAGGCGACCCAGATGGGAGAGTGAGGCCC
ATCCGGCTGGACGAGGCAGTGAAACTGCTGAGAGGCGGGAAGGATCACTCAGACTTCAGCCGGGAAGGAGCTCTGCTGGACGCA
CTGCCCTACTATGGAGCCGTGCTGGAGAGACATGTCGCTTTTGGGACAGGAAACCCCGCAGATCCTGAGGAAAAGCGGGTGGGA
AGAGTCGCCAATCCCACTGTGCACATCGCTCTGAACCAGCTGCGGCATCTGGTGAATGCAATTCTGGCCAGGCACGGCCGCCCT
GAGGAAATCGTGATTGAGCTGGCACGGGACCTGAAAAGGTCTGCCGAAGATCGACGGAGAGAGGACAAGCGGCAGGCCGATAAC
CAGAAAAGAAATGAGGAACGCAAGCGACTGATCCTGAGTCTGGGAGAGCGCCCAACCCCACGAAACCTGCTGAAGCTGCGGCTG
TGGGAGGAACAGGGCCCAGTGGAAAATAGGCGCTGCCCCTACTCTGGGGAGACAATTAGTATGAGAATGCTGCTGAGCGAGCAG
GTGGACATCGATCACATTCTGCCATTCAGCGTGTCCCTGGACGATTCCGCTGCAAACAAGGTGGTCTGTCTGCGGGAGGCCAAC
AGAATCAAGCGGAATAGATCTCCCTGGGAGGCCTTCGGCCATGACAGTGAGAGATGGGCAGGGATTCTGGCACGAGCAGAAGCT
CTGCCCAAGAACAAAAGGTGGCGCTTTGCTCCTGACGCACTGGAGAAACTGGAAGGAGAGGGAGGCCTGCGAGCACGACACCTG
AATGATACAAGGCATCTGAGTCGCCTGGCCGTGGAGTATCTGCGGTGCGTCTGTCCTAAGGTGCGGGTGAGCCCAGGCCGACTG
ACTGCACTGCTGCGACGGAGATGGGGCATCGACGCCATTCTGGCAGAAGCAGATGGACCTCCACCAGAAGTGCCCGCAGAGACA
CTGGACCCTTCCCCAGCTGAGAAGAACCGAGCAGACCACCGGCACCATGCCCTGGATGCTGTGGTCATCGGCTGTATTGATCGC
TCAATGGTGCAGCGAGTCCAGCTGGCCGCTGCAAGCGCAGAAAGAGAGGCCGCTGCAAGGGAGGACAATATCAGGCGCGTGCTG
GAGGGATTCAAAGAGGAACCTTGGGATGGCTTTAGAGCTGAACTGGAGCGACGGGCACGGACCATCGTGGTGAGCCACAGACCA
GAACATGGGATTGGGGGAGCCCTGCATAAGGAGACAGCTTACGGGCCTGTGGACCCTCCAGAGGAAGGATTCAACCTGGTGGTC
AGGAAACCAATCGACGGCCTGTCAAAGGATGAGATTAATAGCGTGCGGGACCCCCGGCTGAGAAGGGCACTGATCGATCGCCTG
GCCATTCGCCGACGGGATGCTAACGACCCTGCTACCGCACTGGCCAAAGCAGCTGAGGATCTGGCAGCACAGCCAGCCTCCCGC
GGCATCAGAAGGGTGCGGGTCCTGAAGAAAGAATCTAACCCCATTAGGGTGGAGCACGGCGGGAATCCAAGTGGACCCCGCTCA
GGAGGCCCTTTTCATAAGCTGCTGCTGGCAGGAGAGGTGCACCATGTGGACGTCGCACTGCGAGCAGATGGCCGCCGATGGGTG
GGACACTGGGTCACACTGTTCGAGGCACATGGGGGACGGGGAGCAGACGGAGCTGCAGCCCCACCTAGACTGGGCGATGGGGAA
AGATTTCTGATGAGGCTGCACAAGGGAGACTGCCTGAAACTGGAGCATAAGGGCAGAGTGAGGGTCATGCAGGTGGTCAAACTG
GAACCTAGTTCAAATAGCGTGGTCGTGGTCGAGCCACACCAGGTGAAAACCGACAGATCCAAACATGTCAAGATCTCTTGTGAT
CAGCTGCGCGCTCGAGGAGCACGGAGAGTGACCGTCGATCCACTGGGACGGGTGAGAGTCCACGCCCCAGGAGCTAGGGTGGGA
ATCGGAGGGGACGCCGGACGAACCGCTATGGAACCCGCAGAGGATATTAGCTAAGAATTC

#17
ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGGGAGAG
AATATGATTGACGAGAGTCTGACCTTCGGCATTGACCTGGGGATTGGGAGTTGTGGGTGGGCAGTGCTGAGGCGGCCAAGCGCC
TTCGGACGGAAGGGCGTGATCGAGGGAATGGGCTCCTGGTGCTTTGACGTCCCCGAGACATCTAAAGAACGGACTCCTACCAAC
CAGATCCGGAGATCCAATAGACTGCTGAGGCGCGTGATCCGACGGAGAAGGAACCGGATGGCCGCTATTCGCCGACTGCTGCAC
GCAGCAGGCCTGCTGCCATCAACCGACAGCGATGCCCTGAAGCGGCCCGGACATGATCCTTGGGAACTGAGGGCACGAGGCCTG
GACAAGCCACTGAAACCCGTGGAGTTCGCTGTGGTCCTGGGCCACATCGCAAAGCGGAGAGGGTTTAAATCCGCTGCAAAGAGA
AAAGCCACAAACATTAGCTCCGACGATAAGAAAATGCTGACAGCCCTGGAGGCTACTCGAGAACGGCTGGGGAGATACAGGACC
GTGGGAGAAATGTTCGCCAGGGACCCTGATTTTGCTTCTAGGCGCCGAAATCGCGAGGGCAAGTATGATAGGACCACAGCTCGC
GACGATCTGGAGCACGAAGTGCACGCCCTGTTCGCAGCTCAGCGGAGACTGGGACAGGGATTTGCCAGTCCAGAACTGGAGGAG
GCCTTCACCGCTTCAGCATTTCATCAGAGGCCCATGCAGGACAGCGAGCGCCTGGTGGGATTCTGCCCTTTTGAGCGAACCGAA
AAGCGGGCAGCCAAACTGACACCCTCTTTCGAGCGCTTTCGACTGCTGGCCCGGCTGCTGAACCTGAGAATCACTACCCCAGAC
GGAGAGCGGCCCCTGACAGTGGATGAAATTGCTCTGGTCACCCGGGACCTGGGCAAGACCGCAAAACTGAGTATCAAGCGGGTG
AGAACTCTGATTGGACTGGAGGACAATCAGAGGTTCACAACTATCCGCCCCGAGGACGAAGATCGAGACATTGTGGCTCGGACA
GGCGGGGCAATGACAGGGACTGCCACCCTGAGGAAGGCACTGGGAGAGGCCCTGTGGACTGATATGCAGGAGCGCCCTGAACAG
CTGGACGCTATCGTGCAGGTCCTGAGCTTCTTTGAGGCCAACGAAACAATCACTGAGAAGCTGAGGGAAATTGGCCTGACTCTG
GCCGTGCTGGACGTCCTGCTGACCGCACTGGATGCCGGAGTGTTCGCCAAGTTTAAAGGCGCTGCACACATCAGCACCAAAGCC
GCTAGGAATCTGCTGCCACATCTGGAGCAGGGCAGGCGCTACGATGAGGCCTGCACAATGGCAGGGTATGACCACGCAGCCTCC
CGCCTGTCTCACCATGGCCAGATCGTGGCAAAGACACAGTTCAACGCCCTGGTCACTGAGATCGGCGAATCCATTGCCAATCCA
ATCGCTCGGAAGGCACTGATCGAGGGGCTGAAACAGATTTGGGCCATGAGAAACCACTGGGGGCTGCCCGGAAGTATCCATGTG
GAGCTGGCCCGGGATGTCGGCAACTCAATTGAAAAGCGACGGGAGATTGAAAAGCACATCGAGAAAAATACTGCCCTGAGGGCT
CGCGAGAGAAGGGAGGTGCATGATCTGCTGGACCTGGAAGATGTCAATGGCGACACCCTGCTGCGATACCGGCTGTGGAAGGAG
CAGGGAGGCAAATGCCTGTATACAGGGAAGGCCATCCACATTCGGCAGATCGCTGCAACTGACAACTCCGTGCAGGTCGATCAT
ATTCTGCCTTGGAGCCGGTTCGGCGACGATAGTTTTAACAACAAGACCCTGTGTCTGGCCTCTGCTAATCAGCAGAAGAAAAGG
TCAACACCATACGAGTGGCTGAGCGGCCAGACTGGGGATGCATGGAACGCCTTCGTGCAGCGCATCGAGACAAATAAGGAACTG
AGAGGGTTTAAGAAAAGGAACTATCTGCTGAAGAATGCTAAAGAGGCAGAGGAAAAATTCAGAAGCAGGAACCTGAATGACACC
AGATACGCCGCTAGGCTGTTCGCAGAGGCCGTGAAGCTGCTGTATGCCTTTGGGGAGAGACAGGAAAAAGGGGGAAACCGCCGA
GTGTTTACTCGGCCTGGAGCACTGACCGCAGCACTGAGACAGGCTTGGGGAGTGGAGAGCCTGAAGAAACAGGATGGGAAGCGC
ATCAATGACGATCGACACCATGCCCTGGATGCTCTGACCGTGGCTGCAGTCGACGAGGCCGAAATTCAGAGGCTGACAAAATCA
TTCCACGAGTGGGAACAGCAGGGCCTGGGGCGGCCTCTGCGGAGAGTGGAGCCACCTTGGGAGAGCTTCCGGGCAGACGTCGAG
GCTACCTACCCTGAAGTGTTTGTCGCACGGCCAGAGAGGCGCCGAGCAAGAGGAGAAGGCCATGCCGCTACCATCCGGCAGGTG
AAGGAGAGAGAATGCACACCAATTGTGTTTGAGAGAAAGGCTGTCTCTAGTCTGAAAGAGGCAGACCTGGAACGAATCAAAGAT
GGCGAGCGCAACGAAGCAATTGTGGAGGCCATCAGGAGCTGGATTGCCACTGGACGCCCAGCTGATGCACCACCACGCTCCCCC
CGAGGCGACATCATTACCAAGATCAGGCTGGCCACCACCATCAAGGCAGCCGTGCCTGTCCGCGGAGGGACCGCAGGAAGGGGA
GAAATGGTGCGCGCAGATGTGTTCAGCAAGCCAAACCGGAGAGGGAAAGACGAGTGGTATCTGGTGCCCGTGTATCCACACCAG
ATCATGAACAGGAAGGCTTGGCCCAAACCTCCAATGCGCTCAATTGTGGCCAATAAGGATGAGGACGAATGGACCGAAGTGGGA
CCTGAACACCAGTTCCGGTTTAGCCTGTACCCTAGATCCAATATCGAGATCATTAGGCCATCTGGAGAAGTGATCGAAGGATAT

FIG. 9L

TTCGTCGGCCTGCATCGCAACACTGGCGCTCTGACCATCAGTGCACACAATGATCCCAAGAGTATCCATTCAGGCATTGGGACC
AAGACACTGCTGGCCATTTCCAAATACCAGGTGGACAGATTCGGCAGAAAGTCTCCAGTGCGCAAAGAGGTCCGAACTTGGCAC
GGGGAAGCCTGTATCTCTCCCACCCCCCCTGGATAAGAATTC

#19

ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGAGAGAG
AACGGGAGTGACGAAAGGCGGAGGAATATGGACGAAAAAATGGATTACAGGATTGGGCTGGACATCGGCATCGCATCAGTGGGG
TGGGCCGTCCTGCAGAACAATTCAGACGATGAGCCTGTGAGGATTGTCGACCTGGGAGTGCGCATTTTCGATACCGCAGAGATC
CCAAAGACAGGCGAAAGCCTGGCAGGACCCCGGAGAGCAGCTCGAACCACAAGGCGCCGACTGCGGAGAAGGAAACACCGCCTG
GACCGAATCAAGTGGCTGTTCGAGAACCAGGGGCTGATCAATATTGACGATTTTCTGAAGAGATACAACATGGCCGGACTGCCA
GATGTGTACCAGCTGCGGTATGAGGCTCTGGACAGAAAACTGACCGATGAGGAACTGGCCCAGGTGCTGCTGCACATCGCTAAG
CATCGGGGCTTCAGAAGCACCAGGAAAGCCGAGACAGCAGCCAAGGAAAACGGGGCAGTGCTGAAGGCCACAGACGAGAATCAG
AAACGAATGCAGGAAAAGGGATACAGGACAGTGGGCGAGATGATCTACCTGGACGAGGCCTTCCGGACTGGCTGCTCTTGGAGT
GAGAAAGGGTACATCCTGACCCCCCGCAACAAGGCTGAAAATTATCAGCACACAATGCTGCGGGCAATGCTGGTGGAGGAAGTC
AAGGAGATTTTCAGCTCCCAGCGCCGACTGGGCAACGAAAAAGCCACTGAGGAACTGGAGGAAAAGTACCTGGAGATCATGACC
TCCCAGCGCTCTTTTGACCTGGGGCCTGGAATGCAGCCAGATGGGAAGCCCTCCCCTTATGCAATGGAGGGCTTCTCTGACAGA
GTGGGGAAATGTACTTTTCTGGGGGATCAGGGAGAGCTGAGGGGCGCTAAGGGGACCTACACAGCCGAATATTTCGTGGCTCTG
CAGAAAATCAACCACACAAAGCTGGTCAATCAGGACGGCGAGACAAGGAATTTCACTGAGGAAGAGCGGAGAGCCCTGACTCTG
CTGCTGTTTACCCAGAAAGAGGTGAAGTACGCTGCAGTCCGCAAGAAACTGGGCCTGCCTGAGGACATCCTGTTCTACAACCTG
AACTACAAGAAGGCCGCTACTAAAGAAGAGCAGCAGAAGGAGAACCAGAATACCGAAAAAGCCAAGTTTATCGGGATGCCATAC
TATCACGATTACAAGAAATGCCTGGAAGAGAGAGTGAAGTATCTGACCGAGAACGAAGTCAGGGACCTGTTTGATGAGATCGGA
ATGATTCTGACTTGTTACAAAAATGACGATTCCCGCACCGAACGACTGGCCAAGCTGGGACTGGTGCCCATCGAGATGGAAGGC
CTGCTGGCTTATACTCCTACCAAATTCCAGCATCTGTCTATGAAGGCAATGCGGAACATCATTCCCTTTCTGGAGAAAGGGATG
ACCTACGACAAGGCTTGCGAAGAGGCAGGATATGACTTCAAAGCCGATAGCAAGGGGACTAAACAGAAGCTGCTGACCGGAGAG
AACGTGAATCAGACAATCAACGAAATTACTAATCCTGTGGTCAAACGCTCAGTGAGCCAGACAGTGAAGGTCATTAACGCCATC
ATTCGGACTTACGGCAGTCCACAGGCTATCAATATTGAGCTGGCAAGAGAAATGTCAAAGACCTTTGAAGAGAGGCGCAAAATC
AAGGGGGACATGGAGAAACGGCAGAAGAACAATGAAGATGTGAAGAAACAGATTCAGGAGCTGGGAAAACTGTCTCCTACAGGC
CAGGACATCCTGAAGTACAGACTGTGGCAGGAGCAGCAGGGGATTTGTATGTATAGTGGAAAAACCATCCCACTGGAAGAGCTG
TTCAAGCCCGGCTACGACATCGATCACATTCTGCCCTATTCAATTACATTCGACGATAGCTTTAGGAACAAAGTGCTGGTCACA
TCCCAGGAGAACAGACAGAAGGGCAATAGGACTCCTTACGAGTATATGGGGAACGACGAACAGCGCTGGAATGAGTTTGAAACC
AGGGTGAAAACTACCATCCGCGATTACAAGAAACAGCAGAAGCTGCTGAAGAAACATTTCTCTGAAGAGGAAAGGAGTGAGTTT
AAAGAACGGAACCTGACAGACACTAAGTACATCACAACCGTGATCTACAACATGATCAGACAGAATCTGGAGATGGCCCCCCTG
AACCGCCCTGAAAAGAAAAAGCAGGTGCGGGCTGTCAATGGCGCAATTACCGCCTACCTGCGAAAACGGTGGGGGCTGCCACAG
AAGAATCGGGAGACAGACACACACCATGCTATGGATGCAGTGGTCATCGCCTGCTGTACCGACGGCATGATCCAGAAAATTAGT
AGATACACAAAGGTGAGAGAGAGGTGCTATTCAAAGGGAACAGAGTTCGTCGATGCAGAGACTGGCGAAATCTTTAGACCCGAG
GACTACAGCAGGGCCGAGTGGGATGAAATTTTCGGCGTGCACATCCCAAAGCCCTGGGAGACATTTCGCGCCGAACTGGACGTC
CGAATGGGGGACGATCCAAAGGGATTCCTGGACACTCATAGCGATGTGGCTCTGGAGCTGGATTATCCCGAGTACATCTACGAA
AACCTGCGGCCTATCTTCGTGAGCAGAATGCCAAATCACAAGGTCACCGGAGCAGCCCATGCTGACACAATTCGGTCCCCAAGA
CACTTTAAAGATGAGGGCATCGTGCTGACTAAGACCGCACTGACCGACCTGAAACTGGACAAGGATGGGGAGATCGACGGATAC
TATAACCCCCAGTCCGATCTGCTGCTGTACGAAGCACTGAAAAAGCAGCTGCTGCTGTATGGCAATGATGCCAAAAAGGCCTTC
GCTCAGGACTTTCATAAACCCAAGGCCGATGGAACTGAGGGCCCTGTGGTCAGGAAGGTGAAGATCCAGAAAAAGCAGACCATG
GGAGTGTTCGTCGACTCTGGCAACGGGATTGCCGAGAATGGCGGGATGGTGCGCATCGATGTGTTCCGAGTCAACGGCAAGTAC
TATTTTGTGCCCGTCTACACCGCTGACGTGGTCAAAAAGGTGCTGCCTAATAGGGCCAGTACAGCTCACAAGCCATACGGCGAG
TGGAAAGTGATGGAGGACAAGGATTTCCTGTTTAGTCTGTATTCACGCGACCTGATCCATATCAAGTCTAAAAAGGATATCCCT
ATTAAGATGGTGAACGGAGGCATGGAGGGGATCAAGGAAACCTACGCATACTATATTGGAGCCGACATCAGCGCTGCAAATATC
CAGGGCATTGCCCACGATTCCAGGTATAAATTCCGCGGACTGGGCATTCAGTCTCTGGACGTGCTGGAGAAGTGTCAGATCGAT
GTGCTGGGACATGTCAGCGTGGTCCGATCCGAAAAGCGGATGGGCTTTAGCTAAGAATTC

#21

ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGAAAAGG
AACTACATTCTGGGGCTGGACATCGGGATTACAAGCGTGGGGTATGGGATTATTGACTATGAAACAAGGGACGTGATCGACGCA
GGCGTCAGACTGTTCAAGGAGGCCAACGTGGAAAACAATGAGGGACGGAGAAGCAAGAGGGGAGCCAGGCGCCTGAAACGACGG
AGAAGGCACAGAATCCAGAGGGTGAAGAAACTGCTGTTCGATTACAACCTGCTGACCGACCATTCTGAGCTGAGTGGAATTAAT
CCTTATGAAGCCAGGGTGAAAGGCCTGAGTCAGAAGCTGTCAGAGGAAGAGTTTTCCGCAGCTCTGCTGCACCTGGCTAAGCGC
CGAGGAGTGCATAACGTCAATGAGGTGGAAGAGGACACCGGCAACGAGCTGTCTACAAAGGAACAGATCTCACGCAATAGCAAA
GCTCTGGAAGAGAAGTATGTCGCAGAGCTGCAGCTGGAACGGCTGAAGAAAGATGGCGAGGTGAGAGGGTCAATTAATAGGTTC
AAGACAAGCGACTACGTCAAAGAAGCCAAGCAGCTGCTGAAAGTGCAGAAGGCTTACCACCAGCTGGATCAGAGCTTCATCGAT
ACTTATATCGACCTGCTGGAGACTCGGAGAACCTACTATGAGGGACCAGGAGAAGGGAGCCCCTTCGGATGGAAAGACATCAAG
GAATGGTACGAGATGCTGATGGGACATTGCACCTATTTTCCAGAAGAGCTGAGAAGCGTCAAGTACGCTTATAACGCAGATCTG
TACAACGCCCTGAATGACCTGAACAACCTGGTCATCACCAGGGATGAAAACGAGAAACTGGAATACTATGAGAAGTTCCAGATC
ATCGAAAACGTGTTTAAGCAGAAGAAAAAGCCTACACTGAAACAGATTGCTAAGGAGATCCTGGTCAACGAAGAGGACATCAAG
GGCTACCGGGTGACAAGCACTGGAAAACCAGAGTTCACCAATCTGAAAGTGTATCACGATATTAAGGACATCACAGCACGGAAA
GAAATCATTGAGAACGCCGAACTGCTGGATCAGATTGCTAAGATCCTGACTATCTACCAGAGCTCCGAGGACATCCAGGAAGAG

FIG. 9M

CTGACTAACCTGAACAGCGAGCTGACCCAGGAAGAGATCGAACAGATTAGTAATCTGAAGGGGTACACCGGAACACACAACCTG
TCCCTGAAAGCTATCAATCTGATTCTGGATGAGCTGTGGCATACAAACGACAATCAGATTGCAATCTTTAACCGGCTGAAGCTG
GTCCCAAAAAAGGTGGACCTGAGTCAGCAGAAAGAGATCCCAACCACACTGGTGGACGATTTCATTCTGTCACCCGTGGTCAAG
CGGAGCTTCATCCAGAGCATCAAAGTGATCAACGCCATCATCAAGAAGTACGGCCTGCCCAATGATATCATTATCGAGCTGGCT
AGGGAGAAGAACAGCAAGGACGCACAGAAGATGATCAATGAGATGCAGAAACGAAACCGGCAGACCAATGAACGCATTGAAGAG
ATTATCCGAACTACCGGGAAAGAGAACGCAAAGTACCTGATTGAAAAAATCAAGCTGCACGATATGCAGGAGGGAAAGTGTCTG
TATTCTCTGGAGGCCATCCCCCTGGAGGACCTGCTGAACAATCCATTCAACTACGAGGTCGATCATATTATCCCCAGAAGCGTG
TCCTTCGACAATTCCTTTAACAACAAGGTGCTGGTCAAGCAGGAAGAGAACTCTAAAAAGGGCAATAGGACTCCTTTCCAGTAC
CTGTCTAGTTCAGATTCCAAGATCTCTTACGAAACCTTTAAAAAGCACATTCTGAATCTGGCCAAAGGAAAGGGCCGCATCAGC
AAGACCAAAAAGGAGTACCTGCTGGAAGAGCGGGACATCAACAGATTCTCCGTCCAGAAGGATTTTATTAACCGGAATCTGGTG
GACACAAGATACGCTACTCGCGGCCTGATGAATCTGCTGCGATCCTATTTCCGGGTGAACAATCTGGATGTGAAAGTCAAGTCC
ATCAACGGCGGGTTCACATCTTTTCTGAGGCGCAAATGGAAGTTTAAAAAGGAGCGCAACAAAGGGTACAAGCACCATGCCGAA
GATGCTCTGATTATCGCAAATGCCGACTTCATCTTTAAGGAGTGGAAAAAGCTGGACAAAGCCAAGAAAGTGATGGAGAACCAG
ATGTTCGAAGAGAAGCAGGCCGAATCTATGCCCGAAATCGAGACAGAACAGGAGTACAAGGAGATTTTCATCACTCCTCACCAG
ATCAAGCATATCAAGGATTTCAAGGACTACAAGTACTCTCACCGGGTGGATAAAAAGCCCAACAGAGAGCTGATCAATGACACC
CTGTATAGTACAAGAAAAGACGATAAGGGGAATACCCTGATTGTGAACAATCTGAACGGACTGTACGACAAAGATAATGACAAG
CTGAAAAAGCTGATCAACAAAAGTCCCGAGAAGCTGCTGATGTACCACCATGATCCTCAGACATATCAGAAACTGAAGCTGATT
ATGGAGCAGTACGGCGACGAGAAGAACCCACTGTATAAGTACTATGAAGAGACTGGGAACTACCTGACCAAGTATAGCAAAAAG
GATAATGGCCCCGTGATCAAGAAGATCAAGTACTATGGGAACAAGCTGAATGCCCATCTGGACATCACAGACGATTACCCTAAC
AGTCGCAACAAGGTGGTCAAGCTGTCACTGAAGCCATACAGATTCGATGTCTATCTGGACAACGGCGTGTATAAATTTGTGACT
GTCAAGAATCTGGATGTCATCAAAAAGGAGAACTACTATGAAGTGAATAGCAAGTGCTACGAAGAGGCTAAAAAGCTGAAAAAG
ATTAGCAACCAGGCAGAGTTCATCGCCTCCTTTTACAACAACGACCTGATTAAGATCAATGGCGAACTGTATAGGGTCATCGGG
GTGAACAATGATCTGCTGAACCGCATTGAAGTGAATATGATTGACATCACTTACCGAGAGTATCTGGAAAACATGAATGATAAG
CGCCCCCCTCGAATTATCAAAACAATTGCCTCTAAGACTCAGAGTATCAAAAAGTACTCAACCGACATTCTGGGAAACCTGTAT
GAGGTGAAGAGCAAAAAGCACCCTCAGATTATCAAAAAGGGCTAAGAATTC

#22

ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGAAAAGG
AACTACATTCTGGGGCTGGACATCGGGATTACAAGCGTGGGGTATGGGATTATTGACTATGAAACAAGGGACGTGATCGACGCA
GGCGTCAGACTGTTCAAGGAGGCCAACGTGGAAAACAATGAGGGACGGAGAAGCAAGAGGGGAGCCAGGCGCCTGAAACGACGG
AGAAGGCACAGAATCCAGAGGGTGAAGAAACTGCTGTTCGATTACAACCTGCTGACCGACCATTCTGAGCTGAGTGGAATTAAT
CCTTATGAAGCCAGGGTGAAAGGCCTGAGTCAGAAGCTGTCAGAGGAAGAGTTTTCCGCAGCTCTGCTGCACCTGGCTAAGCGC
CGAGGAGTGCATAACGTCAATGAGGTGGAAGAGGACACCGGCAACGAGCTGTCTACAAAGGAACAGATCTCACGCAATAGCAAA
GCTCTGGAAGAGAAGTATGTCGCAGAGCTGCAGCTGGAACGGCTGAAGAAAGATGGCGAGGTGAGAGGGTCAATTAATAGGTTC
AAGACAAGCGACTACGTCAAAGAAGCCAAGCAGCTGCTGAAAGTGCAGAAGGCTTACCACCAGCTGGATCAGAGCTTCATCGAT
ACTTATATCGACCTGCTGGAGACTCGGAGAACCTACTATGAGGGACCAGGAGAAGGGAGCCCCTTCGGATGGAAAGACATCAAG
GAATGGTACGAGATGCTGATGGGACATTGCACCTATTTTCCAGAAGAGCTGAGAAGCGTCAAGTACGCTTATAACGCAGATCTG
TACAACGCCCTGAATGACCTGAACAACCTGGTCATCACCAGGGATGAAAACGAGAAACTGGAATACTATGAGAAGTTCCAGATC
ATCGAAAACGTGTTTAAGCAGAAGAAAAAGCCTACACTGAAACAGATTGCTAAGGAGATCCTGGTCAACGAAGAGGACATCAAG
GGCTACCGGGTGACAAGCACTGGAAAACCAGAGTTCACCAATCTGAAAGTGTATCACGATATTAAGGACATCACAGCACGGAAA
GAAATCATTGAGAACGCCGAACTGCTGGATCAGATTGCTAAGATCCTGACTATCTACCAGAGCTCCGAGGACATCCAGGAAGAG
CTGACTAACCTGAACAGCGAGCTGACCCAGGAAGAGATCGAACAGATTAGTAATCTGAAGGGGTACACCGGAACACACAACCTG
TCCCTGAAAGCTATCAATCTGATTCTGGATGAGCTGTGGCATACAAACGACAATCAGATTGCAATCTTTAACCGGCTGAAGCTG
GTCCCAAAAAAGGTGGACCTGAGTCAGCAGAAAGAGATCCCAACCACACTGGTGGACGATTTCATTCTGTCACCCGTGGTCAAG
CGGAGCTTCATCCAGAGCATCAAAGTGATCAACGCCATCATCAAGAAGTACGGCCTGCCCAATGATATCATTATCGAGCTGGCT
AGGGAGAAGAACAGCAAGGACGCACAGAAGATGATCAATGAGATGCAGAAACGAAACCGGCAGACCAATGAACGCATTGAAGAG
ATTATCCGAACTACCGGGAAAGAGAACGCAAAGTACCTGATTGAAAAAATCAAGCTGCACGATATGCAGGAGGGAAAGTGTCTG
TATTCTCTGGAGGCCATCCCCCTGGAGGACCTGCTGAACAATCCATTCAACTACGAGGTCGATCATATTATCCCCAGAAGCGTG
TCCTTCGACAATTCCTTTAACAACAAGGTGCTGGTCAAGCAGGAAGAGAACTCTAAAAAGGGCAATAGGACTCCTTTCCAGTAC
CTGTCTAGTTCAGATTCCAAGATCTCTTACGAAACCTTTAAAAAGCACATTCTGAATCTGGCCAAAGGAAAGGGCCGCATCAGC
AAGACCAAAAAGGAGTACCTGCTGGAAGAGCGGGACATCAACAGATTCTCCGTCCAGAAGGATTTTATTAACCGGAATCTGGTG
GACACAAGATACGCTACTCGCGGCCTGATGAATCTGCTGCGATCCTATTTCCGGGTGAACAATCTGGATGTGAAAGTCAAGTCC
ATCAACGGCGGGTTCACATCTTTTCTGAGGCGCAAATGGAAGTTTAAAAAGGAGCGCAACAAAGGGTACAAGCACCATGCCGAA
GATGCTCTGATTATCGCAAATGCCGACTTCATCTTTAAGGAGTGGAAAAAGCTGGACAAAGCCAAGAAAGTGATGGAGAACCAG
ATGTTCGAAGAGAAGCAGGCCGAATCTATGCCCGAAATCGAGACAGAACAGGAGTACAAGGAGATTTTCATCACTCCTCACCAG
ATCAAGCATATCAAGGATTTCAAGGACTACAAGTACTCTCACCGGGTGGATAAAAAGCCCAACAGAGAGCTGATCAATGACACC
CTGTATAGTACAAGAAAAGACGATAAGGGGAATACCCTGATTGTGAACAATCTGAACGGACTGTACGACAAAGATAATGACAAG
CTGAAAAAGCTGATCAACAAAAGTCCCGAGAAGCTGCTGATGTACCACCATGATCCTCAGACATATCAGAAACTGAAGCTGATT
ATGGAGCAGTACGGCGACGAGAAGAACCCACTGTATAAGTACTATGAAGAGACTGGGAACTACCTGACCAAGTATAGCAAAAAG
GATAATGGCCCCGTGATCAAGAAGATCAAGTACTATGGGAACAAGCTGAATGCCCATCTGGACATCACAGACGATTACCCTAAC
AGTCGCAACAAGGTGGTCAAGCTGTCACTGAAGCCATACAGATTCGATGTCTATCTGGACAACGGCGTGTATAAATTTGTGACT
GTCAAGAATCTGGATGTCATCAAAAAGGAGAACTACTATGAAGTGAATAGCAAGTGCTACGAAGAGGCTAAAAAGCTGAAAAAG
ATTAGCAACCAGGCAGAGTTCATCGCCTCCTTTTACAACAACGACCTGATTAAGATCAATGGCGAACTGTATAGGGTCATCGGG

FIG. 9N

GTGAACAATGATCTGCTGAACCGCATTGAAGTGAATATGATTGACATCACTTACCGAGAGTATCTGGAAAACATGAATGATAAG
CGCCCCCCTCGAATTATCAAAACAATTGCCTCTAAGACTCAGAGTATCAAAAAGTACTCAACCGACATTCTGGGAAACCTGTAT
GAGGTGAAGAGCAAAAAGCACCCTCAGATTATCAAAAAGGGCTAAGAATTC

#23
ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGAACAAT
AGCATCAAATCTAAACCTGAAGTGACCATCGGGCTGGACCTGGGAGTGGGAAGCGTGGGGTGGGCAATCGTGGATAACGAAACA
AACATCATTCACCATCTGGGCTCCAGGCTGTTTTCTCAGGCCAAGACTGCTGAGGATCGGAGATCTTTCCGCGGGGTGAGGCGC
CTGATCCGACGGAGAAAATACAAGCTGAAACGATTCGTCAATCTGATTTGGAAGTACAACAGCTATTTCGGCTTCAAGAACAAA
GAGGACATCCTGAACAATTATCAGGAGCAGCAGAAGCTGCACAATACCGTGCTGAACCTGAAATCAGAGGCACTGAATGCCAAG
ATCGATCCTAAAGCACTGAGCTGGATTCTGCACGACTACCTGAAGAACAGAGGCCATTTTTATGAGGACAATAGGGATTTCAAC
GTGTACCCAACAAAGGAGCTGGCCAAGTACTTCGATAAGTACGGGTACTACAAGGGAATCATTGACAGCAAGGAGGACAATGAT
AACAAACTGGAGGAAGAGCTGACAAAGTACAAATTCTCCAATAAGCACTGGCTGGAAGAGGTGAAGAAAGTCCTGTCTAACCAG
ACTGGCCTGCCAGAAAAGTTTAAAGAAGAGTATGAGTCACTGTTCAGCTACGTGAGAAATTATTCAGAGGGCCCAGGGAGCATC
AACTCTGTCAGTCCCTACGGGATCTACCATCTGGACGAAAAAGAGGGAAAGGTGGTCCAGAAGTACAACAACATCTGGGATAAG
ACAATCGGAAAGTGCAACATCTTCCCTGACGAGTATAGAGCTCCCAAGAACAGTCCTATCGCAATGATTTTCAATGAAATCAAC
GAGCTGTCCACAATCAGGTCATACAGCATCTACCTGACTGGCTGGTTCATTAATCAGGAGTTCAAGAAAGCCTACCTGAACAAG
CTGCTGGATCTGCTGATCAAAACCAACGGAGAGAAGCCAATTGACGCAAGGCAGTTCAAGAAACTGCGCGAAGAGACAATCGCC
GAAAGCATTGGCAAAGAGACACTGAAGGATGTGGAGAATGAAGAGAAACTGGAAAAGGAGGACCACAAGTGGAAACTGAAGGGA
CTGAAGCTGAATACCAACGGCAAAATCCAGTACAACGATCTGAGCTCCCTGGCTAAGTTTGTGCACAAACTGAAGCAGCATCTG
AAACTGGATTTCCTGCTGGAGGACCAGTATGCAACACTGGACAAGATCAATTTCCTGCAGTCCCTGTTTGTGTACCTGGGCAAG
CACCTGAGATATTCCAATAGGGTCGATTCTGCCAACCTGAAGGAATTTTCCGACTCTAACAAACTGTTCGAGCGCATCCTGCAG
AAACAGAAGGATGGGCTGTTCAAGCTGTTTGAACAGACTGACAAAGACGATGAGAAGATCCTGGCCCAGACACATAGTCTGTCA
ACTAAGGCCATGCTGCTGGCTATTACCCGGATGACAAATCTGGACAACGATGAGGACAACCAGAAAAACAATGACAAGGGCTGG
AATTTTGAGGCCATCAAAAACTTCGATCAGAAGTTTATCGACATCACCAAGAAAAACAACAACCTGAGCCTGAAACAGAATAAG
CGCTACCTGGACGATCGATTCATCAACGATGCTATTCTGTCCCCTGGGGTGAAGCGAATCCTGCGGGAGGCAACCAAGGTCTTT
AATGCCATTCTGAAACAGTTCTCTGAAGAGTACGACGTGACAAAGGTGGTCATCGAACTGGCTCGCGAGCTGAGCGAAGAGAAG
GAACTGGAGAACACAAAGAACTACAAGAAACTGATCAAGAAAAACGGCGACAAGATTAGTGAGGGCCTGAAAGCACTGGGGATC
TCAGAAGATGAGATCAAAGACATTCTGAAGAGTCCCACTAAATCATACAAGTTTCTGCTGTGGCTGCAGCAGGACCACATCGAT
CCTTATAGCCTGAAGGAGATCGCCTTCGACGATATTTTTACCAAAACAGAAAAGTTCGAGATCGACCATATCATTCCCTACAGC
ATTTCCTTCGACGATTCTAGTTCAAACAAGCTGCTGGTGCTGGCTGAAAGTAATCAGGCAAAGTCAAACCAGACTCCTTATGAG
TTCATCAGCTCCGGAAACGCAGGCATTAAGTGGGAAGATTACGAGGCCTATTGCCGCAAGTTCAAGGATGGGGACTCTAGTCTG
CTGGACAGCACCCAGCGGTCCAAGAAATTCGCCAAAATGATGAAAACCGATACCTCAAGCAAGTACGACATCGGATTTCTGGCT
CGAAATCTGAACGATACTCGGTACGCAACCATTGTGTTCCGGGACGCCCTGGAGGACTATGCTAATAACCACCTGGTCGAGGAC
AAACCCATGTTTAAGGTGGTCTGTATCAATGGGTCCGTGACCTCTTTCCTGCGGAAGAACTTTGACGATTCCTCTTACGCCAAG
AAAGATAGAGACAAGAATATCCACCATGCTGTGGATGCAAGTATCATCTCAATTTTCAGCAACGAGACAAAGACTCTGTTCAAC
CAGCTGACTCAGTTTGCTGACTATAAACTGTTCAAGAACACCGATGGCAGCTGGAAGAAAATCGACCCTAAGACAGGGGTGGTC
ACTGAAGTGACCGACGAGAATTGGAAGCAGATTAGGGTGCGCAACCAGGTGAGCGAAATCGCCAAAGTCATTGAGAAGTACATC
CAGGATAGCAACATCGAAAGAAAGGCTAGGTATTCCCGCAAAATCGAGAATAAGACTAACATTTCCCTGTTTAATGACACCGTG
TACTCTGCCAAGAAAGTCGGCTATGAGGATCAGATCAAAAGAAAGAACCTGAAAACCCTGGACATTCACGAATCTGCTAAAGAG
AATAAGAACAGTAAAGTGAAGCGGCAGTTTGTCTACAGAAAGCTGGTGAATGTCAGCCTGCTGAATAACGATAAGCTGGCAGAC
CTGTTCGCCGAAAAAGAGGATATCCTGATGTATAGGGCCAATCCATGGGTCATCAACCTGGCTGAGCAGATTTTCAATGAATAC
ACTGAGAACAAGAAAATCAAGTCCCAGAACGTGTTTGAAAAATATATGCTGGACCTGACCAAAGAGTTCCCCGAGAAGTTCAGC
GAGTTTCTGGTGAAGTCCATGCTGAGAAACAAGACCGCCATCATCTACGACGATAAGAAAAACATTGTCCATCGAATCAAACGG
CTGAAGATGCTGAGTTCAGAACTGAAAGAGAATAAGCTGTCTAACGTGATCATTAGGTCTAAGAATCAGAGTGGGACCAAACTG
TCATACCAGGATACAATCAACAGCCTGGCCCTGATGATTATGCGCAGCATCGACCCTACTGCTAAGAAACAGTATATTCGAGTG
CCACTGAATACCCTGAACCTGCACCTGGGAGATCATGACTTTGATCTGCACAATATGGATGCTTACCTGAAGAAACCAAAATTC
GTGAAGTATCTGAAAGCAAACGAAATCGGCGACGAGTACAAGCCCTGGAGGGTCCTGACATCTGGCACTCTGCTGATCCATAAG
AAGGATAAGAAACTGATGTACATCAGCTCCTTCCAGAATCTGAACGACGTGATCGAAATTAAGAATCTGATCGAAACCGAGTAT
AAAGAGAACGACGATTCTGATAGTAAGAAAAAGAAAAAGGCAAACCGCTTTCTGATGACCCTGAGCACAATCCTGAATGACTAC
ATTCTGCTGGACGCCAAGGATAACTTCGACATCCTGGGGCTGTCTAAAAATCGGATCGATGAGATTCTGAACAGTAAGCTGGGA
CTGGACAAGATTGTGAAATAAGAATTC

#24
ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGAGGATT
CTGGGGTTTGACATTGGCATTAACAGCATCGGGTGGGCTTTTGTGGAGAACGACGAACTGAAGGACTGCGGAGTGCGGATCTTC
ACAAAGGCCGAGAACCCAAAAAATAAGGAAAGCCTGGCACTGCCCCGGAGAAATGCACGCAGCTCCAGGCGCCGACTGAAACGG
AGAAAGGCCCGGCTGATCGCTATTAAGAGAATCCTGGCCAAAGAGCTGAAGCTGAACTACAAGGACTATGTCGCAGCTGATGGA
GAGCTGCCAAAGGCCTACGAAGGATCCCTGGCATCTGTGTACGAGCTGCGGTATAAGGCCCTGACACAGAACCTGGAAACTAAA
GATCTGGCCAGAGTGATCCTGCACATTGCTAAGCATAGGGGGTACATGAACAAGAACGAGAAGAAATCAAACGACGCTAAGAAA
GGAAAGATCCTGAGCGCTCTGAAAAACAATGCACTGAAGCTGGAGAACTACCAGAGCGTGGGCGAATACTTCTACAAGGAGTTC
TTTCAGAAATACAAGAAAAACACAAAGAACTTCATCAAGATCCGCAACACTAAGGATAATTACAACAATTGCGTGCTGTCTAGT
GACCTGGAAAAAGAGCTGAAGCTGATCCTGGAAAAACAGAAGGAGTTCGGCTACAACTACTCTGAAGATTTCATCAACGAGATT

FIG. 9O

CTGAAGGTCGCCTTCTTTCAGCGGCCCCTGAAGGACTTCAGTCACCTGGTGGGGGCCTGCACTTTCTTTGAGGAAGAGAAAAGG
GCCTGTAAGAACAGCTACTCTGCCTGGGAGTTTGTGGCTCTGACCAAGATCATTAACGAGATCAAGAGCCTGGAGAAGATCAGC
GGCGAAATTGTGCCAACCCAGACAATCAACGAGGTCCTGAATCTGATCCTGGACAAGGGGTCTATCACCTACAAGAAATTCAGA
AGTTGTATCAATCTGCATGAGAGTATCAGCTTCAAGAGCCTGAAGTATGATAAAGAAAACGCCGAGAATGCTAAACTGATCGAC
TTCCGCAAGCTGGTGGAGTTTAAGAAAGCCCTGGGAGTCCACAGCCTGTCCCGGCAGGAACTGGATCAGATCTCCACTCATATC
ACCCTGATTAAGGACAACGTGAAGCTGAAAACCGTCCTGGAGAAATACAACCTGAGTAATGAACAGATCAACAATCTGCTGGAA
ATTGAGTTCAACGATTATATCAACCTGAGCTTCAAGGCCCTGGGAATGATTCTGCCACTGATGCGCGAGGGCAAACGATACGAC
GAGGCCTGCGAGATCGCCAATCTGAAACCTAAGACCGTGGACGAGAAGAAAGATTTCCTGCCAGCATTTTGTGATTCCATTTTC
GCCCACGAGCTGTCTAACCCCGTGGTCAATAGGGCTATCAGCGAATACCGCAAGGTGCTGAACGCACTGCTGAAGAAATATGGA
AAGGTCCACAAAATTCATCTGGAGCTGGCTCGCGACGTGGGCCTGTCCAAGAAAGCACGAGAGAAGATCGAAAAAGAGCAGAAG
GAAAACCAGGCCGTGAATGCATGGGCCCTGAAGGAATGCGAGAATATTGGCCTGAAGGCCAGCGCAAAGAACATCCTGAAACTG
AAGCTGTGGAAAGAACAGAAGGAGATCTGTATCTACTCCGGAAATAAGATCTCTATTGAGCACCTGAAAGATGAAAAGGCCCTG
GAGGTGGACCATATCTACCCCTATTCTAGGAGTTTCGACGATTCTTTTATCAACAAAGTGCTGGTGTTCACCAAGGAAAATCAG
GAGAAACTGAACAAGACACCTTTCGAGGCCTTTGGCAAGAATATTGAAAAATGGAGCAAGATCCAGACCCTGGCTCAGAACCTG
CCATACAAGAAAAAGAATAAGATTCTGGACGAGAACTTCAAAGATAAGCAGCAGGAGGACTTTATCTCTCGAAATCTGAACGAC
ACCCGGTATATCGCTACACTGATTGCAAAATACACAAAGGAGTATCTGAACTTCCTGCTGCTGAGCGAAAATGAGAACGCCAAT
CTGAAGAGTGGCGAAAAAGGGTCAAAGATCCACGTGCAGACTATTAGCGGGATGCTGACCTCCGTCCTGAGGCACACATGGGGG
TTTGACAAAAAGGATCGCAACAATCATCTGCACCATGCACTGGATGCCATCATTGTGGCCTACAGTACAAATTCAATCATTAAG
GCTTTCAGCGATTTCCGGAAAAACCAGGAGCTGCTGAAGGCCAGATTCTACGCTAAAGAACTGACTTCCGATAACTATAAACAT
CAGGTCAAGTTCTTTGAGCCTTTCAAGAGTTTTAGAGAAAAAATCCTGTCAAAGATCGACGAGATTTTCGTGTCCAAACCACCT
CGAAAGCGAGCTAGGCGCGCACTGCACAAGGATACCTTTCATTCTGAGAACAAGATCATTGACAAGTGCAGCTACAACTCCAAG
GAAGGCCTGCAGATTGCCCTGAGCTGTGGAAGAGTGAGGAAAATCGGCACTAAGTATGTCGAGAATGATACCATCGTGAGGGTC
GACATTTTCAAAAAGCAGAACAAGTTTTACGCTATCCCAATCTACGCAATGGATTTTGCCCTGGGGATCCTGCCCAATAAGATC
GTGATTACTGGAAAAGATAAGAACAATAACCCCAAACAGTGGCAGACCATTGACGAATCATACGAGTTCTGCTTTAGCCTGTAT
AAGAATGACCTGATCCTGCTGCAGAAAAAGAACATGCAGGAACCTGAGTTCGCCTACTATAACGATTTTTCAATCAGCACATCA
AGCATTTGTGTGGAGAAACACGACAACAAGTTCGAAAATCTGACTAGCAACCAGAAGCTGCTGTTTTCCAATGCAAAAGAGGGC
TCTGTGAAGGTCGAAAGTCTGGGGATCCAGAACCTGAAAGTGTTCGAGAAGTACATCATTACCCCCCTGGGAGATAAAATTAAG
GCTGACTTTCAGCCTCGAGAAAACATCAGCCTGAAAACCAGTAAAAAGTATGGCCTGAGGTAAGAATTC

SpCas9 mutation positions

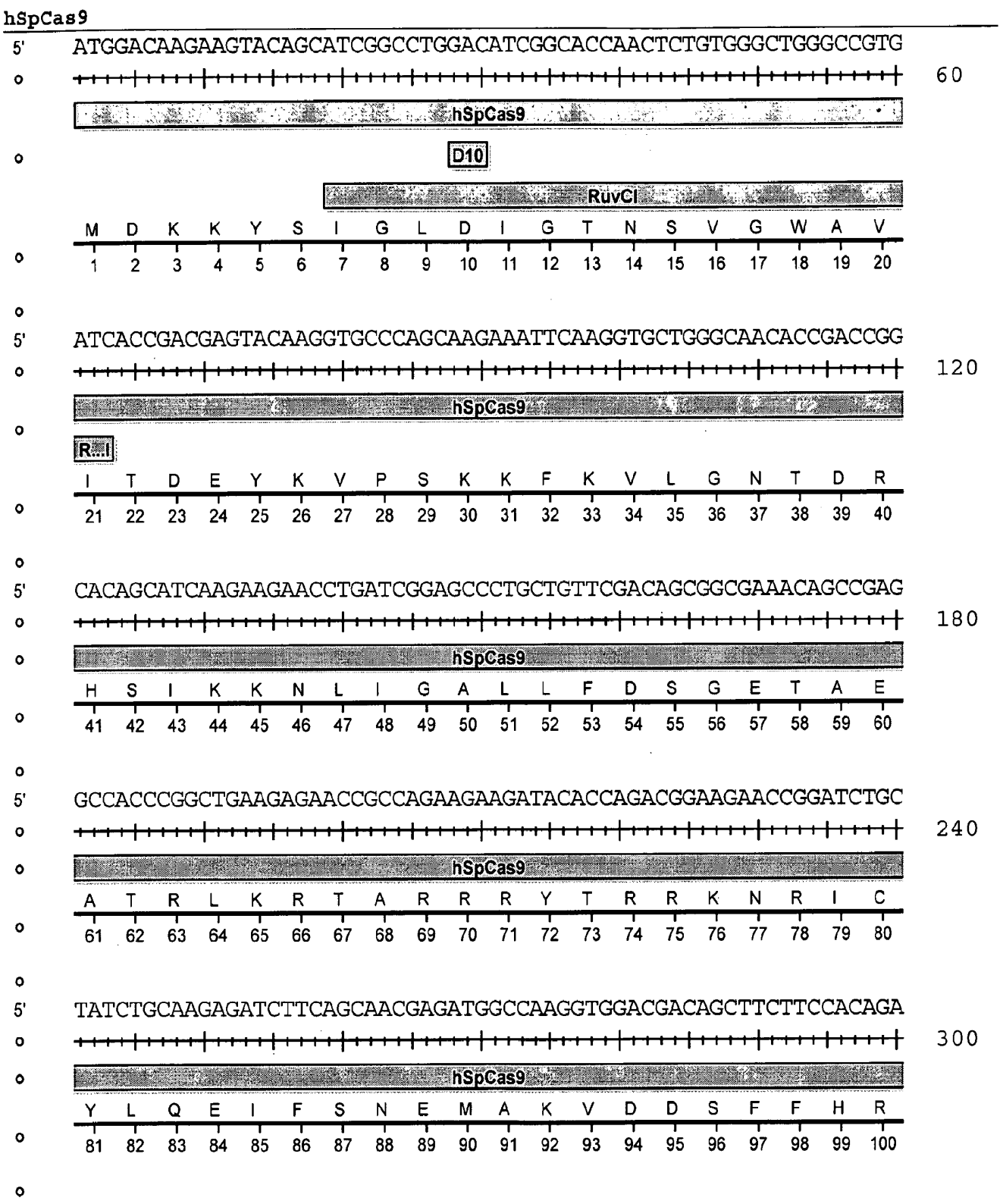


FIG. 10A

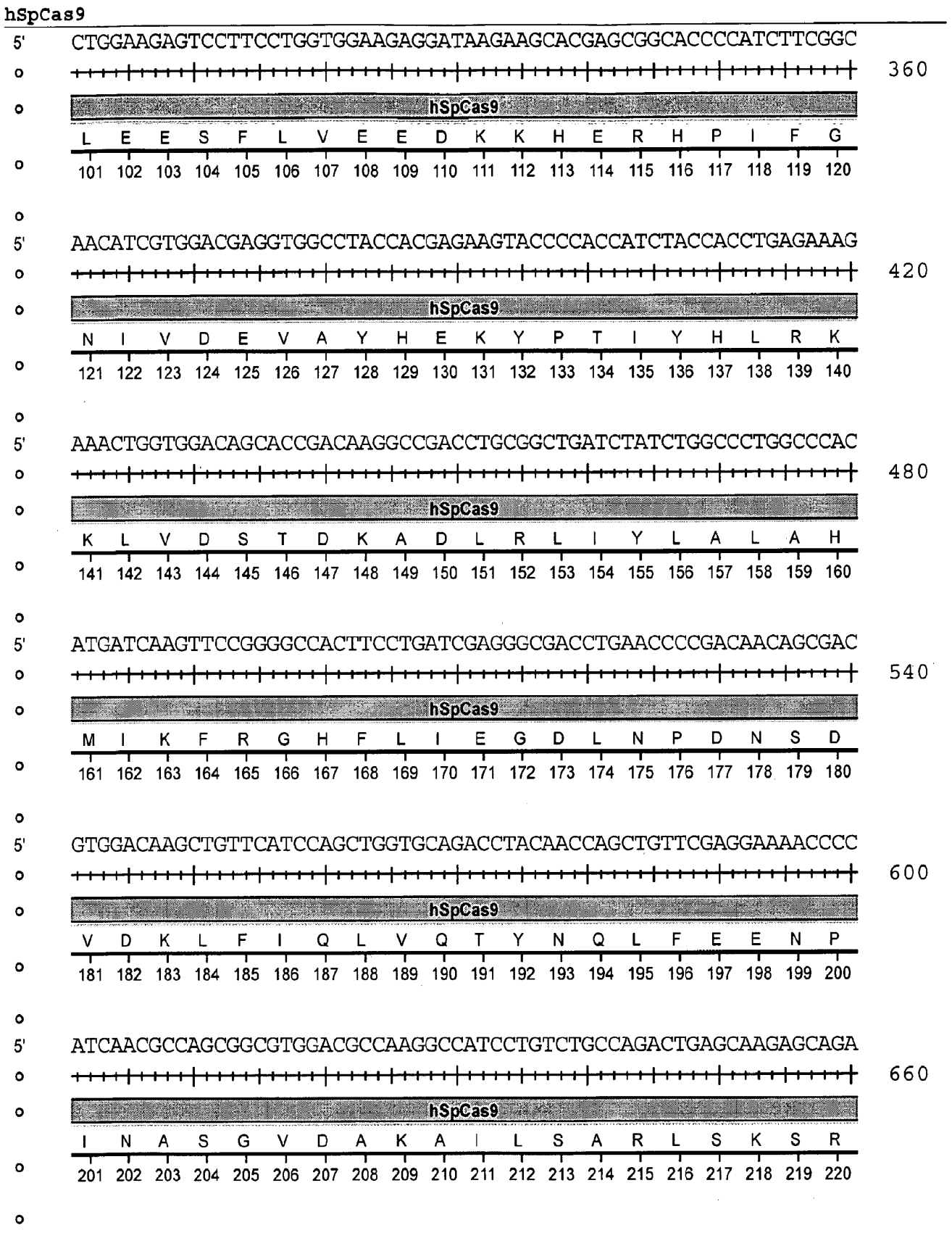

hSpCas9
5' CTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGC
360
hSpCas9
L E E S F L V E E D K K H E R H P I F G
101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
5' AACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAG
420
hSpCas9
N I V D E V A Y H E K Y P T I Y H L R K
121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
5' AAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCAC
480
hSpCas9
K L V D S T D K A D L R L I Y L A L A H
141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
5' ATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGAC
540
hSpCas9
M I K F R G H F L I E G D L N P D N S D
161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
5' GTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCC
600
hSpCas9
V D K L F I Q L V Q T Y N Q L F E E N P
181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
5' ATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGA
660
hSpCas9
I N A S G V D A K A I L S A R L S K S R
201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220

FIG. 10B

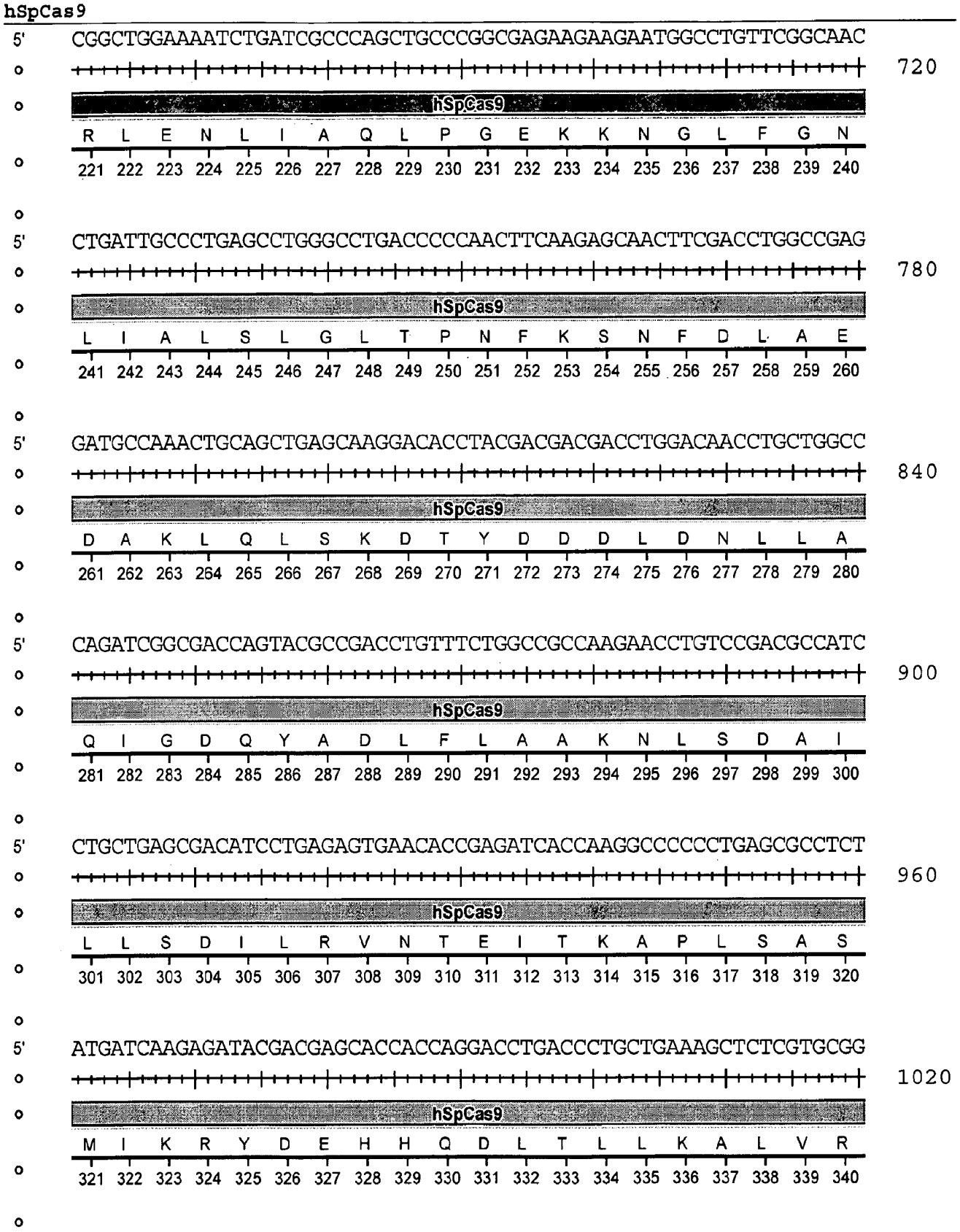

hSpCas9
5' CGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAAC
720
hSpCas9
R L E N L I A Q L P G E K K N G L F G N
221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
5' CTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAG
780
hSpCas9
L I A L S L G L T P N F K S N F D L A E
241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
5' GATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCC
840
hSpCas9
D A K L Q L S K D T Y D D D L D N L L A
261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
5' CAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATC
900
hSpCas9
Q I G D Q Y A D L F L A A K N L S D A I
281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
5' CTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCT
960
hSpCas9
L L S D I L R V N T E I T K A P L S A S
301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
5' ATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGG
1020
hSpCas9
M I K R Y D E H H Q D L T L L K A L V R
321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340

FIG. 10C

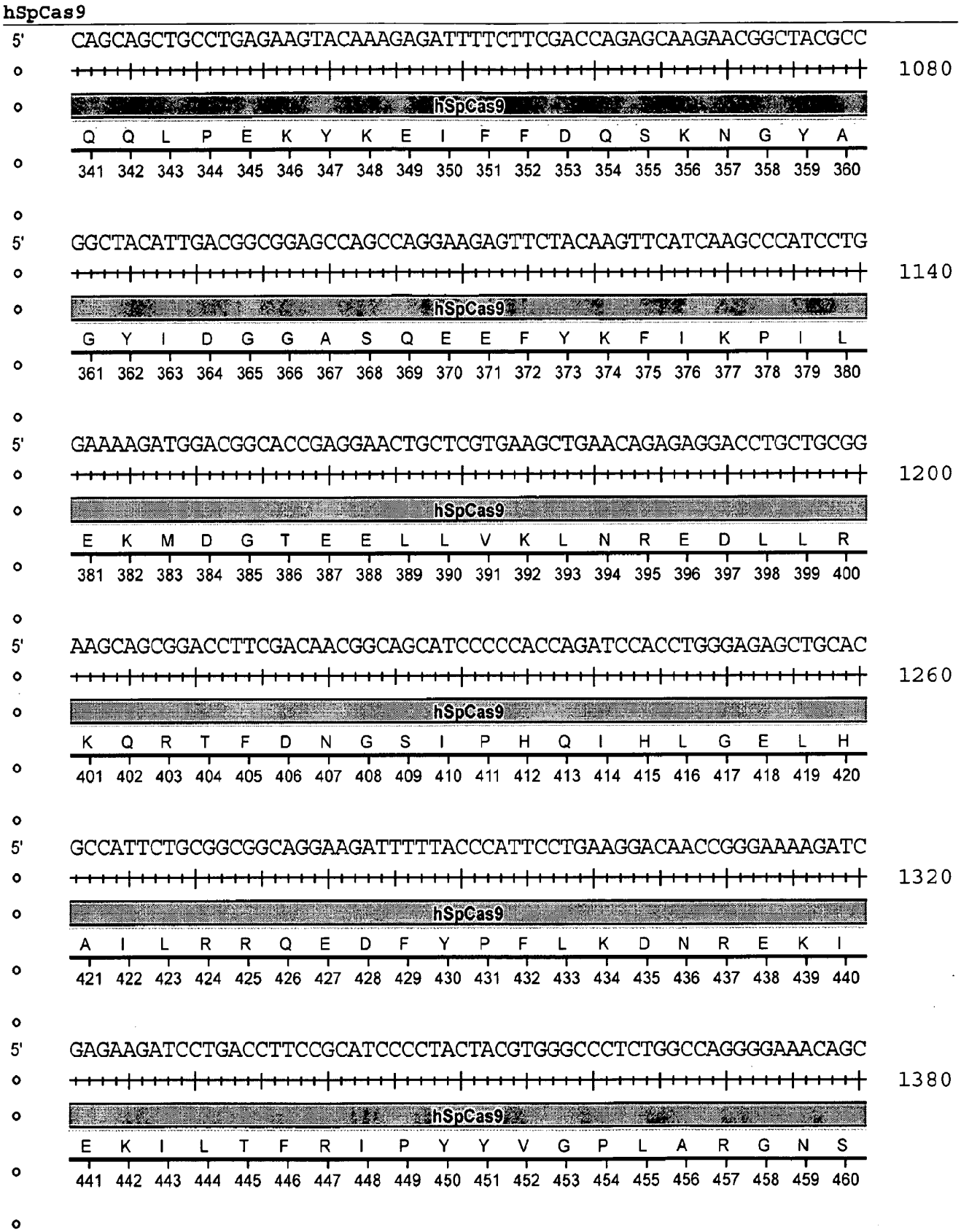

hSpCas9
5' CAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCC
1080
hSpCas9
Q Q L P E K Y K E I F F D Q S K N G Y A
341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
5' GGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTG
1140
hSpCas9
G Y I D G G A S Q E E F Y K F I K P I L
361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
5' GAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGG
1200
hSpCas9
E K M D G T E E L L V K L N R E D L L R
381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
5' AAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCAC
1260
hSpCas9
K Q R T F D N G S I P H Q I H L G E L H
401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
5' GCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATC
1320
hSpCas9
A I L R R Q E D F Y P F L K D N R E K I
421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
5' GAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGC
1380
hSpCas9
E K I L T F R I P Y Y V G P L A R G N S
441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460

FIG. 10D

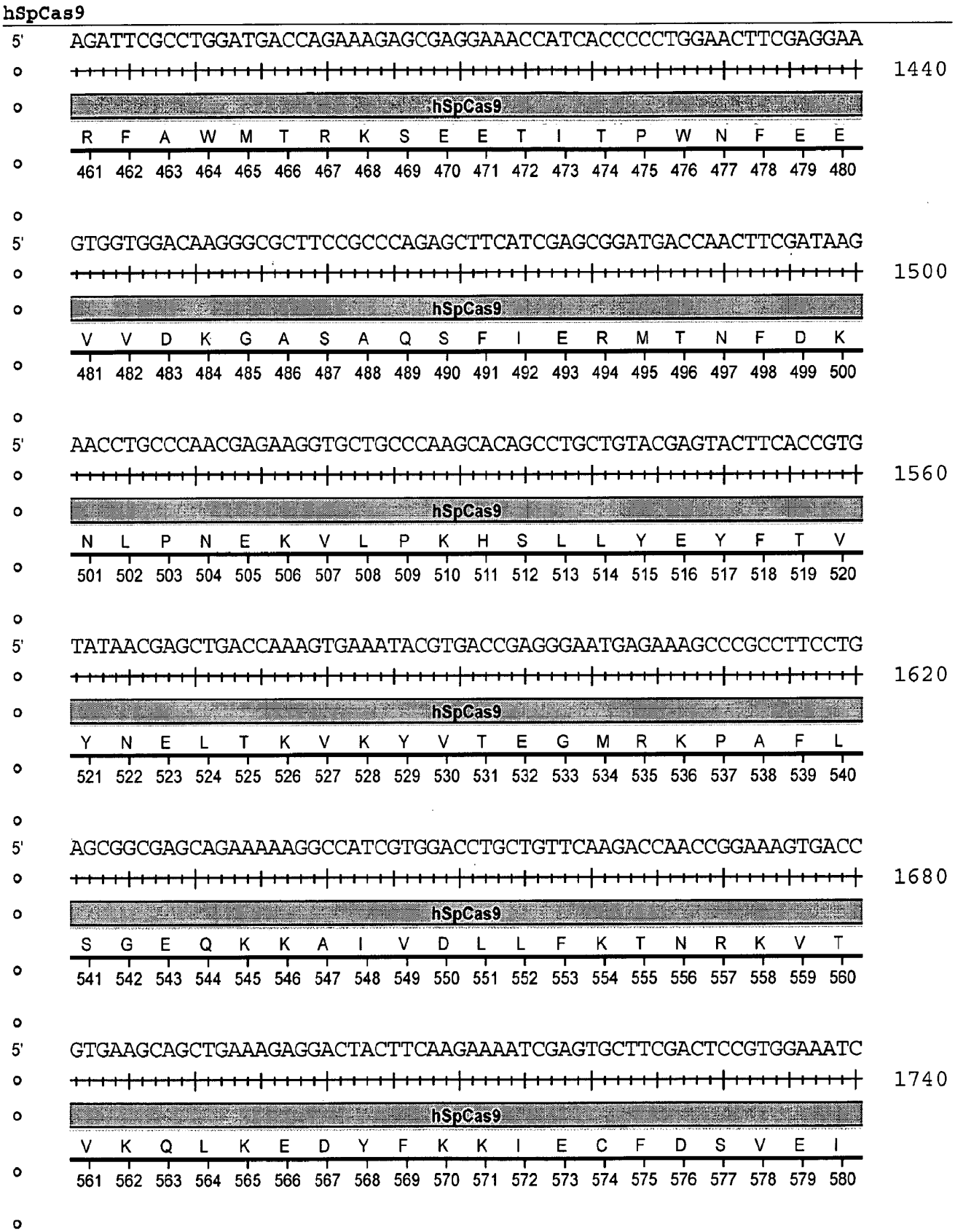

hSpCas9
5' AGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAA
1440
hSpCas9
R F A W M T R K S E E T I T P W N F E E
461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
5' GTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAG
1500
hSpCas9
V V D K G A S A Q S F I E R M T N F D K
481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
5' AACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTG
1560
hSpCas9
N L P N E K V L P K H S L L Y E Y F T V
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520
5' TATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTG
1620
hSpCas9
Y N E L T K V K Y V T E G M R K P A F L
521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540
5' AGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACC
1680
hSpCas9
S G E Q K K A I V D L L F K T N R K V T
541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560
5' GTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATC
1740
hSpCas9
V K Q L K E D Y F K K I E C F D S V E I
561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580

FIG. 10E

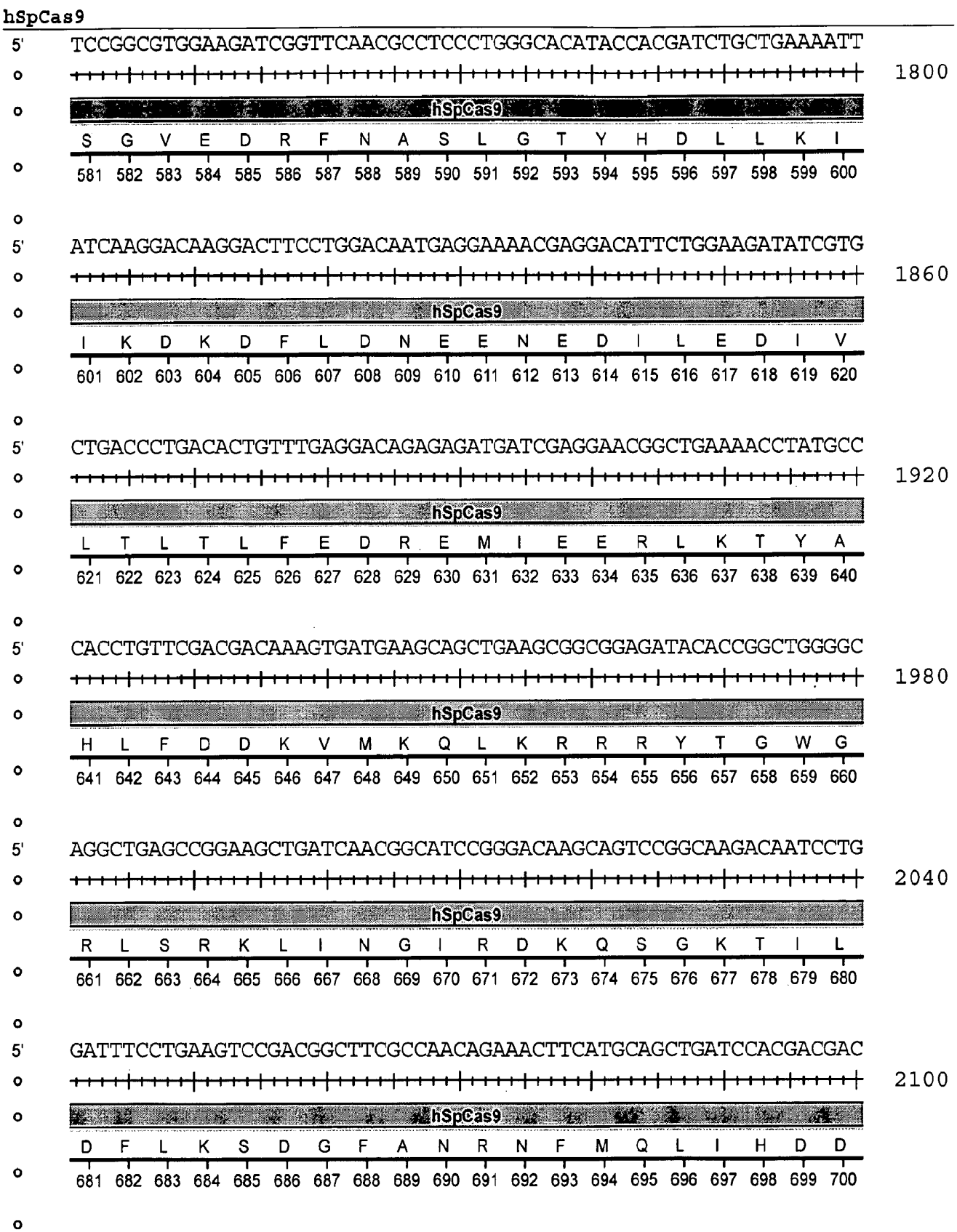

hSpCas9
5' TCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATT
1800
hSpCas9
S G V E D R F N A S L G T Y H D L L K I
581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600
5' ATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTG
1860
hSpCas9
I K D K D F L D N E E N E D I L E D I V
601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620
5' CTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCC
1920
hSpCas9
L T L T L F E D R E M I E E R L K T Y A
621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640
5' CACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGC
1980
hSpCas9
H L F D D K V M K Q L K R R R Y T G W G
641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660
5' AGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTG
2040
hSpCas9
R L S R K L I N G I R D K Q S G K T I L
661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680
5' GATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGAC
2100
hSpCas9
D F L K S D G F A N R N F M Q L I H D D
681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700

FIG. 10F

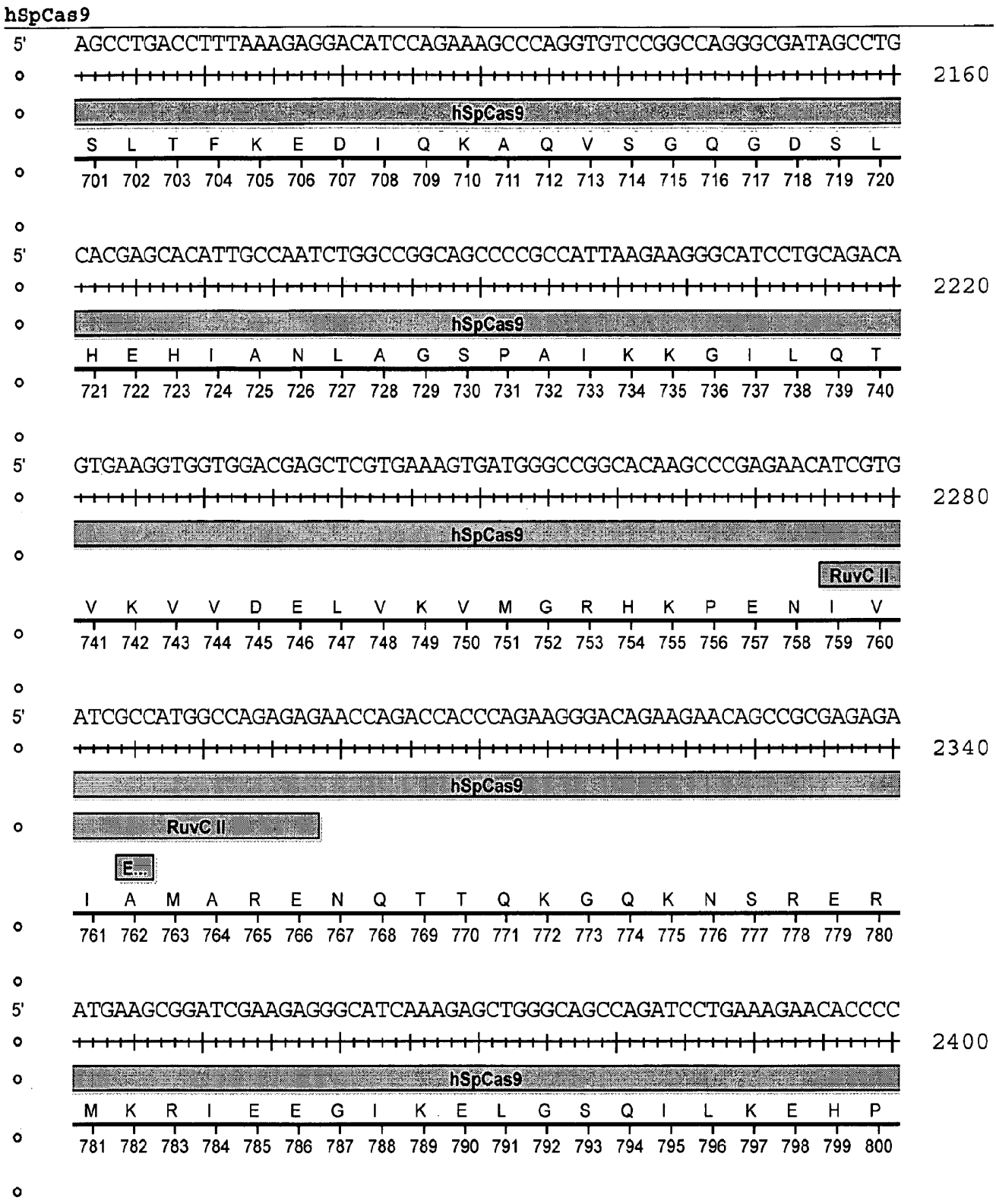

hSpCas9
5' AGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTG
2160
hSpCas9
S L T F K E D I Q K A Q V S G Q G D S L
701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720
5' CACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACA
2220
hSpCas9
H E H I A N L A G S P A I K K G I L Q T
721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740
5' GTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTG
2280
hSpCas9
RuvC II
V K V V D E L V K V M G R H K P E N I V
741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760
5' ATCGCCATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGA
2340
hSpCas9
RuvC II
E...
I A M A R E N Q T T Q K G Q K N S R E R
761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780
5' ATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCC
2400
hSpCas9
M K R I E E G I K E L G S Q I L K E H P
781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800

FIG. 10G

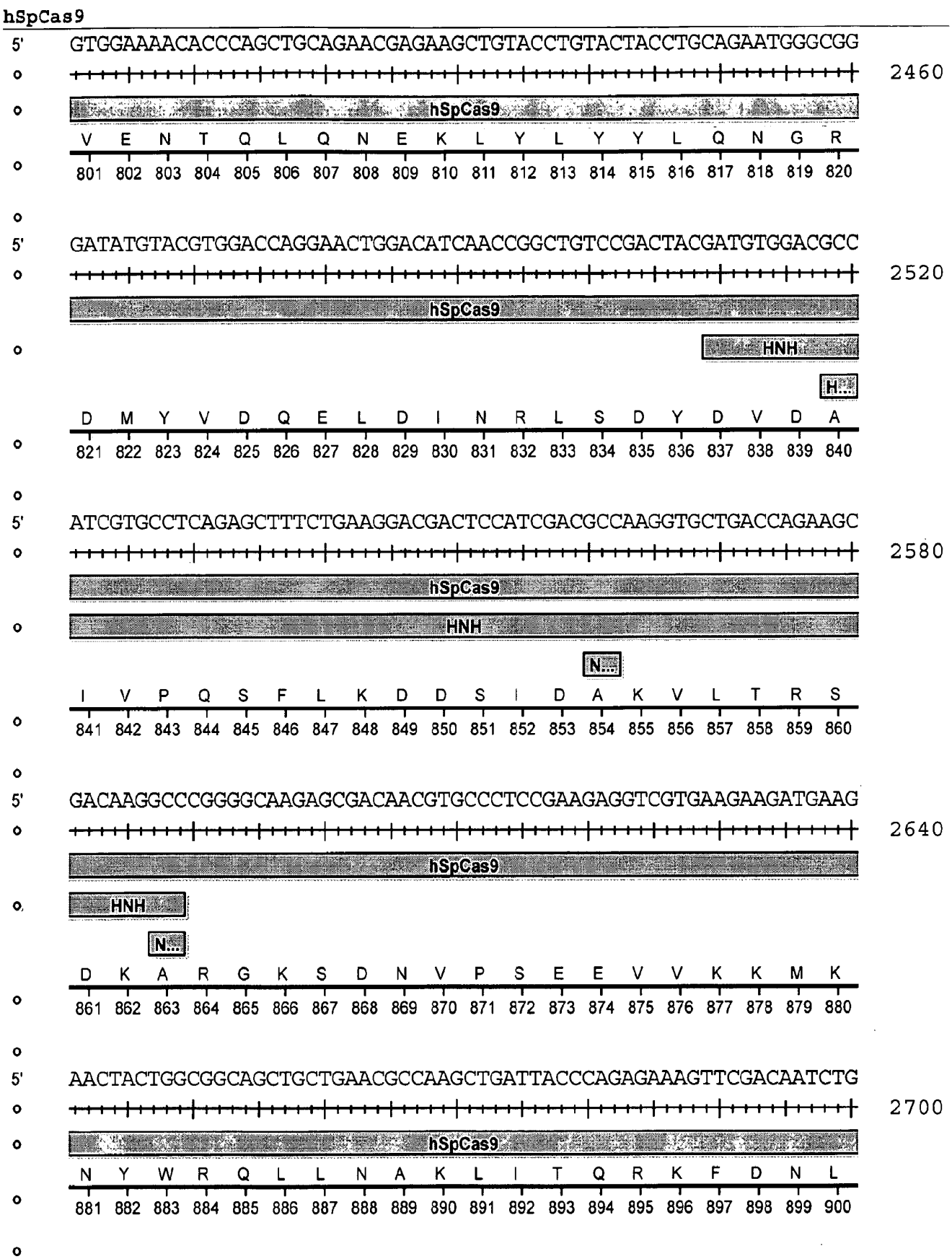

hSpCas9
5' GTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGG
2460
hSpCas9
V E N T Q L Q N E K L Y L Y Y L Q N G R
801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820
5' GATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACGCC
2520
hSpCas9
HNH
H...
D M Y V D Q E L D I N R L S D Y D V D A
821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840
5' ATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACGCCAAGGTGCTGACCAGAAGC
2580
hSpCas9
HNH
N...
I V P Q S F L K D D S I D A K V L T R S
841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860
5' GACAAGGCCCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAG
2640
hSpCas9
HNH
N...
D K A R G K S D N V P S E E V V K K M K
861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880
5' AACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTG
2700
hSpCas9
N Y W R Q L L N A K L I T Q R K F D N L
881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900

FIG. 10H

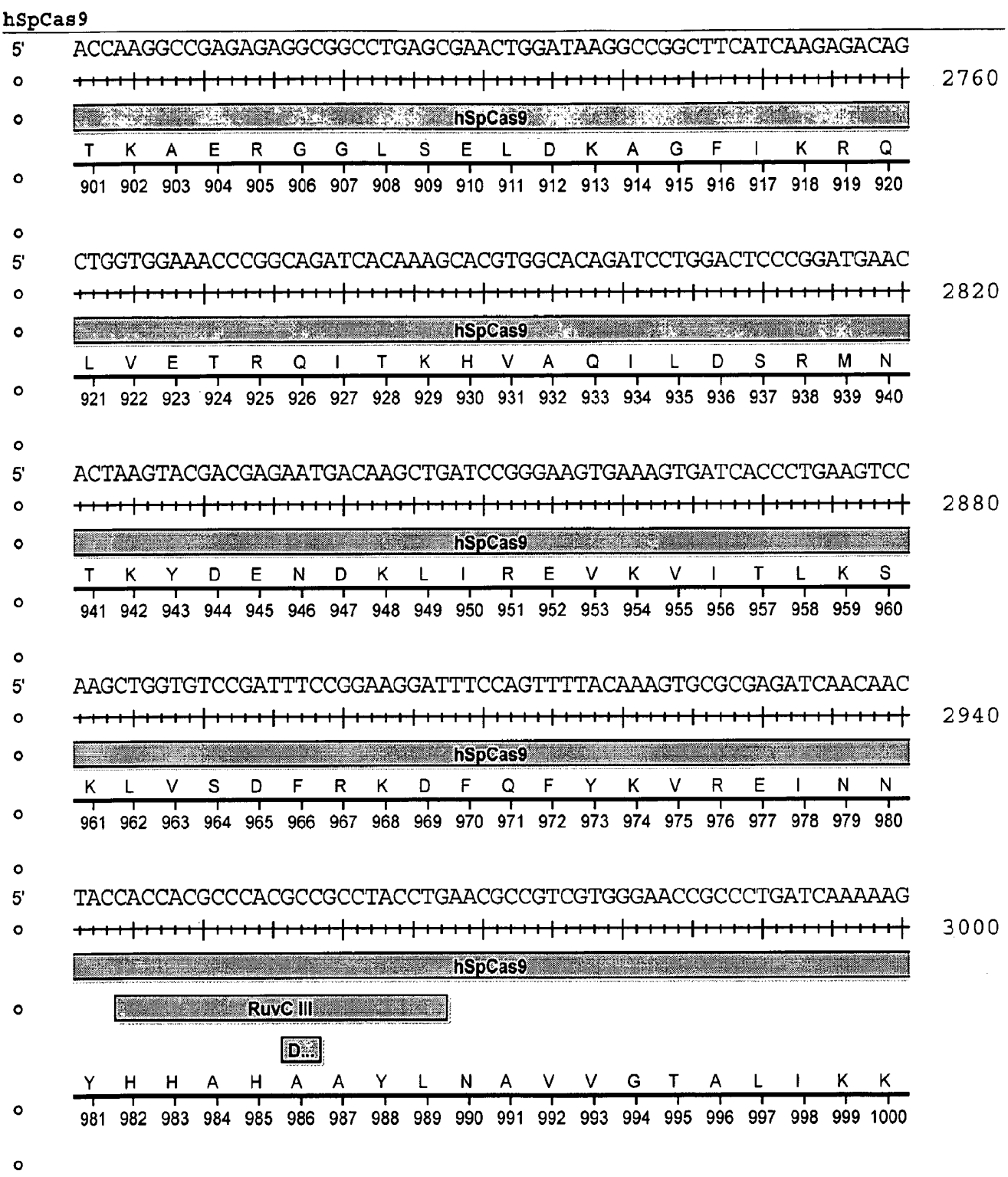

hSpCas9
5' ACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAG
2760
hSpCas9
T K A E R G G L S E L D K A G F I K R Q
901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920
5' CTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAAC
2820
hSpCas9
L V E T R Q I T K H V A Q I L D S R M N
921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940
5' ACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCC
2880
hSpCas9
T K Y D E N D K L I R E V K V I T L K S
941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960
5' AAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAAC
2940
hSpCas9
K L V S D F R K D F Q F Y K V R E I N N
961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980
5' TACCACCACGCCCACGCCGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAG
3000
hSpCas9
RuvC III
D
Y H H A H A A Y L N A V V G T A L I K K
981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

FIG. 10I

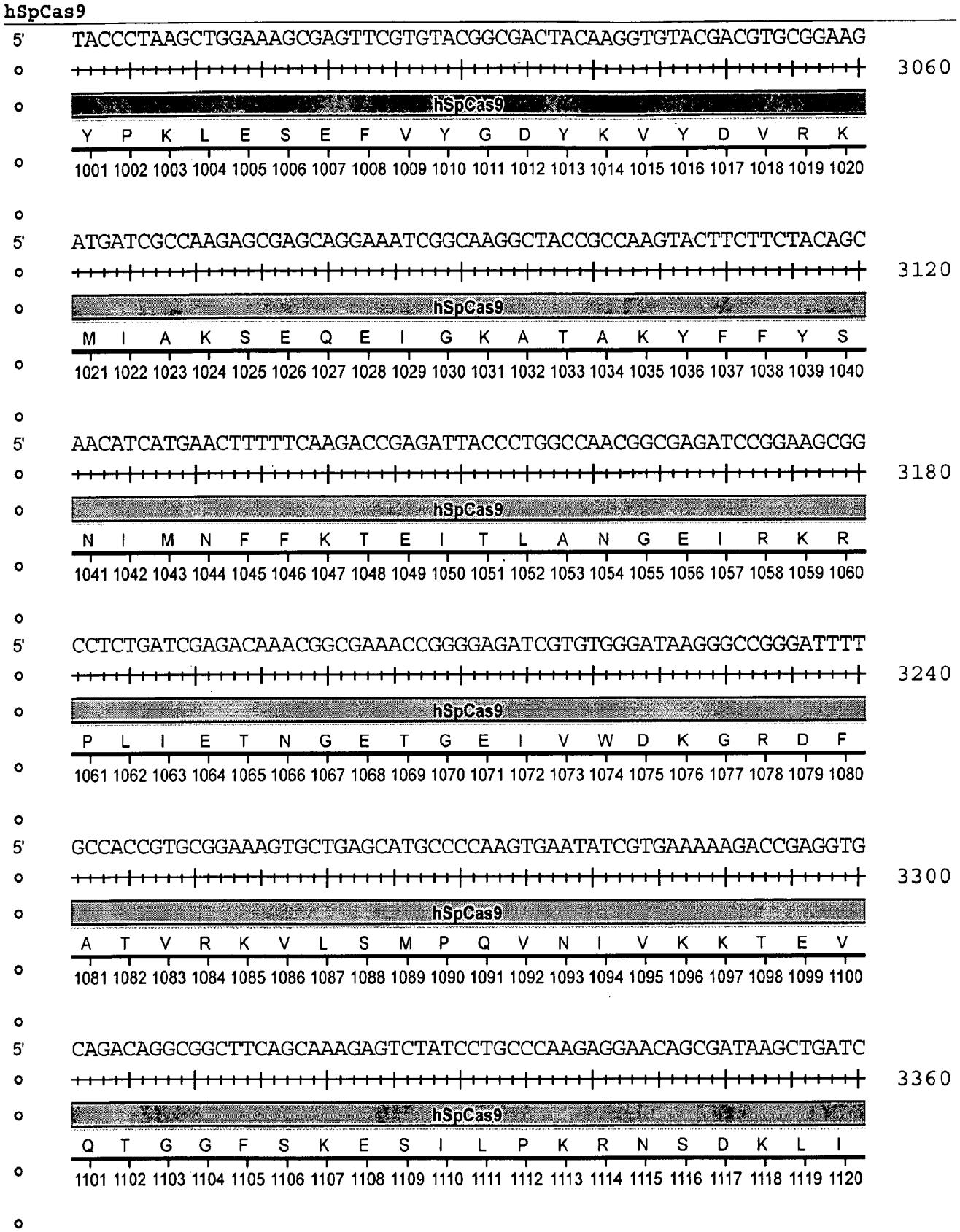

hSpCas9
5' TACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAG
3060
hSpCas9
Y P K L E S E F V Y G D Y K V Y D V R K
1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020
5' ATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGC
3120
hSpCas9
M I A K S E Q E I G K A T A K Y F F Y S
1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040
5' AACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGG
3180
hSpCas9
N I M N F F K T E I T L A N G E I R K R
1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060
5' CCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTT
3240
hSpCas9
P L I E T N G E T G E I V W D K G R D F
1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080
5' GCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTG
3300
hSpCas9
A T V R K V L S M P Q V N I V K K T E V
1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100
5' CAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATC
3360
hSpCas9
Q T G G F S K E S I L P K R N S D K L I
1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120

FIG. 10J

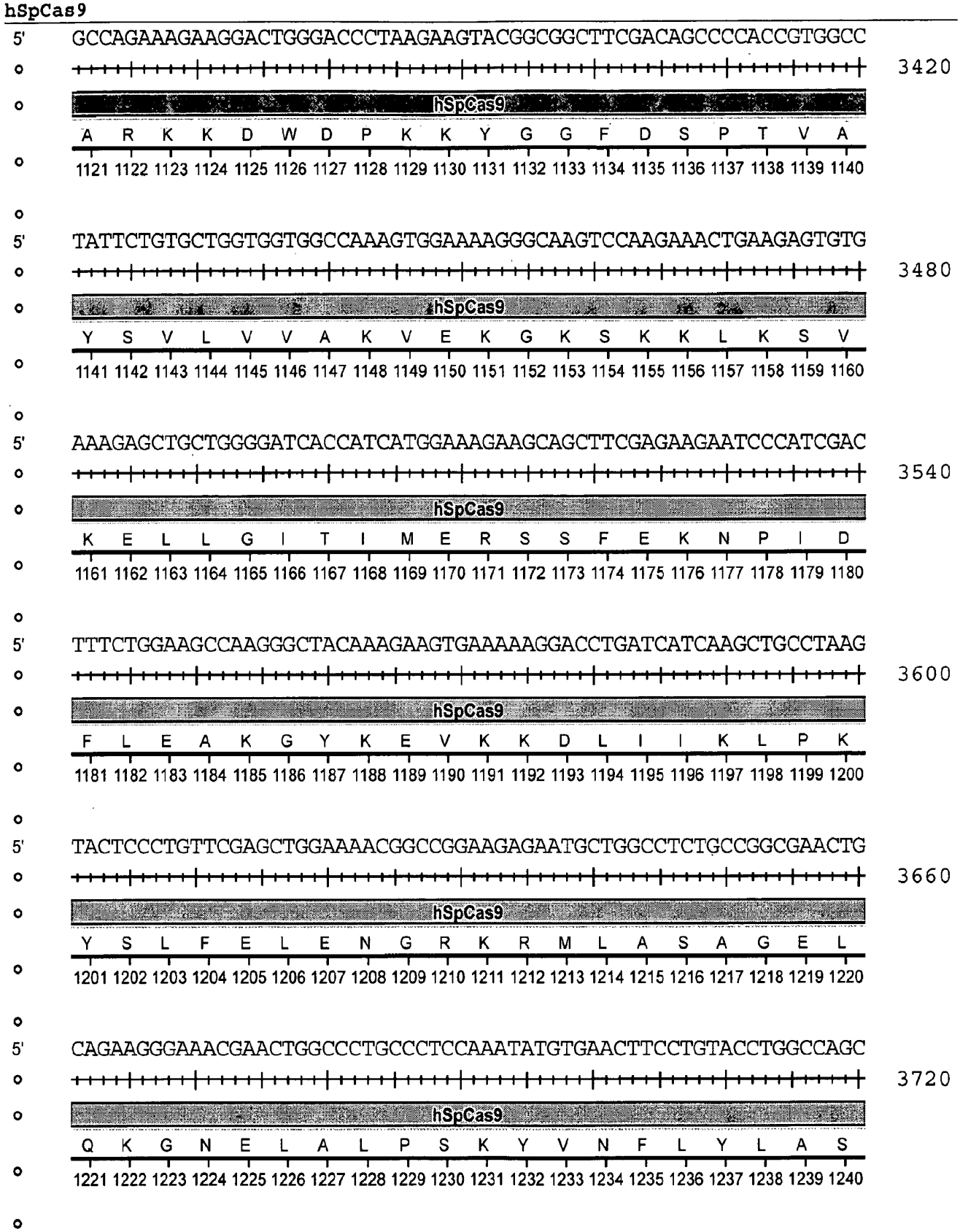

hSpCas9
5' GCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCC
3420
hSpCas9
A R K K D W D P K K Y G G F D S P T V A
1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140
5' TATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTG
3480
hSpCas9
Y S V L V V A K V E K G K S K K L K S V
1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160
5' AAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGAC
3540
hSpCas9
K E L L G I T I M E R S S F E K N P I D
1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180
5' TTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAG
3600
hSpCas9
F L E A K G Y K E V K K D L I I K L P K
1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200
5' TACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTG
3660
hSpCas9
Y S L F E L E N G R K R M L A S A G E L
1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220
5' CAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGC
3720
hSpCas9
Q K G N E L A L P S K Y V N F L Y L A S
1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240

FIG. 10K

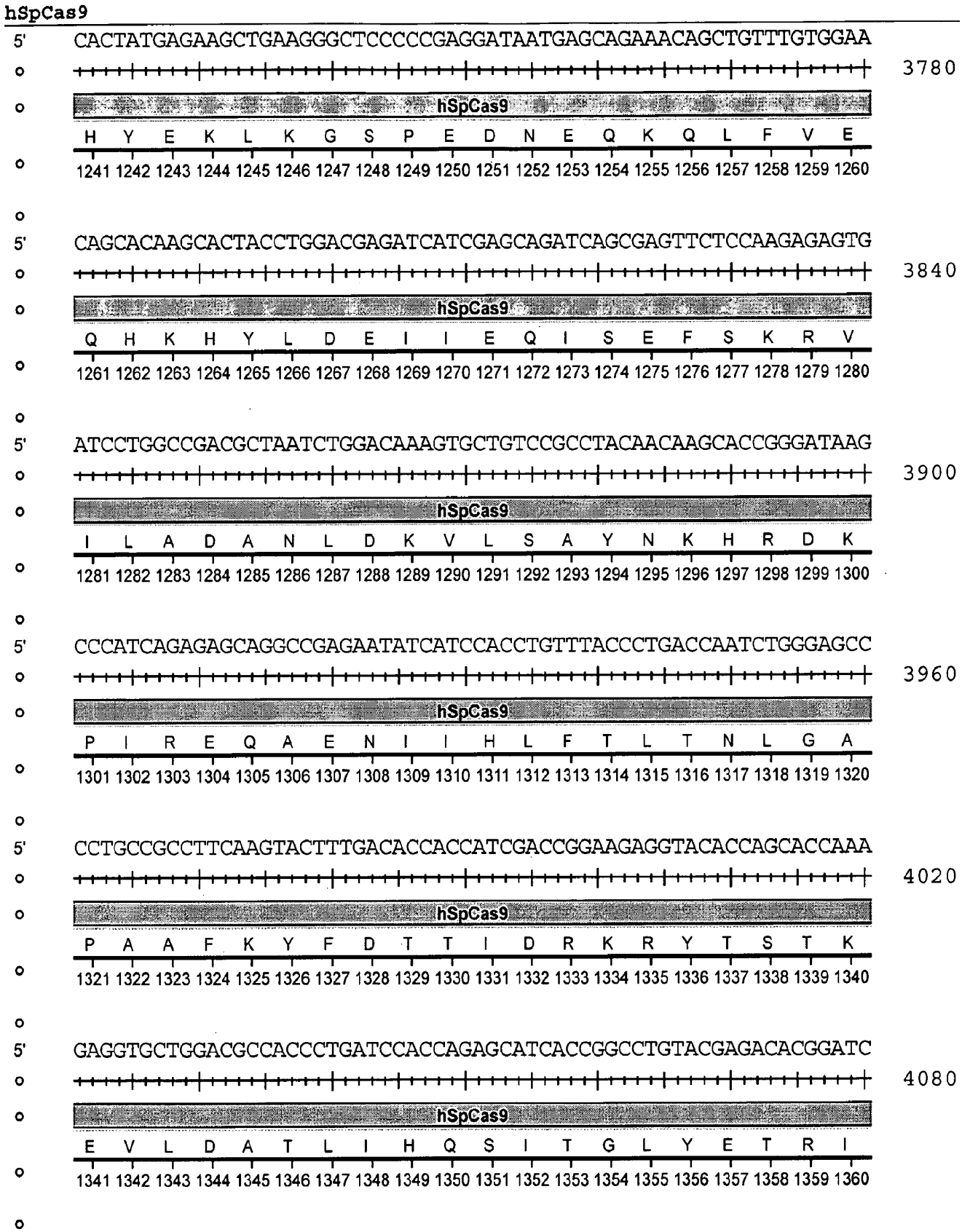

hSpCas9
5' CACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAA 3780
hSpCas9
H Y E K L K G S P E D N E Q K Q L F V E
1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260
5' CAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTG 3840
hSpCas9
Q H K H Y L D E I I E Q I S E F S K R V
1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280
5' ATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAG 3900
hSpCas9
I L A D A N L D K V L S A Y N K H R D K
1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300
5' CCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCC 3960
hSpCas9
P I R E Q A E N I I H L F T L T N L G A
1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320
5' CCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAA 4020
hSpCas9
P A A F K Y F D T T I D R K R Y T S T K
1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340
5' GAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATC 4080
hSpCas9
E V L D A T L I H Q S I T G L Y E T R I
1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360

FIG. 10L

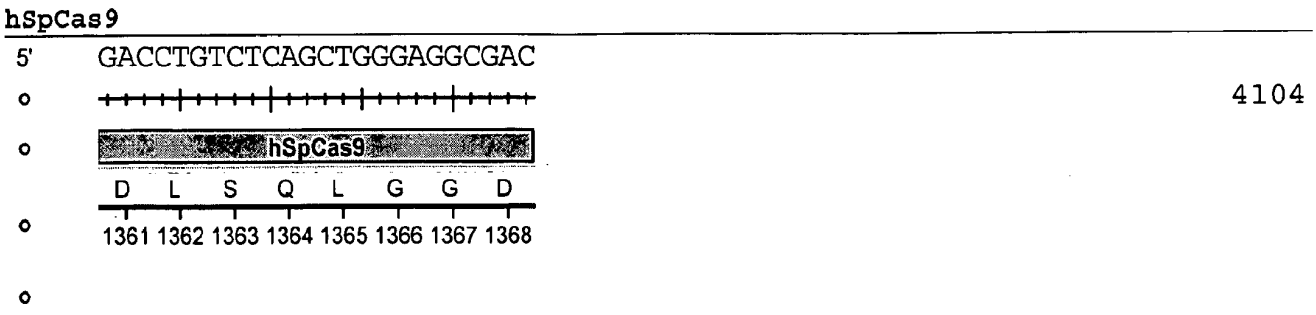

hSpCas9
5'
GACCTGTCTCAGCTGGGAGGCGAC
4104
hSpCas9
D L S Q L G G D
1361 1362 1363 1364 1365 1366 1367 1368

FIG. 10M

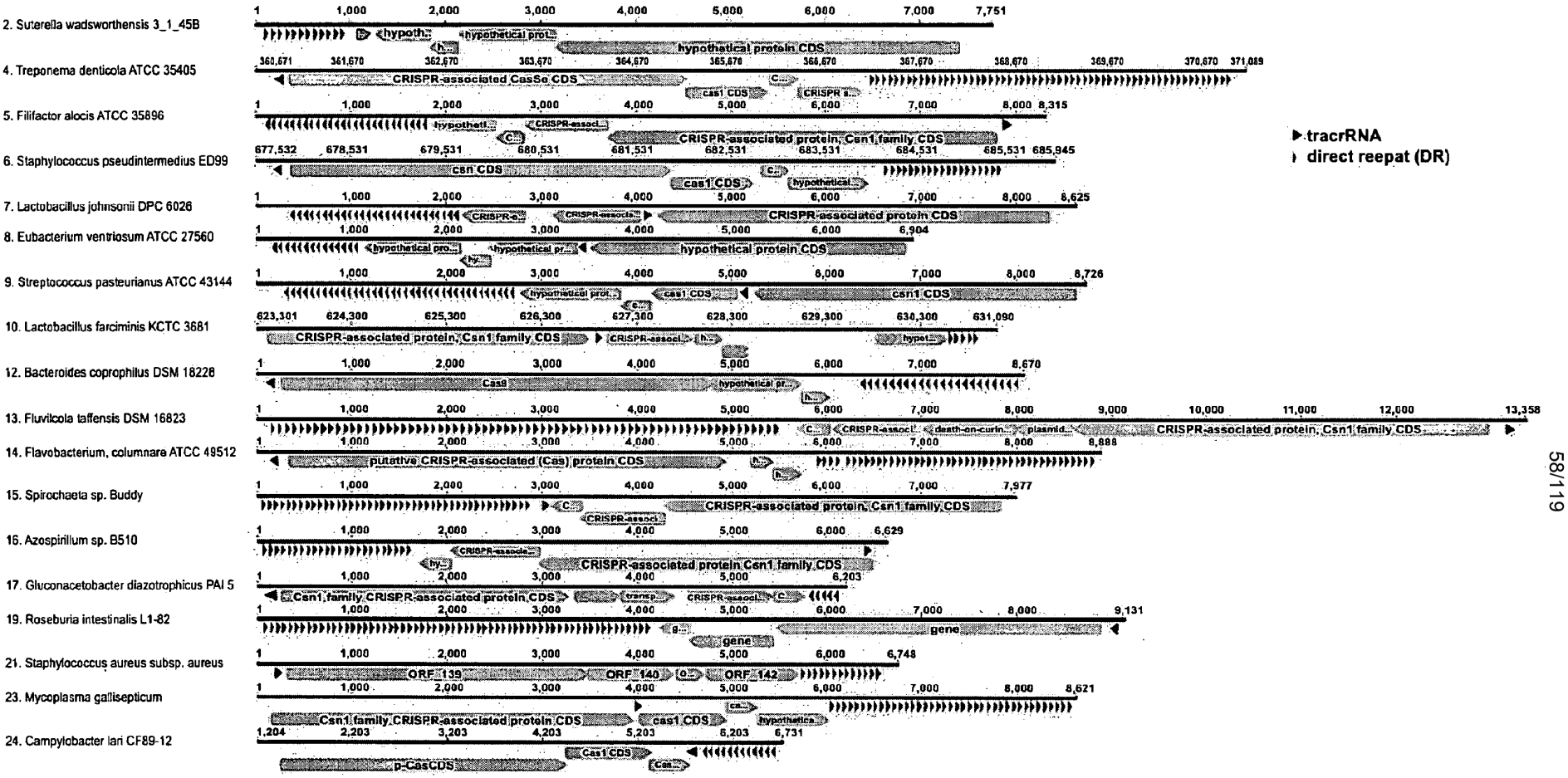

2. Suterella wadsworthensis 3_1_45B
4. Treponema denticola ATCC 35405
5. Filifactor alocis ATCC 35896
6. Staphylococcus pseudintermedius ED99
7. Lactobacillus johnsonii DPC 6026
8. Eubacterium ventriosum ATCC 27560
9. Streptococcus pasteurianus ATCC 43144
10. Lactobacillus farciminis KCTC 3681
12. Bacteroides coprophilus DSM 18228
13. Fluviicola taffensis DSM 16823
14. Flavobacterium, columnare ATCC 49512
15. Spirochaeta sp. Buddy
16. Azospirillum sp. B510
17. Gluconacetobacter diazotrophicus PAI 5
19. Roseburia intestinalis L1-82
21. Staphylococcus aureus subsp. aureus
23. Mycoplasma gallisepticum
24. Campylobacter lari CF89-12
hypothetical protein CDS
CRISPR-associated CasSe CDS
CRISPR-associated protein, Csn1 family CDS
csn CDS
cas1 CDS
CRISPR-associated protein CDS
csn1 CDS
putative CRISPR-associated (Cas) protein CDS
gene
ORF 139
ORF 140
ORF 142
Csn1 family CRISPR-associated protein CDS
p-CasCDS
tracrRNA
direct reepat (DR)

FIG. 11

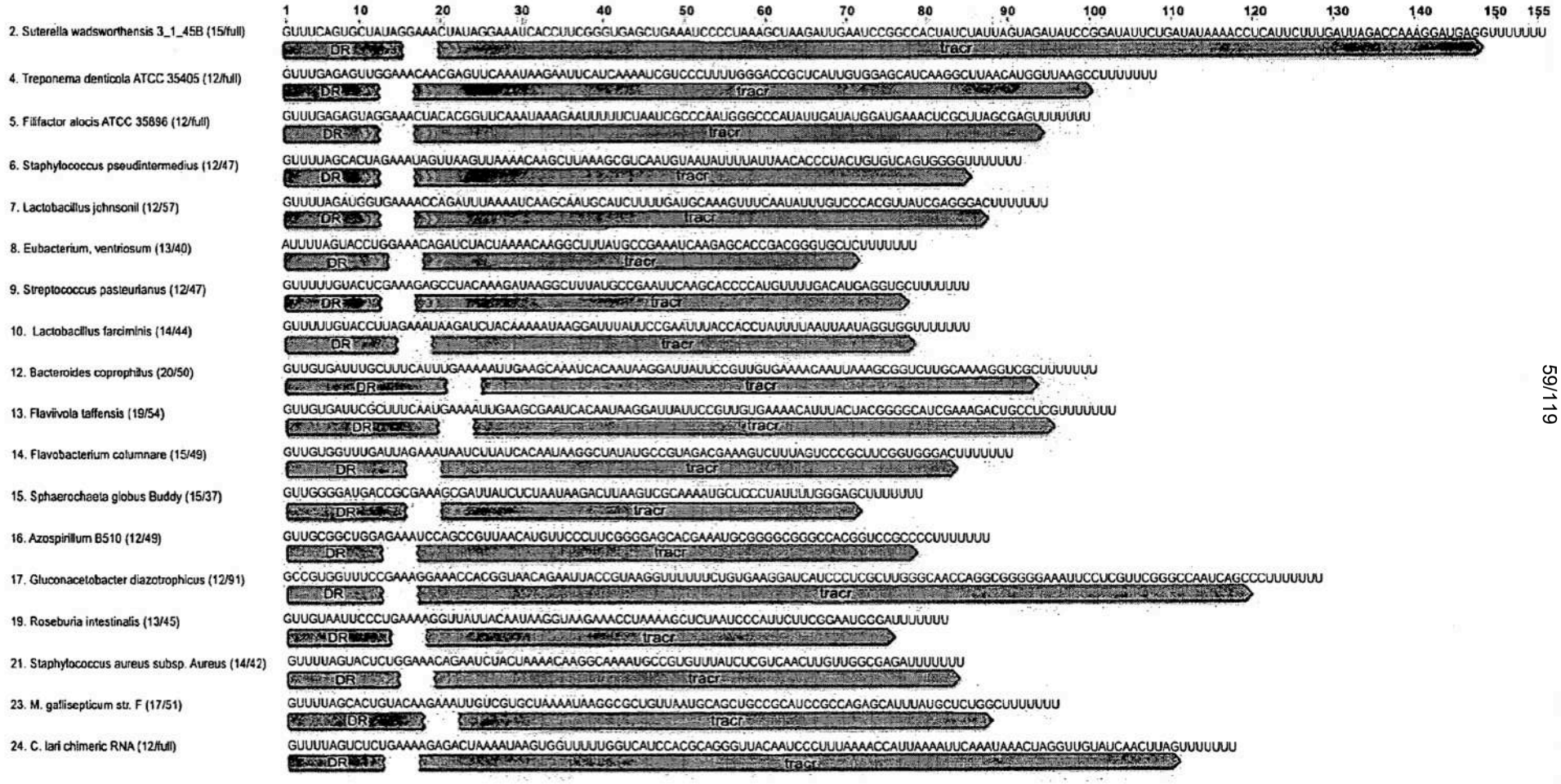

1 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 155
2. Suterella wadsworthensis 3_1_45B (15/full)
GUUUCAGUGCUAUAGGAAACUAUAGGAAAUCACCUUCGGGUGAGCUGAAAUCCCCUAAAGCUAAGAUUGAAUCCGGCCACUAUCUAUUAGUAGAUAUCCGGAUAUUCUGAUAUAAAACCUCAUUCUUUGAUUAGACCAAAGGAUGAGGUUUUUUU
DR
tracr
4. Treponema denticola ATCC 35405 (12/full)
GUUUGAGAGUUGGAAACAACGAGUUCAAAUAAGAAUUCAUCAAAAUCGUCCCUUUUGGGACCGCUCAUUGUGGAGCAUCAAGGCUUAACAUGGUUAAGCCUUUUUUU
DR
tracr
5. Filifactor alocis ATCC 35896 (12/full)
GUUUGAGAGUAGGAAACUACACGGUUCAAAUAAAGAAUUUUUCUAAUCGCCCAAUGGGCCCAUAUUGAUAUGGAUGAAACUCGCUUAGCGAGUUUUUUU
DR
tracr
6. Staphylococcus pseudintermedius (12/47)
GUUUUAGCACUAGAAAUAGUUAAGUUAAAACAAGCUUAAAGCGUCAAUGUAAUAUUUUAUUAACACCCUACUGUGUCAGUGGGGUUUUUUU
DR
tracr
7. Lactobacillus johnsonii (12/57)
GUUUUAGAUGGUGAAAACCAGAUUUAAAAUCAAGCAAUGCAUCUUUUGAUGCAAAGUUUCAAUAUUUGUCCCACGUUAUCGAGGGACUUUUUUU
DR
tracr
8. Eubacterium, ventriosum (13/40)
AUUUUAGUACCUGGAAACAGAUCUACUAAAACAAGGCUUUAUGCCGAAAUCAAGAGCACCGACGGGUGCUCUUUUUUU
DR
tracr
9. Streptococcus pasteurianus (12/47)
GUUUUUGUACUCGAAAGAGCCUACAAAGAUAAGGCUUUAUGCCGAAUUCAAGCACCCCAUGUUUUGACAUGAGGUGCUUUUUUU
DR
tracr
10. Lactobacillus farciminis (14/44)
GUUUUUGUACCUUAGAAAUAAGAUCUACAAAAAUAAGGAUUUAUUCCGAAUUUACCACCUAUUUUAAUUAAUAGGUGGUUUUUUU
DR
tracr
12. Bacteroides coprophilus (20/50)
GUUGUGAUUUGCUUUCAUUUGAAAAAUUGAAGCAAAUCACAAUAAGGAUUAUUCCGUUGUGAAAACAAUUAAAGCGGUCUUGCAAAAGGUCGCUUUUUUU
DR
tracr
13. Flaviivola taffensis (19/54)
GUUGUGAUUCGCUUUCAAUGAAAAUUGAAGCGAAUCACAAUAAGGAUUAUUCCGUUGUGAAAACAUUUACUACGGGGCAUCGAAAGACUGCCUCGUUUUUUU
DR
tracr
14. Flavobacterium columnare (15/49)
GUUGUGGUUUGAUUAGAAAUAAUCUUAUCACAAUAAGGCUAUAUGCCGUAGACGAAAGUCUUUAGUCCCGCUUCGGUGGGACUUUUUUU
DR
tracr
15. Sphaerochaeta globus Buddy (15/37)
GUUGGGGAUGACCGCGAAAGCGAUUAUCUCUAAUAAGACUUAAGUCGCAAAAUGCUCCCUAUUUUGGGAGCUUUUUUU
DR
tracr
16. Azospirillum B510 (12/49)
GUUGCGGCUGGAGAAAUCCAGCCGUUAACAUGUUCCCUUCGGGGAGCACGAAAUGCGGGGCGGGCCACGGUCCGCCCCUUUUUUU
DR
tracr
17. Gluconacetobacter diazotrophicus (12/91)
GCCGUGGUUUCCGAAAGGAAACCACGGUAACAGAAUUACCGUAAGGUUUUUUCUGUGAAGGAUCAUCCCUCGCUUGGGCAACCAGGCGGGGGAAAUUCCUCGUUCGGGCCAAUCAGCCCUUUUUUU
DR
tracr
19. Roseburia intestinalis (13/45)
GUUGUAAUUCCCUGAAAAGGUUAUUACAAUAAGGUAAGAAACCUAAAAGCUCUAAUCCCAUUCUUCGGAAUGCGAUUUUUUU
DR
tracr
21. Staphylococcus aureus subsp. Aureus (14/42)
GUUUUAGUACUCUGGAAACAGAAUCUACUAAAACAAGGCAAAAUGCCGUGUUUAUCUCGUCAACUUGUUGGCGAGAUUUUUUU
DR
tracr
23. M. gallisepticum str. F (17/51)
GUUUUAGCACUGUACAAGAAAUUGUCGUGCUAAAAUAAGGCGCUGUUAAUGCAGCUGCCGCAUCCGCCAGAGCAUUUAUGCUCUGGCUUUUUUU
DR
tracr
24. C. lari chimeric RNA (12/full)
GUUUUAGUCUCUGAAAAGAGACUAAAAUAAGUGGUUUUUGGUCAUCCACGCAGGGUUACAAUCCCUUUAAAACCAUUAAAAUUCAAAUAAACUAGGUUGUAUCAACUUAGUUUUUUU
DR
tracr

FIG. 12

FIG. 13A

Species: Corynebacter_diphtheriae	
Sequencing Barcode	AGCAATTC
ID	1
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNNACUGGGGUUCAGGAAACUGAACCUCAGUAAGCAUUGGCUCGUUUCCAAUG UUGAUUGCUCCGCCGGUGCUCCUUAUUUUUAAGGGCGCCGGCUUUUUUU**
Cas9 Protein Sequence (wildtype, without NLS added)	MKYHVGIDVGTFSVGLAAIEVDDAGMPIKTLSLVSHIHDSGLDPDEIKSAVTRLASSGIARRTRRLYRRKRRRLQQLDK FIQRQGWPVIELEDYSDPLYPWKVRAELAASYIADEKERGEKLSVALRHIARHRGWRNPYAKVSSLYLPDGPSDAFKA IREEIKRASGQPVPETATVGQMVTLCELGTLKLRGEGGVLSARLQQSDYAREIQEICRMQEIGQELYRKIIDVVFAAESP KGSASSRVGKDPLQPGKNRALKASDAFQRYRIAALIGNLRVRVDGEKRILSVEEKNLVFDHLVNLTPKKEPEWVTIAEI LGIDRGQLIGTATMTDDGERAGARPPTHDTNRSIVNSRIAPLVDWWKTASALEQHAMVKALSNAEVDDFDSPEGAKV QAFFADLDDDVHAKLDSLHLPVGRAAYSEDTLVRLTRRMLSDGVDLYTARLQEFGIEPSWTPPTPRIGEPVGNPAVDR VLKTVSRWLESATKTWGAPERVIIEHVREGFVTEKRAREMDGDMRRRAARNAKLFQEMQEKLNVQGKPSRADLWR YQSVQRQNCQCAYCGSPITFSNSEMDHIVPRAGQGSTNTRENLVAVCHRCNQSKGNTPFAIWAKNTSIEGVSVKEAVE RTRHWVTDTGMRSTDFKKFTKAVVERFQRATMDEEIDARSMESVAWMANELRSRVAQHFASHGTTVRVYYRGSLTA EARRASGISGKLKFFDGVGKSRLDRRHHAIDAAVIAFTSDYVAETLAVRSNLKQSQAHRQEAPQWREFTGKDAEHRA AWRVWCQKMEKLSALLTEDLRDDRVVVMSNVRLRLGNGSAHKETIGKLSKVKLSSQLSVSDIDKASSEALWCALTR EPGFDPKEGLPANPERHIRVNGTHVYAGDNIGLFPVSAGSIALRGGYAELGSSFHHARVYKITSGKKPAFAMLRVYTID LLPYRNQDLFSVELKPQTMSMRQAEKKLRDALATGNAEYLGWLVVDDELVVDTSKIATDQVKAVEAELGTIRRWRV DGFFSPSKLRLRPLQMSKEGIKKESAPELSKIIDRPGWLPAVNKLFSDGNVTVVRRDSLGRVRLESTAHLPVTWKVQ
Mammalian Codon Optimized Sequences (AgeI-HA-NLS-Cas9-NLS-taa-EcoRI)	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CATGAAGTACCATGTCGGAATCGATGTCGGAACCTTTTCTGTGGGGCTGGCTGCTATTGAAGTGGATGACGCTGG AATGCCTATTAAGACCCTGAGTCTGGTGTCACACATTCATGACTCAGGACTGGATCCTGACGAGATCAAGAGCGC TGTGACCAGGCTGGCAAGCTCCGGAATCGCCCGGAGAACAAGGCGCCTGTACCGACGGAAGAGAAGGCGCCTGC AGCAGCTGGATAAGTTCATCCAGAGGCAGGGCTGGCCAGTGATCGAGCTGGAAGATTACAGCGACCCCCTGTAT CCTTGGAAGGTGCGCGCCGAACTGGCCGCTTCTTATATTGCTGACGAGAAGGAACGGGGGGAGAAACTGAGTGT GGCTCTGAGACACATCGCAAGGCATCGCGGATGGAGGAACCCTTACGCCAAGGTGTCTAGTCTGTATCTGCCAGA TGGCCCCTCAGACGCCTTCAAGGCTATTAGGGAGGAAATCAAACGCGCTAGCGGCCAGCCTGTGCCAGAGACTG CAACCGTCGGGCAGATGGTGACCCTGTGCGAACTGGGCACACTGAAGCTGCGAGGAGAGGGAGGAGTGCTGAGT GCACGGCTGCAGCAGTCAGATTACGCCCGCGAGATCCAGGAAATTTGTCGAATGCAGGAGATCGGCCAGGAACT GTATCGCAAGATCATTGACGTGGTGTTCGCAGCCGAGTCCCCAAAGGGCTCTGCCTCAAGCCGGGTGGGGAAAG ATCCTCTGCAGCCAGGAAAGAACAGAGCACTGAAAGCCAGCGACGCTTTTCAGCGATACCGGATTGCTGCACTG

FIG. 13B

	ATCGGCAATCTGAGAGTCAGGGTGGATGGGGAGAAGAGGATTCTGAGCGTGGAGGAGAAGAACCTGGTGTTCGA CCACCTGGTGAATCTGACTCCAAAGAAAGAGCCCGAATGGGTGACCATCGCCGAAATTCTGGGCATCGATCGCG GGCAGCTGATCGGAACAGCTACTATGACCGACGATGGAGAGCGAGCAGGAGCCCGACCCCCTACACACGATACT AACAGAAGTATTGTGAACAGCCGGATCGCACCACTGGTCGACTGGTGGAAAACAGCTAGCGCACTGGAGCAGCA CGCCATGGTGAAGGCACTGTCCAACGCCGAAGTCGACGATTTTGATTCTCCCGAGGGAGCAAAAGTGCAGGCAT TCTTTGCCGATCTGGACGATGACGTCCACGCCAAGCTGGACAGCCTGCATCTGCCTGTGGGACGAGCAGCTTACT CCGAGGACACTCTGGTCAGACTGACCCGACGGATGCTGAGTGATGGGGTGGACCTGTATACCGCCCGGCTGCAG GAGTTCGGAATTGAACCTAGCTGGACCCCACCCACACCAAGAATCGGAGAGCCTGTCGGCAATCCAGCCGTCGA CCGGGTGCTGAAAACAGTGAGCAGATGGCTGGAATCCGCAACAAAGACTTGGGGCGCCCCAGAGAGGGTCATCA TTGAGCACGTGCGCGAAGGCTTCGTCACTGAGAAACGCGCTCGAGAAATGGATGGGGACATGAGAAGGCGCGCA GCCCGGAACGCCAAGCTGTTTCAGGAGATGCAGGAAAAGCTGAATGTGCAGGGCAAACCCAGTCGAGCCGATCT GTGGAGATACCAGTCAGTGCAGAGACAGAACTGCCAGTGTGCCTATTGCGGGTCCCCAATTACCTTTTCTAATAG TGAAATGGACCACATCGTGCCCAGAGCAGGGCAGGGATCCACCAACACAAGGGAGAATCTGGTCGCCGTGTGCC ATCGCTGTAACCAGTCTAAGGGCAATACACCCTTCGCTATTTGGGCAAAAAACACTTCTATCGAAGGGGTCAGTG TGAAGGAGGCCGTGGAACGGACCAGACATTGGGTCACTGATACCGGCATGAGAAGCACTGACTTCAAGAAGTTC ACCAAGGCTGTGGTCGAGCGGTTTCAGAGAGCAACAATGGATGAGGAAATCGACGCCAGAAGCATGGAATCCGT CGCCTGGATGGCTAATGAGCTGAGGAGCCGCGTGGCTCAGCACTTCGCATCCCATGGAACCACAGTCAGGGTGT ACCGAGGCAGCCTGACAGCAGAGGCTCGACGGGCATCTGGGATCAGTGGAAAGCTGAAATTCTTTGATGGCGTG GGGAAGTCCAGGCTGGATAGAAGGCACCATGCTATTGACGCTGCAGTGATCGCATTCACCTCTGACTATGTGGCC GAAACACTGGCTGTCCGCTCAAACCTGAAACAGAGCCAGGCCCACCGACAGGAGGCTCCTCAGTGGAGAGAGTT CACCGGCAAGGATGCAGAGCATCGAGCAGCTTGGAGAGTGTGGTGCCAGAAGATGGAAAAACTGAGCGCCCTGC TGACCGAGGACCTGCGAGATGACCGGGTGGTCGTGATGTCTAACGTGCGACTGCGGCTGGGAAATGGCAGTGCC CACAAGGAAACCATTGGCAAACTGTCAAAGGTGAAACTGTCCTCTCAGCTGTCAGTCAGCGATATCGACAAAGC AAGTTCAGAGGCCCTGTGGTGTGCTCTGACCAGAGAGCCCGGATTCGATCCTAAGGAAGGCCTGCCCGCTAACCC TGAGAGACACATCAGGGTGAATGGGACACATGTCTACGCCGGGGACAATATTGGACTGTTTCCAGTGTCAGCAG GAAGCATCGCACTGAGGGGAGGATACGCAGAGCTGGGCAGCTCCTTCCACCATGCTCGCGTGTATAAAATTACTT CCGGCAAGAAACCCGCATTTGCCATGCTGAGGGTGTACACCATCGATCTGCTGCCTTATCGCAACCAGGACCTGT TTAGCGTGGAACTGAAGCCACAGACAATGTCCATGAGGCAGGCTGAGAAGAAACTGCGCGACGCTCTGGCAACT GGGAATGCAGAATATCTGGGATGGCTGGTCGTGGATGACGAGCTGGTCGTGGATACATCTAAGATTGCCACTGA CCAGGTCAAAGCAGTGGAGGCCGAACTGGGGACTATCCGCCGATGGCGGGTGGATGGATTCTTTTCCCCCTCTAA ACTGAGACTGAGGCCTCTGCAGATGTCCAAGGAGGGGATCAAGAAAGAGTCCGCTCCCGAACTGTCTAAAATCA TTGACAGACCAGGATGGCTGCCCGCCGTGAACAAGCTGTTCTCTGATGGAAATGTCACCGTCGTGCGGAGAGACT CTCTGGGACGCGTGCGACTGGAGAGTACAGCCCACCTGCCTGTCACTTGGAAGGTGCAGTAAGAATTC
Length of Protein	1084

FIG. 13C

(without NLS and HA tag)	
PAM	**NGG**
Spacer Length	28
Species: Sutterella_wadsworthensis	
Sequencing Barcode	AACTTGAC
ID	2
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNGUUUCAGUGCUAUAGGAAACUAUAGGAAAUCACCUUCGGGUGAGCUGAAA UCCCCUAAAGCUAAGAUUGAAUCCGGCCACUAUCUAUUAGUAGAUAUCCGGAUAUUCUGAUAUAAAACCU CAUUCUUUGAUUAGACCAAAGGAUGAGGUUUUUUU**
Cas9 Protein Sequence (wildtype, without NLS added)	MTQSERRFSCSIGIDMGAKYTGVFYALFDREELPTNLNSKAMTLVMPETGPRYVQAQRTAVRHRLRGQKRYTLARKL AFLVVDDMIKKQEKRLTDEEWKRGREALSGLLKRRGYSRPNADGEDLTPLENVRADVFAAHPAFSTYFSEVRSLAEQ WEEFTANISNVEKFLGDPNIPADKEFIEFAVAEGLIDKTEKKAYQSALSTLRANANVLTGLRQMGHKPRSEYFKAIEAD LKKDSRLAKINEAFGGAERLARLLGNLSNLQLRAERWYFNAPDIMKDRGWEPDRFKKTLVRAFKFFHPAKDQNKQH LELIKQIENSEDIIETLCTLDPNRTIPPYEDQNNRRPPLDQTLLLSPEKLTRQYGEIWKTWSARLTSAEPTLAPAAEILERS TDRKSRVAVNGHEPLPTLAYQLSYALQRAFDRSKALDPYALRALAAGSKSNKLTSARTALENCIGGQNVKTFLDCAR RYYREADDAKVGLWFDNADGLLERSDLHPPMKKKILPLLVANILQTDETTGQKFLDEIWRKQIKGRETVASRCARIET VRKSFGGGFNIAYNTAQYREVNKLPRNAQDKELLTIRDRVAETADFIAANLGLSDEQKRKFANPFSLAQFYTLIETEVS GFSATTLAVHLENAWRMTIKDAVINGETVRAAQCSRLPAETARPFDGLVRRLVDRQAWEIAKRVSTDIQSKVDFSNGI VDVSIFVEENKFEFSASVADLKKNKRVKDKMLSEAEKLETRWLIKNERIKKASRGTCPYTGDRLAEGGEIDHILPRSLI KDARGIVFNAEPNLIYASSRGNQLKKNQRYSLSDLKANYRNEIFKTSNIAAITAEIEDVVTKLQQTHRLKFFDLLNEHE QDCVRHALFLDDGSEARDAVLELLATQRRTRVNGTQIWMIKNLANKIREELQNWCKTTNNRLHFQAAATNVSDAKN LRLKLAQNQPDFEKPDIQPIASHSIDALCSFAVGSADAERDQNGFDYLDGKTVLGLYPQSCEVIHLQAKPQEEKSHFDS VAIFKEGIYAEQFLPIFTLNEKIWIGYETLNAKGERCGAIEVSGKQPKELLEMLAPFFNKPVGDLSAHATYRILKKPAYE FLAKAALQPLSAEEKRLAALLDALRYCTSRKSLMSLFMAANGKSLKKREDVLKPKLFQLKVELKGEKSFKLNGSLTLP VKQDWLRICDSPELADAFGKPCSADELTSKLARIWKRPVMRDLAHAPVRREFSLPAIDNPSGGFRIRRTNLFGNELYQ VHAINAKKYRGFASAGSNVDWSKGILFNELQHENLTECGGRFITSADVTPMSEWRKVVAEDNLSIWIAPGTEGRRYV RVETTFIQASHWFEQSVENWAITSPLSLPASFKVDKPAEFQKAVGTELSELLGQPRSEIFIENVGNAKHIRFWYIVVSSN KKMNESYNNVSKS

FIG. 13D

Mammalian Codon Optimized Sequences (AgeI-HA-NLS-Cas9-NLS-taa-EcoRI)	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CATGACTCAGAGCGAGCGACGATTTTCTTGCAGCATTGGCATTGACATGGGGGCTAAGTACACTGGGGTGTTCTA CGCACTGTTCGACCGGGAGGAACTGCCCACAAACCTGAACAGCAAGGCCATGACCCTGGTCATGCCTGAGACAG GGCCAAGATACGTGCAGGCACAGAGAACTGCCGTCAGACACAGGCTGCGCGGACAGAAGAGATATACCCTGGCT AGGAAACTGGCATTTCTGGTGGTCGACGATATGATCAAGAAACAGGAAAAGAGGCTGACTGATGAGGAATGGAA ACGAGGACGGGAGGCCCTGTCCGGCCTGCTGAAGCGGAGAGGGTACTCTCGGCCCAACGCTGACGGCGAAGATC TGACCCCTCTGGAGAATGTGAGAGCAGACGTGTTCGCCGCTCATCCTGCCTTCAGCACATATTTTTCCGAAGTGC GCTCTCTGGCTGAGCAGTGGGAGGAGTTCACCGCAAACATCAGCAATGTCGAGAAGTTTCTGGGCGACCCAAAC ATCCCCGCCGATAAAGAGTTCATTGAATTTGCCGTGGCTGAAGGGCTGATTGACAAGACCGAGAAGAAAGCCTA CCAGTCAGCTCTGAGCACCCTGAGGGCAAACGCCAATGTGCTGACAGGACTGCGGCAGATGGGCCACAAGCCTA GATCAGAATATTTTAAAGCAATCGAGGCCGACCTGAAGAAAGATAGCCGCCTGGCCAAGATTAACGAAGCATTC GGAGGAGCAGAGCGCCTGGCTCGACTGCTGGGAAACCTGTCCAATCTGCAGCTGCGGGCAGAAAGATGGTACTT CAATGCCCCCGACATCATGAAGGATAGGGGCTGGGAGCCTGATCGCTTCAAGAAAACACTGGTGCGGGCTTTTA AGTTCTTTCACCCAGCAAAGGACCAGAACAAACAGCATCTGGAACTGATCAAACAGATTGAGAACAGCGAAGAT ATCATTGAGACTCTGTGCACCCTGGACCCAAACAGAACCATCCCCCCTTACGAGGATCAGAACAATAGGCGCCCA CCCCTGGACCAGACTCTGCTGCTGAGTCCCGAAAAGCTGACCCGGCAGTATGGCGAGATCTGGAAAACATGGAG CGCCAGACTGACCTCCGCTGAACCCACACTGGCACCTGCAGCCGAGATTCTGGAAAGATCTACCGACAGGAAGA GTCGCGTGGCAGTCAACGGACACGAGCCACTGCCTACACTGGCTTACCAGCTGAGTTATGCACTGCAGAGAGCCT TCGACAGGTCAAAAGCCCTGGATCCATATGCTCTGAGGGCACTGGCTGCAGGCTCAAAAAGCAATAAGCTGACA TCCGCCCGCACTGCTCTGGAGAACTGCATCGGAGGCCAGAATGTGAAAACCTTCCTGGACTGTGCCCGACGGTAC TATCGGGAAGCAGACGATGCCAAAGTCGGGCTGTGGTTCGACAACGCCGATGGACTGCTGGAGAGATCTGACCT GCATCCTCCAATGAAGAAAAAGATCCTGCCCCTGCTGGTGGCCAATATTCTGCAGACAGATGAAACCACAGGCC AGAAGTTTCTGGACGAGATCTGGCGAAAACAGATTAAGGGGCGGGAAACTGTGGCTAGCCGATGTGCACGGATC GAGACAGTGCGGAAATCCTTCGGGGGAGGCTTTAACATTGCCTACAATACCGCTCAGTATAGGGAGGTGAACAA GCTGCCCCGCAATGCCCAGGATAAAGAACTGCTGACAATCAGAGATAGGGTGGCTGAGACTGCAGACTTCATTG CCGCTAACCTGGGGCTGTCTGACGAGCAGAAAAGAAAGTTCGCCAATCCTTTTAGTCTGGCTCAGTTCTACACCC TGATCGAGACAGAAGTGTCCGGATTTTCTGCAACTACCCTGGCCGTCCACCTGGAGAACGCCTGGAGGATGACAA TCAAGGATGCTGTGATTAATGGGGAAACTGTCAGAGCAGCACAGTGCAGCAGGCTGCCTGCAGAGACAGCTCGC CCATTCGATGGACTGGTGAGAAGGCTGGTCGACAGACAGGCTTGGGAGATCGCAAAGAGGGTGTCAACTGACAT TCAGAGCAAAGTCGATTTCTCCAACGGCATCGTGGACGTCAGCATTTTTGTGGAGGAAAATAAGTTCGAGTTTTC CGCATCTGTGGCCGATCTGAAAAAGAACAAACGGGTCAAAGACAAGATGCTGTCCGAGGCCGAAAAGCTGGAA ACCAGATGGCTGATCAAAAATGAGCGGATCAAGAAGGCCAGCCGGGGAACTTGTCCCTACACCGGCGATAGGCT GGCTGAGGGGGGAGAAATCGACCACATTCTGCCCCGAAGCCTGATCAAGGATGCCCGGGGAATTGTGTTTAACG CTGAGCCTAATCTGATCTATGCAAGCTCCCGCGGCAACCAGCTGAAAAAGAATCAGCGATACAGTCTGTCAGATC TGAAGGCCAACTATCGGAATGAGATCTTCAAAACTAGCAACATCGCTGCAATTACCGCCGAGATTGAGGACGTG

FIG. 13E

	GTCACTAAGCTGCAGCAGACCCATAGACTGAAATTCTTTGATCTGCTGAATGAGCACGAACAGGACTGCGTGCGG CACGCCCTGTTCCTGGACGATGGCAGCGAAGCTCGCGACGCAGTGCTGGAGCTGCTGGCAACACAGCGCCGAAC TCGCGTCAACGGGACACAGATCTGGATGATTAAGAACCTGGCCAACAAGATCCGAGAGGAACTGCAGAATTGGT GTAAGACAACTAACAATAGACTGCACTTTCAGGCCGCTGCAACTAACGTGTCCGATGCAAAGAATCTGAGGCTG AAACTGGCCCAGAACCAGCCCGACTTCGAGAAGCCAGATATCCAGCCCATTGCCAGCCATTCCATCGACGCCCTG TGCTCTTTCGCTGTGGGGAGTGCTGACGCAGAACGCGATCAGAATGGATTTGACTACCTGGATGGCAAGACCGTG CTGGGACTGTATCCACAGAGCTGTGAGGTCATTCACCTGCAGGCCAAGCCCCAGGAGGAAAAAAGTCATTTCGA TTCAGTGGCTATCTTTAAGGAAGGCATCTACGCAGAGCAGTTCCTGCCTATCTTTACCCTGAACGAAAAGATCTG GATTGGATATGAGACACTGAATGCCAAAGGCGAAAGATGCGGGGCTATTGAGGTGAGTGGCAAACAGCCAAAG GAGCTGCTGGAAATGCTGGCCCCCTTCTTTAACAAGCCTGTGGGCGACCTGTCAGCCCACGCTACTTACCGGATC CTGAAAAAGCCTGCATATGAGTTTCTGGCAAAGGCAGCTCTGCAGCCACTGAGCGCAGAGGAAAAAAGACTGGC AGCCCTGCTGGATGCTCTGCGCTACTGTACCAGTCGAAAGTCACTGATGAGCCTGTTCATGGCTGCAAACGGGAA ATCCCTGAAAAAGCGGGAGGACGTGCTGAAACCCAAGCTGTTCCAGCTGAAGGTCGAGCTGAAAGGCGAAAAG AGCTTCAAGCTGAACGGGAGCCTGACCCTGCCTGTGAAGCAGGACTGGCTGAGAATCTGCGATAGCCCAGAACT GGCAGACGCCTTTGGCAAACCCTGTTCCGCCGATGAGCTGACATCTAAGCTGGCTCGCATTTGGAAACGACCTGT GATGCGGGATCTGGCTCATGCACCAGTCCGGAGAGAGTTCAGCCTGCCCGCAATCGACAACCCAAGTGGAGGGT TCAGGATTAGGCGCACCAACCTGTTTGGCAATGAGCTGTACCAGGTGCACGCCATCAACGCTAAAAAGTATCGCG GCTTCGCCTCCGCTGGGTCTAATGTCGACTGGTCCAAGGGGATCCTGTTTAACGAGCTGCAGCATGAAAATCTGA CCGAGTGCGGAGGCAGGTTCATTACAAGCGCCGATGTGACTCCTATGTCCGAATGGCGCAAGGTGGTCGCAGAG GACAACCTGTCTATCTGGATTGCTCCAGGGACAGAAGGACGACGGTACGTGAGGGTCGAGACAACATTCATCCA GGCCAGTCACTGGTTTGAACAGTCAGTGGAGAATTGGGCCATTACTAGTCCTCTGTCACTGCCAGCTTCCTTCAA GGTGGACAAACCAGCTGAGTTTCAGAAGGCAGTCGGAACCGAGCTGTCAGAACTGCTGGGCCAGCCCAGGAGCG AAATCTTCATTGAGAACGTGGGCAATGCCAAGCATATCCGCTTTTGGTACATTGTGGTGAGCAGCAACAAAAAGA TGAACGAGTCTTACAACAATGTGTCTAAGAGTTAAGAATTC
Length of Protein (without NLS and HA tag)	1422
PAM	**NGG**
Spacer Length	35-39
Species: Legionella_pneumophila_Paris	

FIG. 13F

Sequencing Barcode	TGTCGGAT
ID	3
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNGUUUCAGUGGUGAAAACCCUGAAAUCAACAAAAUUAAAGAUUGAAUCGUU UUUCUAUGCUCGUCUUAAUAGCGAGCAUAUAACGAUUUUUUU**
Cas9 Protein Sequence (wildtype, without NLS added)	MESSQILSPIGIDLGGKFTGVCLSHLEAFAELPNHANTKYSVILIDHNNFQLSQAQRRATRHRVRNKKRNQFVKRVALQ LFQHILSRDLNAKEETALCHYLNNRGYTYVDTDLDEYIKDETTINLLKELLPSESEHNFIDWFLQKMQSSEFRKILVSKV EEKKDDKELKNAVKNIKNFITGFEKNSVEGHRHRKVYFENIKSDITKDNQLDSIKKKIPSVCLSNLLGHLSNLQWKNLH RYLAKNPKQFDEQTFGNEFLRMLKNFRHLKGSQESLAVRNLIQQLEQSQDYISILEKTPPEITIPPYEARTNTGMEKDQS LLLNPEKLNNLYPNWRNLIPGIIDAHPFLEKDLEHTKLRDRKRIISPSKQDEKRDSYILQRYLDLNKKIDKFKIKKQLSFL GQGKQLPANLIETQKEMETHFNSSLVSVLIQIASAYNKEREDAAQGIWFDNAFSLCELSNINPPRKQKILPLLVGAILSE DFINNKDKWAKFKIFWNTHKIGRTSLKSKCKEIEEARKNSGNAFKIDYEEALNHPEHSNNKALIKIIQTIPDIIQAIQSHL GHNDSQALIYHNPFSLSQLYTILETKRDGFHKNCVAVTCENYWRSQKTEIDPEISYASRLPADSVRPFDGVLARMMQR LAYEIAMAKWEQIKHIPDNSSLLIPIYLEQNRFEFEESFKKIKGSSSDKTLEQAIEKQNIQWEEKFQRIINASMNICPYKG ASIGGQGEIDHIYPRSLSKKHFGVIFNSEVNLIYCSSQGNREKKEEHYLLEHLSPLYLKHQFGTDNVSDIKNFISQNVANI KKYISFHLLTPEQQKAARHALFLDYDDEAFKTITKFLMSQQKARVNGTQKFLGKQIMEFLSTLADSKQLQLEFSIKQIT AEEVHDHRELLSKQEPKLVKSRQQSFPSHAIDATLTMSIGLKEFPQFSQELDNSWFINHLMPDEVHLNPVRSKEKYNKP NISSTPLFKDSLYAERFIPVWVKGETFAIGFSEKDLFEIKPSNKEKLFTLLKTYSTKNPGESLQELQAKSKAKWLYFPINK TLALEFLHHYFHKEIVTPDDTTVCHFINSLRYYTKKESITVKILKEPMPVLSVKFESSKKNVLGSFKHTIALPATKDWER LFNHPNFLALKANPAPNPKEFNEFIRKYFLSDNNPNSDIPNNGHNIKPQKHKAVRKVFSLPVIPGNAGTMMRIRRKDNK GQPLYQLQTIDDTPSMGIQINEDRLVKQEVLMDAYKTRNLSTIDGINNSEGQAYATFDNWLTLPVSTFKPEIIKLEMKP HSKTRRYIRITQSLADFIKTIDEALMIKPSDSIDDPLNMPNEIVCKNKLFGNELKPRDGKMKIVSTGKIVTYEFESDSTPQ WIQTLYVTQLKKQP
Mammalian Codon Optimized Sequences (AgeI-HA-NLS-Cas9-NLS-taa-EcoRI)	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CATGGAGTCCTCTCAGATCCTGTCACCTATCGGCATCGACCTGGGCGGAAAGTTTACTGGAGTCTGCCTGTCACA CCTGGAAGCATTTGCTGAGCTGCCTAACCACGCAAATACAAAGTACTCAGTGATCCTGATTGATCATAACAATTT CCAGCTGAGCCAGGCACAGCGGAGAGCCACTAGACACAGGGTGCGCAACAAGAAAAGAAATCAGTTCGTGAAG AGGGTCGCCCTGCAGCTGTTTCAGCACATCCTGTCCCGGGACCTGAACGCCAAGGAGGAAACCGCTCTGTGCCAT TACCTGAACAATAGAGGCTACACCTATGTGGACACAGATCTGGACGAGTATATCAAAGATGAAACCACAATTAA CCTGCTGAAGGAGCTGCTGCCAAGCGAGTCCGAACATAATTTCATTGACTGGTTTCTGCAGAAGATGCAGAGCTC CGAGTTCCGGAAGATCCTGGTGAGCAAAGTCGAGGAAAAGAAAGACGATAAAGAACTGAAGAACGCCGTGAAG AACATCAAGAACTTCATCACCGGATTCGAGAAAAATAGCGTGGAAGGCCACCGACATCGGAAGGTCTACTTTGA GAACATCAAGTCCGATATTACAAAAGACAATCAGCTGGATTCTATCAAGAAAAAGATTCCTAGCGTGTGTCTGTC

FIG. 13G

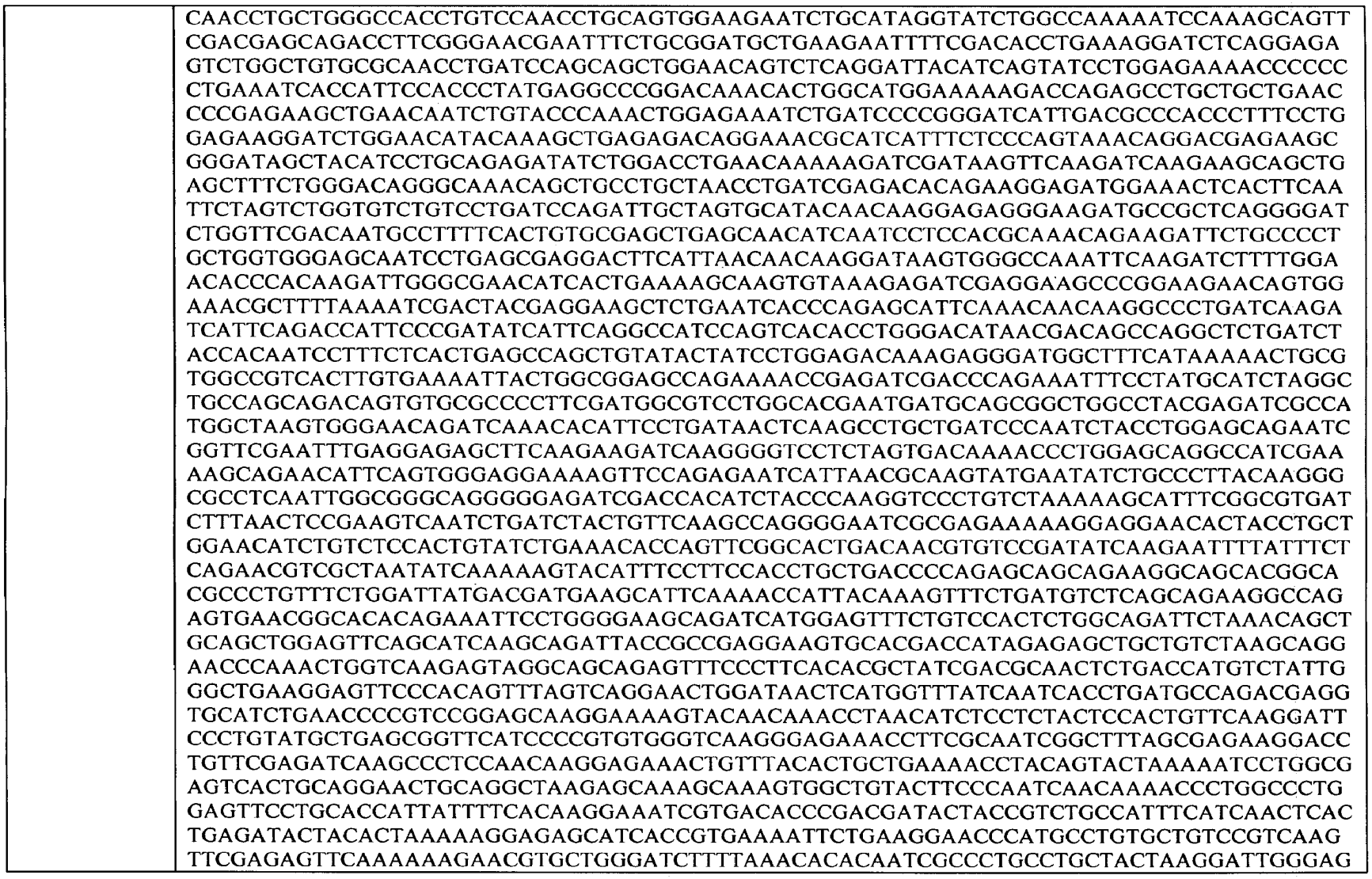

CAACCTGCTGGGCCACCTGTCCAACCTGCAGTGGAAGAATCTGCATAGGTATCTGGCCAAAAATCCAAAGCAGTT
CGACGAGCAGACCTTCGGGAACGAATTTCTGCGGATGCTGAAGAATTTTCGACACCTGAAAGGATCTCAGGAGA
GTCTGGCTGTGCGCAACCTGATCCAGCAGCTGGAACAGTCTCAGGATTACATCAGTATCCTGGAGAAAACCCCCC
CTGAAATCACCATTCCACCCTATGAGGCCCGGACAAACACTGGCATGGAAAAAGACCAGAGCCTGCTGCTGAAC
CCCGAGAAGCTGAACAATCTGTACCCAAACTGGAGAAATCTGATCCCCGGGATCATTGACGCCCACCCTTTCCTG
GAGAAGGATCTGGAACATACAAAGCTGAGAGACAGGAAACGCATCATTTCTCCCAGTAAACAGGACGAGAAGC
GGGATAGCTACATCCTGCAGAGATATCTGGACCTGAACAAAAAGATCGATAAGTTCAAGATCAAGAAGCAGCTG
AGCTTTCTGGGACAGGGCAAACAGCTGCCTGCTAACCTGATCGAGACACAGAAGGAGATGGAAACTCACTTCAA
TTCTAGTCTGGTGTCTGTCCTGATCCAGATTGCTAGTGCATACAACAAGGAGAGGGAAGATGCCGCTCAGGGGAT
CTGGTTCGACAATGCCTTTTCACTGTGCGAGCTGAGCAACATCAATCCTCCACGCAAACAGAAGATTCTGCCCCT
GCTGGTGGGAGCAATCCTGAGCGAGGACTTCATTAACAACAAGGATAAGTGGGCCAAATTCAAGATCTTTTGGA
ACACCCACAAGATTGGGCGAACATCACTGAAAAGCAAGTGTAAAGAGATCGAGGAAGCCCGGAAGAACAGTGG
AAACGCTTTTAAAATCGACTACGAGGAAGCTCTGAATCACCCAGAGCATTCAAACAACAAGGCCCTGATCAAGA
TCATTCAGACCATTCCCGATATCATTCAGGCCATCCAGTCACACCTGGGACATAACGACAGCCAGGCTCTGATCT
ACCACAATCCTTTCTCACTGAGCCAGCTGTATACTATCCTGGAGACAAAGAGGGATGGCTTTCATAAAAACTGCG
TGGCCGTCACTTGTGAAAATTACTGGCGGAGCCAGAAAACCGAGATCGACCCAGAAATTTCCTATGCATCTAGGC
TGCCAGCAGACAGTGTGCGCCCCTTCGATGGCGTCCTGGCACGAATGATGCAGCGGCTGGCCTACGAGATCGCCA
TGGCTAAGTGGGAACAGATCAAACACATTCCTGATAACTCAAGCCTGCTGATCCCAATCTACCTGGAGCAGAATC
GGTTCGAATTTGAGGAGAGCTTCAAGAAGATCAAGGGGTCCTCTAGTGACAAAACCCTGGAGCAGGCCATCGAA
AAGCAGAACATTCAGTGGGAGGAAAAGTTCCAGAGAATCATTAACGCAAGTATGAATATCTGCCCTTACAAGGG
CGCCTCAATTGGCGGGCAGGGGGAGATCGACCACATCTACCCAAGGTCCCTGTCTAAAAAGCATTTCGGCGTGAT
CTTTAACTCCGAAGTCAATCTGATCTACTGTTCAAGCCAGGGGAATCGCGAGAAAAAGGAGGAACACTACCTGCT
GGAACATCTGTCTCCACTGTATCTGAAACACCAGTTCGGCACTGACAACGTGTCCGATATCAAGAATTTTATTTCT
CAGAACGTCGCTAATATCAAAAAGTACATTTCCTTCCACCTGCTGACCCCAGAGCAGCAGAAGGCAGCACGGCA
CGCCCTGTTTCTGGATTATGACGATGAAGCATTCAAAACCATTACAAAGTTTCTGATGTCTCAGCAGAAGGCCAG
AGTGAACGGCACACAGAAATTCCTGGGGAAGCAGATCATGGAGTTTCTGTCCACTCTGGCAGATTCTAAACAGCT
GCAGCTGGAGTTCAGCATCAAGCAGATTACCGCCGAGGAAGTGCACGACCATAGAGAGCTGCTGTCTAAGCAGG
AACCCAAACTGGTCAAGAGTAGGCAGCAGAGTTTCCCTTCACACGCTATCGACGCAACTCTGACCATGTCTATTG
GGCTGAAGGAGTTCCCACAGTTTAGTCAGGAACTGGATAACTCATGGTTTATCAATCACCTGATGCCAGACGAGG
TGCATCTGAACCCCGTCCGGAGCAAGGAAAAGTACAACAAACCTAACATCTCCTCTACTCCACTGTTCAAGGATT
CCCTGTATGCTGAGCGGTTCATCCCCGTGTGGGTCAAGGGAGAAACCTTCGCAATCGGCTTTAGCGAGAAGGACC
TGTTCGAGATCAAGCCCTCCAACAAGGAGAAACTGTTTACACTGCTGAAAACCTACAGTACTAAAAATCCTGGCG
AGTCACTGCAGGAACTGCAGGCTAAGAGCAAAGCAAAGTGGCTGTACTTCCCAATCAACAAAACCCTGGCCCTG
GAGTTCCTGCACCATTATTTTCACAAGGAAATCGTGACACCCGACGATACTACCGTCTGCCATTTCATCAACTCAC
TGAGATACTACACTAAAAAGGAGAGCATCACCGTGAAAATTCTGAAGGAACCCATGCCTGTGCTGTCCGTCAAG
TTCGAGAGTTCAAAAAAGAACGTGCTGGGATCTTTTAAACACACAATCGCCCTGCCTGCTACTAAGGATTGGGAG

FIG. 13H

	AGGCTGTTCAACCATCCAAATTTTCTGGCACTGAAGGCCAACCCAGCTCCCAATCCTAAAGAGTTCAATGAGTTC ATCCGCAAGTACTTCCTGAGCGACAACAATCCCAACTCCGATATCCCTAACAATGGCCACAATATCAAGCCCCAG AAACATAAGGCCGTGCGAAAGGTCTTTAGCCTGCCAGTGATCCCCGGGAACGCTGGAACCATGATGCGCATTAG GCGCAAAGACAATAAGGGACAGCCACTGTATCAGCTGCAGACAATCGACGATACTCCCAGCATGGGCATCCAGA TTAACGAGGATCGCCTGGTGAAACAGGAAGTCCTGATGGACGCCTACAAGACACGAAATCTGAGCACTATCGAT GGGATTAACAATTCCGAGGGACAGGCATATGCCACATTCGACAACTGGCTGACCCTGCCCGTGAGCACCTTCAAG CCTGAGATCATCAAGCTGGAAATGAAGCCTCACTCTAAAACCCGACGGTACATCAGAATTACACAGAGTCTGGC CGACTTCATCAAAACTATTGATGAGGCTCTGATGATCAAGCCCAGTGACTCAATTGACGATCCTCTGAACATGCC AAATGAGATCGTGTGTAAAAACAAGCTGTTCGGGAATGAACTGAAGCCTAGGGATGGAAAAATGAAGATCGTGA GCACTGGCAAGATTGTCACCTACGAGTTTGAAAGCGACTCCACCCCCCAGTGGATCCAGACCCTGTATGTGACAC AGCTGAAAAAGCAGCCTTAAGAATTC
Length of Protein (without NLS and HA tag)	1372
PAM	**NGG**
Spacer Length	34-36
Species: Treponema_denticola	
Sequencing Barcode	TCGCCTTG
ID	4
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNGUUUGAGAGUUGGAAACAACGAGUUCAAAUAAGAAUUCAUCAAAAUCGUC CCUUUUGGGACCGCUCAUUGUGGAGCAUCAAGGCUUAACAUGGUUAAGCCUUUUUUU**
Cas9 Protein Sequence (wildtype, without NLS added)	MKKEIKDYFLGLDVGTGSVGWAVTDTDYKLLKANRKDLWGMRCFETAETAEVRRLHRGARRRIERRKKRIKLLQEL FSQEIAKTDEGFFQRMKESPFYAEDKTILQENTLFNDKDFADKTYHKAYPTINHLIKAWIENKVKPDPRLLYLACHNIIK KRGHFLFEGDFDSENQFDTSIQALFEYLREDMEVDIDADSQKVKEILKDSSLKNSEKQSRLNKILGLKPSDKQKKAITN LISGNKINFADLYDNPDLKDAEKNSISFSKDDFDALSDDLASILGDSFELLLKAKAVYNCSVLSKVIGDEQYLSFAKVKI YEKHKTDLTKLKNVIKKHFPKDYKKVFGYNKNEKNNNNYSGYVGVCKTKSKKLIINNSVNQEDFYKFLKTILSAKSEI KEVNDILTEIETGTFLPKQISKSNAEIPYQLRKMELEKILSNAEKHFSFLKQKDEKGLSHSEKIIMLLTFKIPYYIGPINDN

FIG. 13I

	HKKFFPDRCWVVKKEKSPSGKTTPWNFFDHIDKEKTAEAFITSRTNFCTYLVGESVLPKSSLLYSEYTVLNEINNLQIII DGKNICDIKLKQKIYEDLFKKYKKITQKQISTFIKHEGICNKTDEVIILGIDKECTSSLKSYIELKNIFGKQVDEISTKNML EEIIRWATIYDEGEGKTILKTKIKAEYGKYCSDEQIKKILNLKFSGWGRLSRKFLETVTSEMPGFSEPVNIITAMRETQNN LMELLSSEFTFTENIKKINSGFEDAEKQFSYDGLVKPLFLSPSVKKMLWQTLKLVKEISHITQAPPKKIFIEMAKGAELEP ARTKTRLKILQDLYNNCKNDADAFSSEIKDLSGKIENEDNLRLRSDKLYLYYTQLGKCMYCGKPIEIGHVFDTSNYDID HIYPQSKIKDDSISNRVLVCSSCNKNKEDKYPLKSEIQSKQRGFWNFLQRNNFISLEKLNRLTRATPISDDETAKFIARQL VETRQATKVAAKVLEKMFPETKIVYSKAETVSMFRNKFDIVKCREINDFHHAHDAYLNIVVGNVYNTKFTNNPWNFI KEKRDNPKIADTYNYYKVFDYDVKRNNITAWEKGKTIITVKDMLKRNTPIYTRQAACKKGELFNQTIMKKGLGQHPL KKEGPFSNISKYGGYNKVSAAYYTLIEYEEKGNKIRSLETIPLYLVKDIQKDQDVLKSYLTDLLGKKEFKILVPKIKINSL LKINGFPCHITGKTNDSFLLRPAVQFCCSNNEVLYFKKIIRFSEIRSQREKIGKTISPYEDLSFRSYIKENLWKKTKNDEIG EKEFYDLLQKKNLEIYDMLLTKHKDTIYKKRPNSATIDILVKGKEKFKSLIIENQFEVILEILKLFSATRNVSDLQHIGGS KYSGVAKIGNKISSLDNCILIYQSITGIFEKRIDLLKV
Mammalian Codon Optimized Sequences (AgeI-HA-NLS-Cas9-NLS-taa-EcoRI)	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CATGAAAAAAGAAATCAAAGACTACTTCCTGGGGCTGGATGTGGGGACTGGGAGCGTGGGGTGGGCTGTGACCG ATACTGACTACAAACTGCTGAAGGCTAACCGAAAAGACCTGTGGGGCATGAGATGCTTCGAGACAGCCGAAACT GCTGAGGTGCGGAGACTGCACAGGGGAGCCAGGCGCCGAATCGAGCGGAGAAAGAAACGCATTAAGCTGCTGC AGGAGCTGTTCTCTCAGGAAATCGCCAAAACCGATGAGGGCTTCTTTCAGAGAATGAAGGAAAGCCCCTTTTACG CTGAGGACAAAACAATCCTGCAGGAAAACACTCTGTTCAATGACAAGGATTTTGCTGATAAGACTTACCACAAA GCATATCCTACCATTAATCATCTGATCAAGGCTTGGATTGAGAACAAGGTGAAACCAGACCCCCGACTGCTGTAC CTGGCATGTCACAACATCATTAAGAAAAGGGGACATTTCCTGTTTGAAGGCGACTTCGATTCAGAGAATCAGTTT GATACCAGCATCCAGGCACTGTTCGAGTATCTGCGCGAGGACATGGAGGTGGACATCGATGCCGACAGCCAGAA GGTCAAAGAGATTCTGAAGGATAGCTCCCTGAAGAACTCTGAAAAACAGAGTCGGCTGAATAAGATCCTGGGGC TGAAGCCTTCCGACAAACAGAAGAAAGCCATCACAAACCTGATTTCTGGAAACAAGATCAATTTCGCCGATCTGT ACGACAATCCAGATCTGAAGGACGCTGAGAAAAACTCAATCAGCTTCTCCAAGGACGATTTTGATGCACTGAGT GACGATCTGGCCTCAATTCTGGGCGACAGCTTTGAACTGCTGCTGAAGGCCAAAGCTGTCTATAACTGCTCTGTG CTGAGTAAGGTCATCGGGGACGAGCAGTACCTGAGCTTCGCCAAGGTGAAAATCTACGAAAAGCACAAAACCGA TCTGACAAAGCTGAAAAACGTGATCAAGAAACATTTCCCCAAGGACTACAAGAAGGTCTTTGGATACAACAAGA ACGAGAAAAACAACAACAATTACTCCGGCTATGTGGGAGTCTGTAAGACCAAGAGTAAGAAACTGATCATTAAC AACTCAGTCAACCAGGAAGATTTCTACAAGTTTCTGAAAACTATCCTGTCAGCCAAGAGCGAGATCAAGGAAGT GAATGACATCCTGACCGAGATTGAAACTGGCACCTTTCTGCCAAAGCAGATCTCTAAAAGTAACGCAGAGATTCC CTATCAGCTGAGGAAAATGGAGCTGGAAAAGATCCTGTCCAATGCCGAAAAGCACTTCTCTTTTCTGAAGCAGAA AGACGAAAAAGGACTGTCACATAGCGAGAAGATCATTATGCTGCTGACCTTCAAGATCCCTTACTATATTGGCCC AATCAACGATAATCACAAGAAATTCTTTCCCGACAGATGCTGGGTGGTCAAGAAAGAGAAATCCCCTTCTGGCA AGACCACACCATGGAACTTCTTTGATCATATCGACAAGGAAAAAACAGCAGAGGCCTTCATTACTTCTAGGACCA ATTTTTGCACTTACCTGGTGGGAGAGAGCGTCCTGCCTAAGTCTAGTCTGCTGTACTCCGAATATACCGTGCTGAA

FIG. 13J

	CGAGATCAACAATCTGCAGATCATTATCGATGGCAAGAATATTTGTGACATCAAGCTGAAACAGAAGATCTACG AGGACCTGTTCAAGAAGTACAAGAAAATTACCCAGAAGCAGATCAGCACCTTCATCAAGCACGAAGGCATCTGC AACAAAACCGATGAAGTGATCATCCTGGGGATTGACAAGGAATGTACATCAAGCCTGAAAAGCTACATCGAGCT GAAAAACATTTTCGGCAAGCAGGTGGATGAGATCTCCACTAAGAATATGCTGGAGGAAATTATCAGATGGGCTA CCATCTACGACGAGGGGGAAGGAAAGACCATCCTGAAAACAAAGATCAAGGCAGAATACGGAAAGTATTGCTC CGACGAGCAGATTAAGAAAATCCTGAACCTGAAGTTCTCCGGCTGGGGGCGACTGTCTCGGAAATTTCTGGAGA CAGTGACTAGTGAAATGCCCGGCTTCTCAGAACCTGTCAATATTATCACCGCCATGAGGGAGACACAGAACAATC TGATGGAGCTGCTGTCCTCTGAGTTCACCTTCACCGAGAACATTAAGAAAATCAATTCTGGATTCGAAGATGCCG AGAAGCAGTTTAGTTACGACGGCCTGGTGAAACCACTGTTTCTGAGTCCCTCAGTCAAGAAAATGCTGTGGCAGA CCCTGAAGCTGGTGAAAGAGATTAGCCATATCACACAGGCCCCCCCTAAGAAAATTTTCATCGAAATGGCAAAG GGGGCCGAGCTGGAACCTGCTCGGACTAAGACCAGACTGAAAATCCTGCAGGATCTGTATAACAATTGTAAGAA CGATGCTGACGCCTTCAGCTCAGAGATCAAAGACCTGAGCGGAAAGATTGAGAACGAAGATAATCTGAGGCTGC GCTCCGACAAGCTGTACCTGTACTATACTCAGCTGGGGAAATGCATGTATTGTGGAAAGCCAATTGAGATCGGCC ACGTGTTCGATACCTCAAACTACGATATTGACCATATCTATCCCCAGAGCAAGATCAAAGACGATAGCATTTCCA ATCGGGTGCTGGTCTGCAGCTCCTGTAACAAGAACAAGGAGGACAAGTACCCACTGAAATCAGAGATCCAGAGC AAGCAGCGCGGCTTCTGGAACTTTCTGCAGCGAAACAATTTCATTTCTCTGGAGAAGCTGAATAGACTGACAAGG GCCACTCCAATCAGTGACGATGAGACAGCCAAGTTTATTGCTAGGCAGCTGGTGGAAACTCGCCAGGCTACCAA GGTGGCCGCTAAAGTCCTGGAAAAGATGTTCCCCGAGACAAAAATCGTGTACAGCAAGGCCGAGACTGTCTCCA TGTTCCGGAACAAGTTTGATATCGTGAAGTGCAGAGAAATTAACGATTTTCACCATGCTCACGACGCATACCTGA ATATCGTGGTCGGCAACGTGTATAATACCAAGTTCACAAACAATCCTTGGAACTTTATCAAGGAGAAAAGAGAT AATCCAAAGATTGCTGACACCTACAACTACTATAAGGTGTTCGATTATGACGTCAAAAGGAACAATATCACAGCA TGGGAGAAGGGGAAAACTATTATCACCGTGAAAGACATGCTGAAGAGAAACACACCAATCTACACTAGGCAGGC AGCCTGTAAGAAAGGGGAGCTGTTCAATCAGACCATTATGAAGAAAGGACTGGGCCAGCACCCCCTGAAGAAAG AAGGACCTTTTTCCAATATCTCTAAATACGGCGGGTATAACAAGGTGAGCGCTGCATACTATACACTGATTGAGT ATGAGGAAAAGGGCAACAAAATCCGCTCCCTGGAAACTATTCCCCTGTACCTGGTGAAAGATATCCAGAAGGAT CAGGACGTCCTGAAGTCTTATCTGACAGACCTGCTGGGGAAGAAAGAGTTCAAGATCCTGGTGCCCAAGATCAA GATCAACAGCCTGCTGAAGATCAATGGGTTTCCTTGCCATATTACAGGAAAAACTAACGATAGTTTCCTGCTGCG CCCTGCCGTGCAGTTTTGCTGTTCAAACAATGAGGTCCTGTACTTCAAGAAAATTATCCGGTTTTCCGAAATCCGC TCTCAGCGAGAGAAGATCGGGAAAACAATTAGCCCATACGAGGACCTGAGCTTCCGGTCATATATCAAGGAGAA CCTGTGGAAGAAAACTAAGAACGATGAAATCGGAGAGAAGGAATTTTACGACCTGCTGCAGAAGAAAAACCTG GAGATCTATGATATGCTGCTGACTAAGCACAAAGACACCATCTACAAGAAACGCCCTAATTCTGCCACCATTGAT ATCCTGGTGAAGGGGAAAGAGAAGTTCAAAAGCCTGATTATCGAAAACCAGTTTGAAGTGATCCTGGAGATCCT GAAGCTGTTTTCTGCAACACGGAATGTCAGTGACCTGCAGCATATCGGAGGCAGCAAGTACTCCGGCGTGGCCA AAATCGGGAACAAGATCTCTAGTCTGGATAACTGTATCCTGATCTATCAGTCCATCACCGGCATCTTCGAGAAAC GGATCGACCTGCTGAAGGTGTAAGAATTC

FIG. 13K

Length of Protein (without NLS and HA tag)	1395
PAM	**NYAAAT**
Spacer Length	30
Species: Filifactor_alocis	
Sequencing Barcode	AAGACACT
ID	5
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNGUUUGAGAGUAGGAAACUACACGGUUCAAAUAAAGAAUUUUUCUAAUCGC CCAAUGGGCCCAUAUUGAUAUGGAUGAAACUCGCUUAGCGAGUUUUUUU**
Cas9 Protein Sequence (wildtype, without NLS added)	MTKEYYLGLDVGTNSVGWAVTDSQYNLCKFKKKDMWGIRLFESANTAKDRRLQRGNRRRLERKKQRIDLLQEIFSP EICKIDPTFFIRLNESRLHLEDKSNDFKYPLFIEKDYSDIEYYKEFPTIFHLRKHLIESEEKQDIRLIYLALHNIIKTRGHFLI DGDLQSAKQLRPILDTFLLSLQEEQNLSVSLSENQKDEYEEILKNRSIAKSEKVKKLKNLFEISDELEKEEKKAQSAVIE NFCKFIVGNKGDVCKFLRVSKEELEIDSFSFSEGKYEDDIVKNLEEKVPEKVYLFEQMKAMYDWNILVDILETEEYISF AKVKQYEKHKTNLRLLRDIILKYCTKDEYNRMFNDEKEAGSYTAYVGKLKKNNKKYWIEKKRNPEEFYKSLGKLLD KIEPLKEDLEVLTMMIEECKNHTLLPIQKNKDNGVIPHQVHEVELKKILENAKKYYSFLTETDKDGYSVVQKIESIFRFR IPYYVGPLSTRHQEKGSNVWMVRKPGREDRIYPWNMEEIIDFEKSNENFITRMTNKCTYLIGEDVLPKHSLLYSKYMV LNELNNVKVRGKKLPTSLKQKVFEDLFENKSKVTGKNLLEYLQIQDKDIQIDDLSGFDKDFKTSLKSYLDFKKQIFGEE IEKESIQNMIEDIIKWITIYGNDKEMLKRVIRANYSNQLTEEQMKKITGFQYSGWGNFSKMFLKGISGSDVSTGETFDIIT AMWETDNNLMQILSKKFTFMDNVEDFNSGKVGKIDKITYDSTVKEMFLSPENKRAVWQTIQVAEEIKKVMGCEPKKI FIEMARGGEKVKKRTKSRKAQLLELYAACEEDCRELIKEIEDRDERDFNSMKLFLYYTQFGKCMYSGDDIDINELIRGN SKWDRDHIYPQSKIKDDSIDNLVLVNKTYNAKKSNELLSEDIQKKMHSFWLSLLNKKLITKSKYDRLTRKGDFTDEEL SGFIARQLVETRQSTKAIADIFKQIYSSEVVYVKSSLVSDFRKKPLNYLKSRRVNDYHHAKDAYLNIVVGNVYNKKFTS NPIQWMKKNRDTNYSLNKVFEHDVVINGEVIWEKCTYHEDTNTYDGGTLDRIRKIVERDNILYTEYAYCEKGELFNA TIQNKNGNSTVSLKKGLDVKKYGGYFSANTSYFSLIEFEDKKGDRARHIIGVPIYIANMLEHSPSAFLEYCEQKGYQNV RILVEKIKKNSLLIINGYPLRIRGENEVDTSFKRAIQLKLDQKNYELVRNIEKFLEKYVEKKGNYPIDENRDHITHEKMN QLYEVLLSKMKKFNKKGMADPSDRIEKSKPKFIKLEDLIDKINVINKMLNLLRCDNDTKADLSLIELPKNAGSFVVKK NTIGKSKIILVNQSVTGLYENRREL

FIG. 13L

Mammalian Codon Optimized Sequences (AgeI-HA-NLS-Cas9-NLS-taa-EcoRI)	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CATGACCAAGGAGTATTACCTGGGGCTGGATGTGGGGACCAATTCCGTGGGATGGGCAGTGACCGATTCTCAGT ACAACCTGTGCAAGTTTAAGAAAAAGGATATGTGGGGCATCCGGCTGTTCGAAAGCGCCAACACAGCAAAGGAC CGGAGACTGCAGAGAGGGAATAGGCGCCGACTGGAGCGGAAAAAGCAGAGAATTGATCTGCTGCAGGAAATCT TCTCCCCAGAGATCTGCAAGATTGACCCCACTTTCTTTATCCGACTGAACGAATCCCGGCTGCACCTGGAGGACA AGTCTAACGATTTCAAATACCCACTGTTTATTGAGAAGGACTATTCTGATATCGAGTACTATAAAGAGTTCCCCA CCATTTTTCACCTGAGGAAGCATCTGATCGAGAGTGAGGAAAAACAGGATATCCGGCTGATCTACCTGGCCCTGC ACAACATCATTAAGACCCGAGGACATTTTCTGATTGACGGCGATCTGCAGAGCGCCAAGCAGCTGAGGCCCATCC TGGATACATTCCTGCTGTCCCTGCAGGAGGAACAGAACCTGTCAGTGAGCCTGTCCGAAAATCAGAAGGACGAG TATGAGGAAATTCTGAAAAACCGCAGCATCGCCAAGTCCGAAAAAGTGAAAAAGCTGAAGAATCTGTTTGAGAT CTCAGACGAGCTGGAAAAAGAGGAGAAGAAGGCCCAGAGCGCCGTGATCGAGAACTTCTGCAAGTTTATCGTGG GAAATAAGGGCGATGTCTGTAAATTCCTGCGGGTGTCTAAGGAGGAACTGGAGATTGACTCTTTCAGTTTTTCAG AGGGCAAGTACGAGGACGACATCGTGAAAAACCTGGAGGAAAAAGTGCCTGAAAAGGTCTACCTGTTTGAGCAG ATGAAGGCAATGTATGATTGGAATATTCTGGTCGACATCCTGGAAACCGAGGAATACATCAGCTTCGCCAAAGTG AAGCAGTATGAGAAACACAAGACTAACCTGCGGCTGCTGAGAGACATCATTCTGAAATACTGCACCAAGGATGA GTATAATCGGATGTTTAACGACGAGAAGGAAGCTGGCAGCTACACCGCATATGTGGGGAAACTGAAAAAGAACA ACAAGAAGTACTGGATCGAGAAAAAGAGAAATCCCGAGGAGTTCTACAAATCCCTGGGCAAGCTGCTGGATAAA ATTGAGCCTCTGAAGGAGGACCTGGAAGTGCTGACTATGATGATCGAGGAGTGTAAGAACCACACCCTGCTGCC AATTCAGAAAAATAAGGACAACGGCGTGATCCCCCACCAGGTGCATGAGGTCGAACTGAAAAAGATCCTGGAAA ATGCCAAAAAGTACTATTCCTTCCTGACCGAGACAGACAAGGATGGGTACTCAGTGGTCCAGAAAATCGAGAGC ATTTTCAGGTTTCGCATCCCCTACTATGTGGGGCCTCTGAGTACCCGGCACCAGGAAAAGGGATCAAACGTGTGG ATGGTCAGAAAACCTGGCAGGGAGGATCGCATCTACCCATGGAATATGGAGGAAATCATTGACTTTGAGAAGAG CAACGAAAATTTCATTACACGGATGACTAACAAATGTACATATCTGATCGGGGAAGATGTCCTGCCCAAGCATTC TCTGCTGTACAGTAAATATATGGTGCTGAATGAGCTGAACAATGTGAAGGTCAGAGGAAAAAAGCTGCCTACAT CTCTGAAACAGAAGGTGTTCGAGGACCTGTTTGAAAACAAATCCAAAGTGACTGGAAAGAATCTGCTGGAGTAC CTGCAGATCCAGGACAAAGATATCCAGATTGACGATCTGTCTGGCTTCGACAAGGACTTCAAGACCAGCCTGAA GAGCTATCTGGACTTCAAAAAGCAGATTTTTGGGGAGGAAATCGAGAAGGAAAGCATTCAGAACATGATCGAAG ATATCATTAAGTGGATCACCATCTACGGCAATGACAAGGAGATGCTGAAACGAGTGATTCGGGCTAATTATAGC AACCAGCTGACAGAGGAACAGATGAAAAAGATCACTGGATTTCAGTACAGTGGCTGGGGGAACTTCTCAAAGAT GTTTCTGAAAGGGATCAGCGGATCCGACGTGAGCACCGGCGAAACATTCGACATCATTACCGCAATGTGGGAGA CAGACAACAATCTGATGCAGATCCTGTCAAAAAAGTTCACCTTTATGGACAACGTCGAGGACTTCAACAGCGGC AAGGTCGGGAAAATCGACAAGATTACTTACGATAGCACCGTGAAGGAAATGTTCCTGTCCCCTGAGAACAAAAG GGCCGTCTGGCAGACCATTCAGGTGGCTGAGGAGATCAAGAAAGTGATGGGCTGCGAGCCAAAAAAGATCTTTA TTGAAATGGCACGGGGCGGGGAGAAGGTGAAAAAGAGGACAAAATCTCGCAAGGCCCAGCTGCTGGAGCTGTA CGCCGCTTGCGAGGAAGATTGTAGAGAACTGATCAAGGAGATTGAGGACCGGGACGAGAGGGACTTCAATAGCA

FIG. 13M

	TGAAGCTGTTTCTGTACTATACCCAGTTCGGGAAATGTATGTATTCCGGCGACGACATCGATATTAACGAGCTGA TTCGCGGCAATTCTAAGTGGGACCGAGATCACATCTACCCCCAGAGCAAAATTAAGGACGATTCCATCGATAACC TGGTGCTGGTCAATAAGACATATAATGCCAAAAAGTCCAATGAGCTGCTGTCTGAGGACATCCAGAAAAAGATG CATTCATTCTGGCTGAGCCTGCTGAACAAAAAGCTGATCACTAAAAGCAAGTACGACCGCCTGACTCGAAAGGG CGACTTTACCGATGAGGAACTGAGTGGGTTCATCGCTAGACAGCTGGTGGAAACAAGGCAGTCAACTAAGGCAA TCGCCGATATCTTCAAGCAGATCTACAGCTCCGAGGTGGTCTATGTGAAGAGCAGCCTGGTGAGCGACTTCAGGA AAAAGCCACTGAACTACCTGAAGTCTCGGAGAGTCAATGATTACCACCATGCAAAAGACGCCTATCTGAACATT GTGGTCGGGAACGTGTACAACAAAAAGTTTACCAGTAATCCCATCCAGTGGATGAAAAAGAATCGCGATACAAA CTATAGCCTGAACAAGGTGTTCGAACACGACGTGGTCATTAACGGAGAAGTGATCTGGGAAAAGTGCACATACC ATGAGGACACTAATACCTATGATGGAGGCACTCTGGACCGAATCCGGAAGATTGTGGAACGCGATAACATTCTG TACACCGAGTACGCTTATTGTGAGAAGGGCGAACTGTTTAATGCAACCATCCAGAACAAAAATGGAAACTCCAC AGTCTCTCTGAAAAAGGGCCTGGACGTGAAAAAGTACGGGGGATACTTCAGCGCCAACACAAGTTACTTCTCACT GATCGAGTTTGAGGACAAGAAGGGGGATAGAGCAAGGCACATCATTGGAGTGCCTATCTATATTGCAAACATGC TGGAGCATTCTCCAAGTGCCTTCCTGGAGTACTGCGAACAGAAGGGGTATCAGAATGTGCGGATTCTGGTCGAGA AAATCAAAAAGAACAGCCTGCTGATCATTAATGGATACCCTCTGCGCATTCGAGGCGAGAACGAAGTGGATACT TCCTTTAAGAGGGCCATCCAGCTGAAGCTGGACCAGAAAAACTATGAGCTGGTCCGCAATATCGAGAAGTTCCTG GAAAAATACGTGGAGAAAAAGGGAAACTATCCAATTGACGAGAATAGAGATCACATCACACATGAAAAGATGA ACCAGCTGTACGAGGTGCTGCTGTCCAAAATGAAAAAGTTCAACAAGAAGGGCATGGCCGACCCCTCTGATAGG ATCGAAAAGAGTAAGCCTAAATTCATCAAGCTGGAGGACCTGATCGATAAGATTAATGTGATCAACAAAATGCT GAACCTGCTGCGCTGTGACAATGATACTAAGGCCGACCTGTCTCTGATTGAGCTGCCCAAAAACGCTGGGAGTTT CGTGGTCAAAAAGAATACCATCGGAAAGTCAAAAATCATCCTGGTGAATCAGAGCGTGACTGGACTGTACGAGA ATAGACGGGAACTGTAAGAATTC
Length of Protein (without NLS and HA tag)	1365
PAM	**NNNAAGC**
Spacer Length	30-31
Species: Staphylococcus_pseudintermedius	
Sequencing	TCTCGGTC

FIG. 13N

Barcode	
ID	6
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNNGUUUUAGCACUAGAAAUAGUUAAGUUAAAACAAGCUUAAAGCGUCAAUGU AAUAUUUUAUUAACACCCUACUGUGUCAGUGGGGUUUUUUU**
Cas9 Protein Sequence (wildtype, without NLS added)	MGRKPYILSLDIGTGSVGYACMDKGFNVLKYHDKDALGVYLFDGALTAQERRQFRTSRRRKNRRIKRLGLLQELLAP LVQNPNFYQFQRQFAWKNDNMDFKNKSLSEVLSFLGYESKKYPTIYHLQEALLLKDEKFDPELIYMALYHLVKYRGH FLFDHLKIENLTNNDNMHDFVELIETYENLNNIKLNLDYEKTKVIYEILKDNEMTKNDRAKRVKNMEKKLEQFSIMLL GLKFNEGKLFNHADNAEELKGANQSHTFADNYEENLTPFLTVEQSEFIERANKIYLSLTLQDILKGKKSMAMSKVAAY DKFRNELKQVKDIVYKADSTRTQFKKIFVSSKKSLKQYDATPNDQTFSSLCLFDQYLIRPKKQYSLLIKELKKIIPQDSE LYFEAENDTLLKVLNTTDNASIPMQINLYEAETILRNQQKYHAEITDEMIEKVLSLIQFRIPYYVGPLVNDHTASKFGW MERKSNESIKPWNFDEVVDRSKSATQFIRRMTNKCSYLINEDVLPKNSLLYQEMEVLNELNATQIRLQTDPKNRKYRM MPQIKLFAVEHIFKKYKTVSHSKFLEIMLNSNHRENFMNHGEKLSIFGTQDDKKFASKLSSYQDMTKIFGDIEGKRAQI EEIIQWITIFEDKKILVQKLKECYPELTSKQINQLKKLNYSGWGRLSEKLLTHAYQGHSIIELLRHSDENFMEILTNDVY GFQNFIKEENQVQSNKIQHQDIANLTTSPALKKGIWSTIKLVRELTSIFGEPEKIIMEFATEDQQKGKKQKSRKQLWDD NIKKNKLKSVDEYKYIIDVANKLNNEQLQQEKLWLYLSQNGKCMYSGQSIDLDALLSPNATKHYEVDHIFPRSFIKDD SIDNKVLVIKKMNQTKGDQVPLQFIQQPYERIAYWKSLNKAGLISDSKLHKLMKPEFTAMDKEGFIQRQLVETRQISV HVRDFLKEEYPNTKVIPMKAKMVSEFRKKFDIPKIRQMNDAHHAIDAYLNGVVYHGAQLAYPNVDLFDFNFKWEKV REKWKALGEFNTKQKSRELFFFKKLEKMEVSQGERLISKIKLDMNHFKINYSRKLANIPQQFYNQTAVSPKTAELKYE SNKSNEVVYKGLTPYQTYVVAIKSVNKKGKEKMEYQMIDHYVFDFYKFQNGNEKELALYLAQRENKDEVLDAQIVY SLNKGDLLYINNHPCYFVSRKEVINAKQFELTVEQQLSLYNVMNNKETNVEKLLIEYDFIAEKVINEYHHYLNSKLKE KRVRTFFSESNQTHEDFIKALDELFKVVTASATRSDKIGSRKNSMTHRAFLGKGKDVKIAYTSISGLKTTKPKSLFKLA ESRNEL
Mammalian Codon Optimized Sequences (AgeI-HA-NLS-Cas9-NLS-taa-EcoRI)	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CATGGGGAGGAAACCTTACATTCTGTCTCTGGATATTGGAACTGGGTCCGTCGGCTACGCTTGCATGGATAAAGG ATTCAACGTGCTGAAGTACCACGACAAAGATGCCCTGGGAGTGTATCTGTTCGACGGCGCTCTGACTGCACAGGA GCGGAGACAGTTTAGGACCTCCAGGCGCCGAAAGAACCGGAGAATCAAACGCCTGGGCCTGCTGCAGGAACTGC TGGCACCCCTGGTGCAGAACCCTAATTTCTACCAGTTTCAGCGGCAGTTCGCCTGGAAGAACGACAATATGGATT TTAAGAACAAGAGCCTGTCTGAGGTGCTGAGCTTCCTGGGATATGAATCCAAGAAATACCCTACCATCTACCACC TGCAGGAGGCTCTGCTGCTGAAAGACGAGAAGTTTGATCCAGAACTGATCTACATGGCACTGTATCATCTGGTGA AATACAGAGGCCACTTTCTGTTCGATCATCTGAAGATCGAGAACCTGACTAACAATGACAATATGCACGATTTCG TGGAGCTGATTGAAACCTATGAGAACCTGAACAATATCAAGCTGAATCTGGACTACGAGAAAACCAAAGTGATC TATGAGATTCTGAAAGACAACGAAATGACTAAGAATGATAGAGCCAAAAGGGTCAAGAACATGGAGAAGAAAC TGGAACAGTTCTCTATCATGCTGCTGGGGCTGAAGTTCAATGAGGGAAAACTGTTTAACCACGCCGATAATGCTG AGGAACTGAAGGGGGCTAACCAGAGCCATACATTTGCAGACAACTACGAGGAAAATCTGACTCCCTTCCTGACC

FIG. 13O

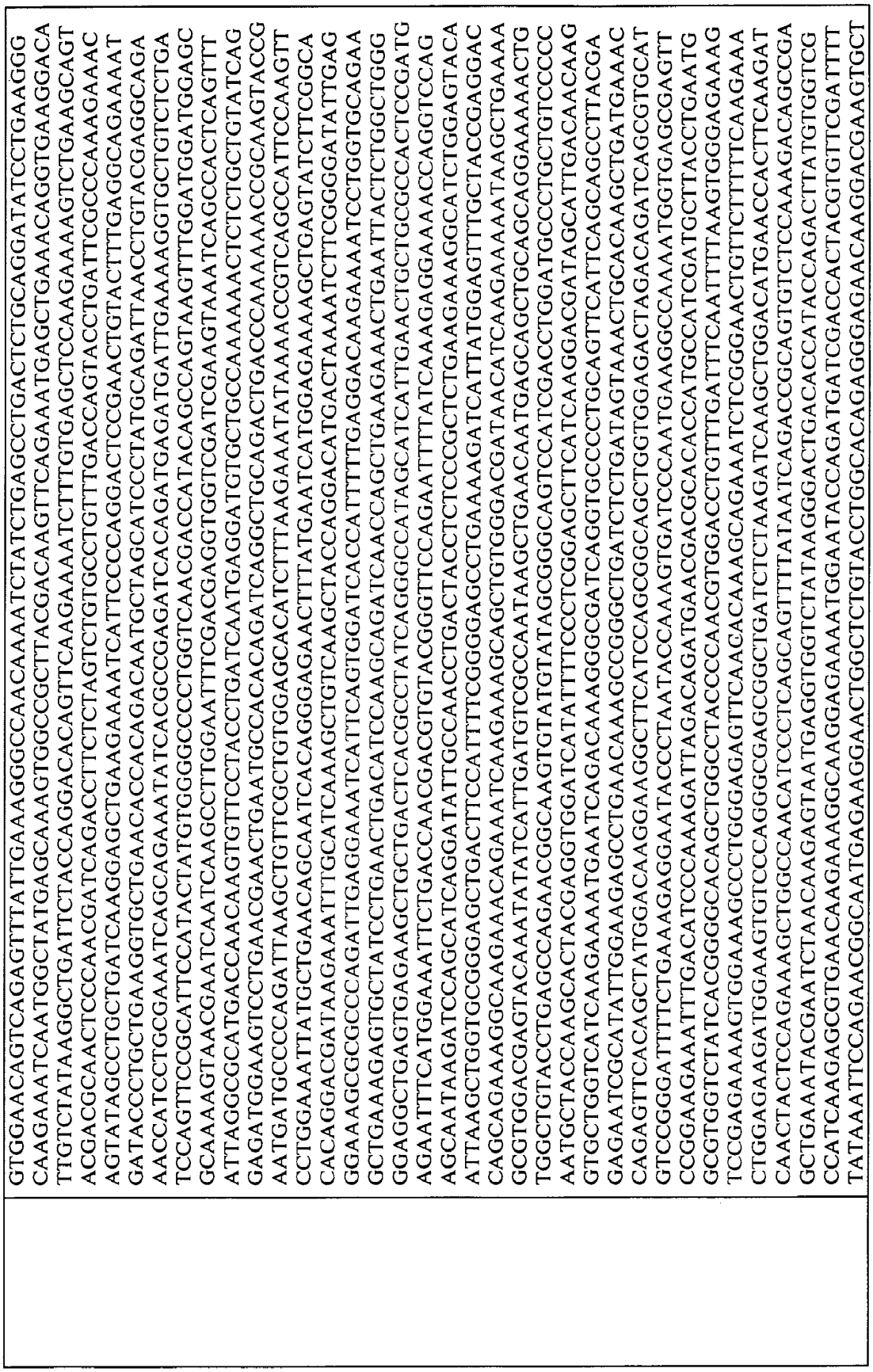

GTGGAACAGTCAGAGTTTATTGAAAGGGCCAACAAAATCTATCTGAGCCTGACTCTGCAGGATATCCTGAAGGG
CAAGAAATCAATGGCTATGAGCAAAGTGGCCGCTTACGACAAGTTCAGAAATGAGCTGAAACAGGTGAAGGACA
TTGTCTATAAGGCTGATTCTACCAGGACACAGTTCAAGAAAATCTTTGTGAGCTCCAAGAAAAGTCTGAAGCAGT
ACGACGCAACTCCCAACGATCAGACCTTCTCTAGTCTGTGCCTGTTTGACCAGTACCTGATTCGCCCAAAGAAAC
AGTATAGCCTGCTGATCAAGGAGCTGAAGAAAATCATTCCCCAGGACTCCGAACTGTACTTTGAGGCAGAAAAT
GATACCCTGCTGAAGGTGCTGAACACCACAGACAATGCTAGCATCCCTATGCAGATTAACCTGTACGAGGCAGA
AACCATCCTGCGAAATCAGCAGAAATATCACGCCGAGATCACAGATGAGATGATTGAAAAGGTGCTGTCTCTGA
TCCAGTTCCGCATTCCATACTATGTGGGGCCCCTGGTCAACGACCATACAGCCAGTAAGTTTGGATGGAGGAGC
GCAAAAGTAACGAATCAATCAAGCCTTGGAATTTCGACGAGGTGGTCGATCGAAGTAAATCAGCCACTCAGTTT
ATTAGGCGCATGACCAACAAGTGTTCCTACCTGATCAATGAGGATGTGCTGCCAAAAAACTCTCTGCTGTATCAG
GAGATGGAAGTCCTGAACGAACTGAATGCCACACAGATCAGGCTGCAGACTGACCCAAAAAACCGCAAGTACCG
AATGATGCCCCCAGATTAAGCTGTTCGCTGTGGAGCACATCTTTTAAGAAATATAAAACCGTCAGCCATTCCAAGTT
CCTGGAAATTATATGCTGAACAGCAATCACAGGGAGAAGAACTTTATGAATCATGGAGAAAAGCTGAGTATCTTCGGCA
CACAGGACGATAAGAAATTTGCATCAAAGCTGTCAAGCTACCAGGACATGACTAAAATCTTCGGGGATATTGAG
GGAAAGCGCGCCCCAGATTGAGGAAATCATTCAGTGGATCACCATTTTTGAGGACAAGAAAATCCTGGTGCAGAA
GCTGAAAGAGTGCTATCCTGAACTGACATCCAAGCAGATCAACCAGCTGAAGAAACTGAATTACTCTGGCTGGG
GGAGGCTGAGTGAGAAGCTGCTGACTCACGCCTATCAGGGCCATAGCATCATTGAACTGCTGCGCCACTCCGATG
AGAATTTCATGGAAATTCTGACCAACGACGTGTACGGGTTCCAGAATTTTTATCAAAGAGGAAAAACCAGGTCCAG
AGCAATAAGATCCAGCATCAGGATATTGCCAACCTGACTACCTCTCCCGCTCTGAAGAAAGGCATCTGGAGTACA
ATTAAGCTGGTGCGGGAGCTGACTTCCATTTTCGGGGAGCCTGAAAAGATCATTATGGAGTTTGCTACCGAGGAC
CAGCAGAAAGGCAAGAAACAGAAATCAAGAAAGCAGCTGTGGGACGATAACATCAAGAAAAATAAGCTGAAAA
GCGTGGACGAGTACAAATATATCATTGATGTCGCCAATAAGCTGAACAATGAGCAGCTGCAGCAGGAAAAACTG
TGGCTGTACCTGAGCCAGAACGGCAAGTGTATGTATAGCGGGGCAGTCCATCGACCTGGATGCCCTGCTGTCCCCC
AATGCTACCAAGCACTACGAGGTGGATCATATTTTCCCTCGGAGCTTCATCAAGGACGATAGCATTGACAACAAG
GTGCTGGTCATCAAGAAAATGAATCAGACAAAGGGCGATCAGGTGCCCCTGCAGTTCATTCAGCAGCCTTACGA
GAGAATCGCATATTGGAAGAGCCTGAACAAAGCCGGGCTGATCTCTGATAGTAAACTGCACACAAGCTGATGAAAC
CAGAGTTCACAGCTATGGACAAGGAAGGCTTCATCCAGCGGCAGCTGGTGGAGACTAGACAGATCAGCGTGCAT
GTCCGGGATTTTCTGAAAGAGGAATACCCTAATACCAAAGTGATCCCAAATGGTGAGCGAGTT
CCGGAAGAAATTTGACATCCCAAAGATTAGACAGATGAACGACGCACACCATGCCATCGATGCTTACCTGAATG
GCGTGGTCTATCACGGGGCACAGCTGGCCTACCCCAACGTGGACCTGTTTGATTTCAATTTTAAGTGGGAGAAAG
TCCGAGAAAAGTGGAAAGCCCTGGGAGAGTTCAACACAAAGCAGAAATCTCGGGAACTGTTCTTTTTCAAGAAA
CTGGAGAAGATGGAAGTGTCCCAGGGCGAGCGGCTGATCTCTAAGATCAAGCTGGACATGAACCACTTCAAGAT
CAACTACTCCAGAAAGCTGGCCAACATCCCTCAGCAGTTTTATAATCAGACCGCAGTGTCTCCAAAGACAGCCGA
GCTGAAATACGAATCTAACAAGAGTAATGAGGTGGTCTATAAGGGACTGACACCATACCAGACTTATGTGGTCG
CCATCAAGAGCGTGAACAAGAAAGGCAAGGAGAAAATGGAATACCCAGATGATCGATTTTCGATTTT
TATAAATTCCAGAACGGCAATGAGAAGGAACTGGCTCTGTACCTGGCACAGAGGGAGAACAAGGACGAAGTGCT

FIG. 13P

	GGATGCTCAGATTGTCTATAGTCTGAATAAGGGGGATCTGCTGTACATCAACAATCATCCCTGCTATTTCGTGTCA CGCAAAGAGGTCATCAACGCAAAGCAGTTTGAGCTGACCGTGGAACAGCAGCTGTCTCTGTACAACGTGATGAA CAACAAGGAGACAAATGTCGAAAAGCTGCTGATCGAGTATGACTTCATTGCCGAGAAAGTGATCAACGAATACC ACCATTATCTGAATAGCAAGCTGAAAGAAAAGCGAGTCCGGACCTTTTTCTCAGAGAGCAACCAGACACACGAG GACTTCATCAAGGCCCTGGACGAGCTGTTTAAGGTGGTCACCGCATCCGCCACAAGGTCTGATAAAATCGGGAGT CGCAAGAACAGCATGACTCATCGAGCCTTCCTGGGAAAAGGCAAGGACGTGAAGATTGCTTACACCTCCATCTCT GGACTGAAAACAACTAAACCTAAGAGTCTGTTTAAGCTGGCCGAGTCAAGAAACGAACTGTAAGAATTC
Length of Protein (without NLS and HA tag)	1334
PAM	**NGG**
Spacer Length	29
Species: Lactobacillus_johnsonii	
Sequencing Barcode	AATGTTCT
ID	7
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNGUUUUAGAUGGUGAAAACCAGAUUUAAAAUCAAGCAAUGCAUCUUUUGAU GCAAAGUUUCAAUAUUUGUCCCACGUUAUCGAGGGACUUUUUUU**
Cas9 Protein Sequence (wildtype, without NLS added)	MTKIKDDYIVGLDIGTDSCGWVAMNSNNDILKLQGKTAIGSRLFEGGKSAAERRLFRTTHRRIKRRRWRLKLLEEFFD PYMAEVDPYFFARLKESGLSPLDKRKTVSSIVFPTSAEDKKFYDDYPTIYHLRYKLMTEDEKFDLREVYLAIHHIIKYR GNFLYNTSVKDFKASKIDVKSSIEKLNELYENLGLDLNVEFNISNTAEIEKVLKDKQIFKRDKVKKIAELFAIKTDNKEQ SKRIKDISKQVANAVLGYKTRFDTIALKEISKDELSDWNFKLSDIDADSKFEALMGNLDENEQAILLTIKELFNEVTLNG IVEDGNTLSESMINKYNDHRDDLKLLKEVIENHIDRKKAKELALAYDLYVNNRHGQLLQAKKKLGKIKPRSKEDFYK VVNKNLDDSRASKEIKKKIELDSFMPKQRTNANGVIPYQLQQLELDKIIENQSKYYPFLKEINPVSSHLKEAPYKLDELI RFRVPYYVGPLISPNESTKDIQTKKNQNFAWMIRKEEGRITPWNFDQKVDRIESANKFIKRMTTKDTYLFGEDVLPANS LLYQKFTVLNELNNIRINGKRISVDLKQEIYENLFKKHTTVTVKKLENYLKENHNLVKVEIKGLADEKKFNSGLTTYN RFKNLNIFDNQIDDLKYRNDFEKIIEWSTIFEDKSIYKEKLRSIDWLNEKQINALSNIRLQGWGRLSKKLLAQLHDHNGQ TIIEQLWDSQNNFMQIVTQADFKDAIAKANQNLLVATSVEDILNNAYTSPANKKAIRQVIKVVDDIVKAASGKVPKQI

FIG. 13Q

	AIEFTRDADENPKRSQTRGSKLQKVYKDLSTELASKTIAEELNEAIKDKKLVQDKYYLYFMQLGRDAYTGEPINIDEIQ KYDIDHILPQSFIKDDALDNRVLVSRAVNNGKSDNVPVKLFGNEMAANLGMTIRKMWEEWKNIGLISKTKYNNLLTD PDHINKYKSAGFIRRQLVETSQIIKLVSTILQSRYPNTEIITVKAKYNHYLREKFDLYKSREVNDYHHAIDAYLSAICGNL LYQNYPNLRPFFVYGQYKKFSSDPDKEKAIFNKTRKFSFISQLLKNKSENSKEIAKKLKRAYQFKYMLVSRETETRDQE MFKMTVYPRFSHDTVKAPRNLIPKKMGMSPDIYGGYTNNSDAYMVIVRIDKKKGTEYKILGIPTRELVNLKKAEKED HYKSYLKEILTPRILYNKNGKRDKKITSFEIVKSKIPYKQVIQDGDKKFMLGSSTYVYNAKQLTLSTESMKAITNNFDK DSDENDALIKAYDEILDKVDKYLPLFDINKFREKLHSGREKFIKLSLEDKKDTILKVLEGLHDNAVMTKIPTIGLSTPLG FMQFPNGVILSENAKLIYQSPTGLFKKSVKISDL
Mammalian Codon Optimized Sequences (AgeI-HA-NLS-Cas9-NLS-taa-EcoRI)	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CATGACAAAAATCAAAGACGACTACATCGTGGGACTGGACATCGGCACAGACTCCTGCGGGTGGGTGGCTATGA ACAGCAATAATGACATTCTGAAACTGCAGGGCAAGACCGCAATCGGGTCACGCCTGTTCGAGGGAGGGAAGAGC GCAGCTGAACGGAGACTGTTTCGCACCACACACAGGCGCATCAAACGACGGAGATGGCGACTGAAGCTGCTGGA GGAGTTCTTCGACCCCTACATGGCAGAGGTGGATCCTTATTTCTTTGCCCGGCTGAAGGAATCTGGCCTGAGTCC ACTGGACAAAAGAAAGACCGTGAGCTCCATTGTGTTCCCCACATCCGCCGAGGATAAGAAGTTCTACGACGATT ACCCTACAATCTACCATCTGAGGTATAAACTGATGACTGAGGACGAAAAGTTCGATCTGCGCGAAGTGTACCTGG CTATCCACCATATCATTAAGTACCGAGGAAACTTCCTGTATAATACCAGTGTGAAAGACTTCAAGGCATCAAAGA TCGATGTCAAATCTAGTATCGAGAAGCTGAACGAGCTGTATGAAAATCTGGGCCTGGACCTGAACGTGGAGTTCA ACATTAGCAATACTGCCGAGATCGAAAAGGTGCTGAAAGACAAGCAGATCTTCAAGCGGGATAAAGTCAAGAAA ATTGCCGAGCTGTTTGCTATCAAAACCGACAACAAGGAACAGAGCAAGAGAATCAAAGATATTTCCAAACAGGT GGCCAATGCTGTCCTGGGGTACAAGACCAGGTTCGACACAATCGCTCTGAAAGAGATTTCCAAGGACGAACTGT CTGATTGGAACTTCAAACTGTCAGACATCGATGCAGACAGCAAGTTTGAGGCCCTGATGGGAAACCTGGATGAG AATGAACAGGCCATCCTGCTGACTATTAAGGAGCTGTTTAACGAAGTGACCCTGAATGGAATTGTCGAGGACGG CAACACCCTGAGCGAATCCATGATCAACAAGTACAATGATCACCGGGACGATCTGAAGCTGCTGAAAGAAGTGA TCGAAAATCATATTGACAGAAAGAAAGCCAAGGAGCTGGCACTGGCCTACGATCTGTATGTCAACAATAGGCAC GGACAGCTGCTGCAGGCTAAGAAAAAGCTGGGCAAAATCAAGCCCCGCTCTAAGGAGGACTTCTACAAAGTGGT CAACAAGAATCTGGACGATTCACGGGCAAGCAAGGAGATCAAAAAGAAAATTGAACTGGACAGCTTTATGCCTA AGCAGAGAACCAACGCCAATGGCGTGATCCCATACCAGCTGCAGCAGCTGGAGCTGGATAAGATCATCGAAAAC CAGTCTAAGTACTATCCATTCCTGAAGGAGATTAATCCCGTGTCAAGCCACCTGAAAGAGGCCCCCTATAAGCTG GACGAACTGATCCGATTTCGGGTGCCTTACTATGTCGGCCCCCTGATTTCTCCTAACGAGAGTACCAAGGATATC CAGACAAAGAAAAACCAGAATTTCGCCTGGATGATTCGCAAAGAGGAAGGGCGAATCACACCTTGGAACTTTGA CCAGAAGGTGGATCGAATTGAGAGCGCCAATAAGTTCATCAAACGGATGACTACCAAGGACACTTACCTGTTTG GGGAGGATGTGCTGCCAGCTAACAGCCTGCTGTATCAGAAGTTCACCGTCCTGAACGAACTGAACAACATCCGG ATTAATGGAAAAAGAATCTCCGTGGACCTGAAGCAGGAGATCTACGAAAACCTGTTTAAGAAACACACAACTGT GACCGTCAAGAAACTGGAGAATTATCTGAAGGAAAACCATAATCTGGTGAAAGTCGAGATCAAGGGGCTGGCCG ATGAAAAGAAATTCAACAGCGGACTGACCACATACAATAGATTCAAGAACCTGAACATCTTTGACAACCAGATT

FIG. 13R

	GACGATCTGAAGTACAGGAACGATTTCGAGAAGATCATCGAATGGTCTACAATTTTTGAGGACAAGAGTATCTAC AAAGAAAAGCTGAGGAGCATCGATTGGCTGAACGAGAAGCAGATTAACGCTCTGTCTAATATCAGACTGCAGGG GTGGGGAAGGCTGAGTAAGAAACTGCTGGCACAGCTGCACGACCATAATGGCCAGACCATCATTGAGCAGCTGT GGGATTCCCAGAACAATTTCATGCAGATTGTGACACAGGCCGACTTTAAAGATGCTATCGCAAAGGCCAACCAG AATCTGCTGGTGGCTACCTCAGTCGAGGACATTCTGAACAATGCATACACAAGCCCCGCAAACAAGAAAGCCAT CAGACAGGTCATCAAGGTGGTCGACGATATCGTGAAGGCAGCCTCCGGAAAGGTCCCAAAACAGATCGCCATTG AGTTCACTAGGGATGCTGACGAAAATCCCAAGAGAAGTCAGACCAGGGGCTCAAAGCTGCAGAAAGTGTACAAG GACCTGAGCACTGAGCTGGCCTCCAAGACCATTGCTGAGGAACTGAACGAAGCAATCAAAGACAAGAAACTGGT GCAGGATAAGTACTATCTGTACTTTATGCAGCTGGGGCGGGACGCCTATACAGGAGAGCCTATCAATATCGATGA AATCCAGAAGTACGATATCGACCACATTCTGCCACAGTCTTTCATCAAGGACGATGCCCTGGACAACAGGGTGCT GGTGAGCCGGGCTGTGAACAATGGCAAATCTGATAATGTGCCTGTCAAGCTGTTTGGCAACGAGATGGCTGCAA ATCTGGGGATGACTATCAGGAAAATGTGGGAGGAATGGAAGAACATCGGCCTGATTAGCAAAACAAAGTACAAC AATCTGCTGACTGATCCCGACCACATTAACAAGTATAAGAGTGCCGGGTTCATCAGGCGCCAGCTGGTGGAGAC ATCACAGATCATCAAGCTGGTGAGCACTATCCTGCAGAGTCGCTACCCTAACACTGAAATCATTACCGTGAAGGC TAAGTACAATCATTATCTGCGGGAGAAATTTGACCTGTATAAGAGCAGAGAAGTCAACGACTACCACCATGCTAT TGATGCATATCTGTCCGCCATCTGCGGAAATCTGCTGTACCAGAACTATCCAAATCTGCGGCCCTTCTTTGTGTAC GGCCAGTATAAGAAATTCTCCTCTGATCCTGACAAAGAGAAGGCCATTTTTAACAAAACCCGCAAGTTCTCCTTT ATCTCTCAGCTGCTGAAAAACAAGAGTGAGAACAGCAAGGAAATCGCTAAGAAACTGAAACGGGCATACCAGTT CAAGTATATGCTGGTGTCTCGAGAGACTGAAACCCGGGACCAGGAGATGTTCAAAATGACCGTGTACCCCCGGTT CAGCCACGATACAGTCAAGGCTCCTAGGAACCTGATTCCAAAGAAAATGGGCATGTCCCCTGACATCTACGGAG GCTATACAAACAATTCTGACGCATACATGGTCATCGTCCGCATTGATAAGAAAAAGGGAACTGAGTATAAGATC CTGGGCATTCCAACCCGGGAACTGGTGAATCTGAAAAAGGCCGAGAAGGAGGACCATTACAAAAGCTATCTGAA GGAGATCCTGACACCAAGGATTCTGTACAACAAAAATGGGAAGCGCGATAAAAAGATCACTTCCTTCGAAATTG TGAAATCTAAGATCCCCTATAAGCAGGTCATCCAGGATGGGGACAAAAAGTTTATGCTGGGAAGTTCAACATAC GTGTATAACGCAAAGCAGCTGACACTGAGCACTGAGTCCATGAAAGCCATCACTAACAATTTCGATAAGGACAG CGATGAGAACGACGCTCTGATTAAGGCATACGATGAAATCCTGGACAAAGTGGATAAGTATCTGCCACTGTTCG ACATCAACAAGTTCCGGGAGAAGCTGCACAGTGGGCGAGAAAAGTTCATCAAGCTGAGCCTGGAGGACAAAAA GGATACCATCCTGAAAGTGCTGGAAGGACTGCATGATAACGCTGTCATGACAAAGATCCCTACTATTGGCCTGTC CACACCACTGGGGTTCATGCAGTTTCCCAACGGCGTGATTCTGAGCGAGAATGCCAAACTGATCTACCAGTCCCC CACCGGGCTGTTCAAAAAGTCAGTGAAGATCAGCGACCTGTAAGAATTC
Length of Protein (without NLS and HA tag)	1375
PAM	**NTAAA**

FIG. 13S

Spacer Length	30
Species: Neisseria_cinerea	
Sequencing Barcode	ATTGTCTG
ID	18
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNGUUGUAGCUCCCAUUCUCGAAAGAGAACCGUUGCUACAAUAAGGCCGUCU GAAAAGAUGUGCCGCAACGCUCUGCCCCUUAAAGCUUCUGCUUUAAGGGGCAUCGUUUAUUUCGGUUAA AAAUGCCGUCUGAAACCGGUUUUUAGGUUUCAGACGGCA**
Cas9 Protein Sequence (wildtype, without NLS added)	MAAFKPNPMNYILGLDIGIASVGWAIVEIDEEENPIRLIDLGVRVFERAEVPKTGDSLAAARRLARSVRRLTRRRAHRL LRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELG ALLKGVADNTHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTFNRKDLQAELNLLFEKQKEFGNPHVSDGLKE GIETLLMTQRPALSGDAVQKMLGHCTFEPTEPKAAKNTYTAERFVWLTKLNNLRILEQGSERPLTDTERATLMDEPYR KSKLTYAQARKLLDLDDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSPELQDEIGTAFSLF KTDEDITGRLKDRVQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGNRYDEACTEIYGDHYGKKNTEEKIYLPPIPA DEIRNPVVLRALSQARKVINGVVRRYGSPARIHIETARDVGKSFKDRKEIEKRQEENRKDREKSAAKFREYFPNFVGEP KSKDILKLRLYEQQHGKCLYSGKEINLGRLNEKGYVEIDHALPFSRTWDDSFNNKVLALGSENQNKGNQTPYEYFNG KDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDGFKERNLNDTRYINRFLCQFVADHMLLTGKGKRRVFASNGQIT NLLRGFWGLRKVRAENDRHHALDAVVVACSTIAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQKAHFPQPWEFF AQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHKYVTPLFISRAPNRKMSGQGHMETVKSAKRLDEGI SVLRVPLTQLKLKDLEKMVNREREPKLYEALKARLEAHKDDPAKAFAEPFYKYDKAGNRTQQVKAVRVEQVQKTG VWVHNHNGIADNATIVRVDVFEKGGKYYLVPIYSWQVAKGILPDRAVVQGKDEEDWTVMDDSFEFKFVLYANDLIK LTAKKNEFLGYFVSLNRATGAIDIRTHDTDSTKGKNGIFQSVGVKTALSFQKYQIDELGKEIRPCRLKKRPPVR
Mammalian Codon Optimized Sequences (AgeI-HA-NLS-Cas9-NLS-taa-	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CATGGCTGCCTTCAAACCTAATCCTATGAACTACATCCTGGGCCTGGACATTGGAATCGCTTCTGTCGGGTGGGCT ATCGTGGAAATCGACGAGGAAGAGAACCCTATCAGACTGATTGATCTGGGAGTCAGAGTGTTTGAAAGGGCAGA GGTGCCAAAGACCGGCGACTCCCTGGCCGCTGCACGGAGACTGGCTCGGTCTGTCAGGCGCCTGACACGACGGA GAGCACACAGGCTGCTGCGAGCTAGGCGCCTGCTGAAGAGAGAGGGCGTGCTGCAGGCCGCTGACTTCGATGAA AACGGCCTGATCAAGAGCCTGCCCAATACTCCTTGGCAGCTGAGAGCAGCCGCTCTGGACAGGAAGCTGACCCC ACTGGAGTGGTCTGCCGTGCTGCTGCACCTGATCAAGCATCGCGGCTACCTGAGTCAGCGAAAAAATGAAGGGG

FIG. 13T

EcoRI)	AGACAGCAGATAAGGAGCTGGGAGCACTGCTGAAAGGAGTGGCCGACAACACTCATGCTCTGCAGACCGGCGAT TTTAGGACACCCGCTGAGCTGGCACTGAATAAGTTCGAAAAAGAGAGTGGACACATTCGAAACCAGCGGGGCGA CTATTCACATACCTTCAACCGCAAGGATCTGCAGGCCGAGCTGAATCTGCTGTTTGAAAAGCAGAAAGAGTTCGG GAATCCCCACGTGTCCGACGGGCTGAAAGAAGGAATCGAGACACTGCTGATGACTCAGAGGCCTGCACTGTCTG GCGATGCCGTGCAGAAGATGCTGGGGCATTGCACCTTTGAACCAACAGAGCCCAAGGCAGCCAAAAACACCTAC ACAGCCGAGAGGTTCGTGTGGCTGACAAAGCTGAACAATCTGCGCATCCTGGAACAGGGCAGTGAGCGGCCCCT GACTGACACCGAAAGAGCCACACTGATGGATGAGCCTTACAGGAAGTCTAAACTGACTTATGCCCAGGCTCGCA AGCTGCTGGACCTGGACGATACTGCCTTCTTTAAGGGCCTGAGGTACGGGAAAGATAATGCAGAAGCCAGCACC CTGATGGAGATGAAGGCCTATCACGCTATCTCCCGCGCCCTGGAAAAAGAGGGCCTGAAGGACAAGAAATCTCC CCTGAACCTGAGTCCTGAACTGCAGGATGAGATTGGGACCGCTTTTAGCCTGTTCAAGACTGACGAGGATATCAC CGGACGCCTGAAAGACCGAGTGCAGCCCGAAATTCTGGAGGCACTGCTGAAGCACATCAGTTTTGATAAATTCGT GCAGATTTCACTGAAGGCCCTGCGACGGATCGTCCCTCTGATGGAGCAGGGCAATCGGTACGACGAGGCCTGCA CCGAGATCTACGGAGATCATTATGGCAAGAAAAACACAGAAGAGAAAATCTATCTGCCCCCTATTCCTGCCGAC GAGATCCGGAATCCAGTGGTCCTGAGAGCTCTGTCACAGGCAAGAAAAGTGATCAACGGAGTGGTCAGAAGGTA CGGCAGCCCTGCTAGGATCCACATTGAAACCGCACGCGAAGTGGGAAAGTCCTTTAAAGACCGCAAGGAAATCG AGAAGCGACAGGAAGAGAATAGAAAAGATAGGGAAAAGTCTGCTGCAAAATTCAGGGAGTACTTTCCAAACTTC GTGGGCGAACCCAAGAGTAAAGACATCCTGAAGCTGCGCCTGTACGAGCAGCAGCACGGGAAGTGTCTGTATAG CGGAAAAGAAATTAACCTGGGCCGGCTGAATGAAAAGGGCTATGTGGAGATCGATCATGCACTGCCCTTTTCCA GAACATGGGACGATTCTTTCAACAATAAGGTCCTGGCTCTGGGGAGCGAGAACCAGAACAAGGGAAATCAGACT CCTTACGAATATTTCAACGGGAAGGACAATAGCCGAGAATGGCAGGAGTTTAAAGCCCGCGTGGAGACAAGCCG GTTCCCACGAAGCAAGAAACAGCGGATTCTGCTGCAGAAGTTTGACGAAGATGGATTCAAAGAGAGAAACCTGA ATGACACCCGGTACATCAACAGATTTCTGTGCCAGTTCGTGGCTGATCACATGCTGCTGACCGGAAAGGGCAAAC GCCGAGTCTTTGCAAGCAACGGCCAGATCACAAATCTGCTGAGGGGCTTCTGGGGGCTGCGGAAGGTGAGAGCC GAGAATGACCGCCACCATGCACTGGATGCCGTGGTCGTGGCTTGTTCCACTATTGCAATGCAGCAGAAGATCACC AGGTTTGTGCGCTATAAAGAGATGAACGCCTTCGACGGAAAGACAATTGATAAAGAAACTGGCGAGGTGCTGCA CCAGAAGGCACATTTTCCTCAGCCATGGGAGTTCTTCGCCCAGGAAGTGATGATCCGGGTCTTTGGGAAGCCTGA CGGAAAACCAGAGTTCGAAGAGGCCGATACCCCAGAAAAGCTGCGGACACTGCTGGCTGAAAAACTGAGCTCCA GACCCGAGGCAGTGCACAAGTACGTCACCCCCCTGTTCATTAGCAGGGCCCCTAATCGCAAAATGTCCGGGCAG GGACATATGGAGACTGTGAAATCAGCTAAGCGGCTGGACGAAGGCATCAGCGTGCTGAGAGTCCCACTGACCCA GCTGAAGCTGAAAGATCTGGAGAAGATGGTGAACCGGGAAAGAGAGCCCAAGCTGTATGAAGCTCTGAAAGCA AGACTGGAGGCCCACAAGGACGATCCAGCTAAAGCATTTGCCGAGCCCTTCTACAAATATGACAAGGCCGGCAA TCGGACACAGCAGGTGAAGGCTGTCAGAGTGGAGCAGGTCCAGAAAACTGGGGTCTGGGTGCACAACCATAATG GAATTGCCGACAACGCTACAATCGTCCGGGTGGATGTGTTCGAGAAAGGCGGGAAGTACTATCTGGTGCCTATCT ACTCCTGGCAGGTCGCCAAGGGAATCCTGCCAGATAGAGCTGTCGTGCAGGGCAAAGACGAAGAGGATTGGACT GTGATGGACGATTCTTTCGAGTTTAAGTTCGTCCTGTACGCAAACGACCTGATCAAGCTGACAGCCAAGAAAAAT GAATTTCTGGGGTATTTCGTGTCACTGAACAGGGCAACTGGAGCCATCGATATTCGCACACATGACACTGATAGC

FIG. 13U

	ACCAAGGGAAAAAACGGCATCTTTCAGTCTGTGGGCGTCAAGACCGCCCTGAGTTTCCAGAAATATCAGATTGAC GAACTGGGGAAGGAGATCCGACCCTGTCGGCTGAAGAAACGACCACCCGTGCGGTAAGAATTC
Length of Protein (without NLS and HA tag)	1082
PAM	**NNNNGTAA**
Spacer Length	30
Species: Parvibaculum_lavamentivorans	
Sequencing Barcode	CGCCTTCC
ID	20
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNGCUGCGGAUUGCGGGAAAUCGCUUUUCGCAAGCAAAUUGACCCCUUGUGC GGGCUCGGCAUCCCAAGGUCAGCUGCCGGUUAUUAUCGAAAAGGCCCACCGCAAGCAGCGCGUGGGCCU UUUUUU**
Cas9 Protein Sequence (wildtype, without NLS added)	MERIFGFDIGTTSIGFSVIDYSSTQSAGNIQRLGVRIFPEARDPDGTPLNQQRRQKRMMRRQLRRRRIRRKALNETLHEA GFLPAYGSADWPVVMADEPYELRRRGLEEGLSAYEFGRAIYHLAQHRHFKGRELEESDTPDPDVDDEKEAANERAAT LKALKNEQTTLGAWLARRPPSDRKRGIHAHRNVVAEEFERLWEVQSKFHPALKSEEMRARISDTIFAQRPVFWRKNT LGECRFMPGEPLCPKGSWLSQQRRMLEKLNNLAIAGGNARPLDAEERDAILSKLQQQASMSWPGVRSALKALYKQR GEPGAEKSLKFNLELGGESKLLGNALEAKLADMFGPDWPAHPRKQEIRHAVHERLWAADYGETPDKKRVIILSEKDR KAHREAAANSFVADFGITGEQAAQLQALKLPTGWEPYSIPALNLFLAELEKGERFGALVNGPDWEGWRRTNFPHRNQ PTGEILDKLPSPASKEERERISQLRNPTVVRTQNELRKVVNNLIGLYGKPDRIRIEVGRDVGKSKREREEIQSGIRRNEKQ RKKATEDLIKNGIANPSRDDVEKWILWKEGQERCPYTGDQIGFNALFREGRYEVEHIWPRSRSFDNSPRNKTLCRKDV NIEKGNRMPFEAFGHDEDRWSAIQIRLQGMVSAKGGTGMSPGKVKRFLAKTMPEDFAARQLNDTRYAAKQILAQLK RLWPDMGPEAPVKVEAVTGQVTAQLRKLWTLNNILADDGEKTRADHRHHAIDALTVACTHPGMTNKLSRYWQLRD DPRAEKPALTPPWDTIRADAEKAVSEIVVSHRVRKKVSGPLHKETTYGDTGTDIKTKSGTYRQFVTRKKIESLSKGELD EIRDPRIKEIVAAHVAGRGGDPKKAFPPYPCVSPGGPEIRKVRLTSKQQLNLMAQTGNGYADLGSNHHIAIYRLPDGKA DFEIVSLFDASRRLAQRNPIVQRTRADGASFVMSLAAGEAIMIPEGSKKGIWIVQGVWASGQVVLERDTDADHSTTTR PMPNPILKDDAKKVSIDPIGRVRPSND

FIG. 13V

Mammalian Codon Optimized Sequences (AgeI-HA-NLS-Cas9-NLS-taa-EcoRI)	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CATGGAGAGGATTTTCGGCTTTGACATCGGCACAACAAGTATCGGATTCAGCGTGATTGATTACAGTAGCACCCA GTCCGCAGGCAACATCCAGAGGCTGGGCGTGCGCATTTTCCCTGAGGCAAGGGACCCAGATGGGACCCCCCTGA ACCAGCAGCGGAGACAGAAACGCATGATGAGGCGCCAGCTGCGACGGAGAAGGATTCGCCGAAAGGCACTGAA TGAGACACTGCACGAAGCCGGCTTTCTGCCAGCTTACGGGTCTGCAGATTGGCCCGTGGTCATGGCCGACGAGCC TTATGAACTGCGGAGAAGGGGACTGGAGGAAGGCCTGAGTGCTTACGAGTTCGGACGGGCAATCTATCATCTGG CCCAGCACCGGCATTTTAAAGGCAGAGAACTGGAGGAATCCGATACACCCGACCCTGATGTGGACGATGAGAAG GAAGCCGCTAACGAGAGAGCAGCCACTCTGAAGGCCCTGAAAAATGAACAGACCACACTGGGAGCATGGCTGG CCCGCCGACCCCCTTCTGACCGCAAGCGAGGAATCCACGCCCATAGGAACGTGGTCGCTGAGGAGTTCGAGCGC CTGTGGGAAGTGCAGTCCAAGTTTCACCCCGCTCTGAAATCTGAGGAAATGCGGGCAAGAATCAGTGATACAATT TTCGCCCAGAGGCCTGTGTTTTGGCGCAAGAACACTCTGGGAGAGTGCAGATTCATGCCTGGCGAACCACTGTGT CCCAAGGGGTCCTGGCTGTCTCAGCAGCGGAGAATGCTGGAGAAACTGAACAATCTGGCTATCGCAGGCGGGAA TGCTAGGCCACTGGATGCAGAGGAACGCGACGCCATTCTGAGTAAGCTGCAGCAGCAGGCCAGCATGTCCTGGC CAGGCGTGCGGTCAGCTCTGAAGGCACTGTACAAACAGAGAGGCGAGCCCGGGGCTGAAAAGAGCCTGAAATTC AACCTGGAGCTGGGAGGCGAATCCAAGCTGCTGGGAAATGCCCTGGAGGCTAAACTGGCAGATATGTTTGGCCC TGACTGGCCAGCTCACCCCCGAAAGCAGGAGATCCGGCACGCAGTGCATGAACGGCTGTGGGCTGCAGATTACG GCGAGACACCCGACAAGAAAAGAGTCATCATTCTGTCCGAGAAGGATCGAAAAGCTCATCGGGAAGCCGCTGCA AACTCTTTCGTGGCAGACTTTGGAATTACTGGCGAGCAGGCAGCTCAGCTGCAGGCCCTGAAGCTGCCAACCGGC TGGGAACCTTATAGCATCCCAGCACTGAACCTGTTCCTGGCCGAGCTGGAAAAGGGGGAGAGGTTTGGAGCCCT GGTGAATGGACCTGATTGGGAAGGCTGGAGGCGCACAAACTTCCCCCACCGCAATCAGCCTACTGGGGAGATCC TGGACAAGCTGCCAAGTCCCGCCTCAAAAGAGGAAAGGGAACGCATTAGCCAGCTGCGCAACCCAACCGTGGTC CGAACACAGAATGAGCTGAGAAAGGTGGTCAACAATCTGATCGGGCTGTATGGAAAACCCGATCGAATCCGGAT TGAAGTGGGCCGGGACGTCGGGAAGTCCAAAAGAGAAAGGGAGGAAATCCAGTCTGGCATTCGACGGAACGAG AAGCAGAGAAAGAAAGCCACTGAAGATCTGATCAAAAACGGAATTGCTAATCCTAGCCGGGACGATGTGGAGA AGTGGATCCTGTGGAAAGAGGGCCAGGAAAGATGCCCATACACCGGCGACCAGATTGGCTTCAATGCCCTGTTT AGAGAAGGCAGATATGAGGTGGAACACATCTGGCCTCGCTCTCGAAGTTTTGATAACAGCCCAAGGAATAAGAC ACTGTGTCGCAAAGACGTGAACATCGAGAAGGGAAATAGGATGCCTTTCGAGGCATTTGGCCATGACGAAGATC GGTGGAGCGCCATCCAGATTAGACTGCAGGGCATGGTGTCAGCCAAAGGGGGAACTGGGATGAGCCCCGGAAA GGTCAAACGCTTCCTGGCTAAGACCATGCCTGAGGATTTTGCAGCCCGGCAGCTGAACGACACAAGATACGCTGC AAAGCAGATCCTGGCCCAGCTGAAAAGGCTGTGGCCAGACATGGGACCTGAGGCTCCAGTGAAGGTCGAAGCAG TGACTGGACAGGTCACCGCCCAGCTGCGCAAACTGTGGACTCTGAACAATATTCTGGCTGACGATGGGGAGAAA ACCAGAGCAGATCACAGGCACCATGCCATCGACGCTCTGACAGTGGCCTGCACTCATCCTGGAATGACCAACAA GCTGAGCAGGTATTGGCAGCTGCGCGACGATCCACGAGCAGAGAAGCCAGCTCTGACTCCACCCTGGGATACCA TCCGCGCCGACGCTGAGAAAGCCGTGTCTGAAATTGTGGTCAGTCACCGGGTGAGAAAGAAAGTCAGCGGCCCA CTGCATAAGGAGACTACCTACGGCGATACAGGGACTGACATTAAGACCAAATCCGGCACATATAGACAGTTCGT

FIG. 13W

	GACCAGGAAGAAAATCGAGTCACTGAGCAAGGGGGAGCTGGATGAAATTCGCGACCCCCGAATCAAAGAAATT GTGGCAGCTCACGTCGCAGGACGAGGAGGCGACCCCAAGAAGGCCTTCCCTCCATACCCCTGTGTGTCTCCCGGA GGCCCTGAGATCCGGAAGGTCAGACTGACCAGTAAACAGCAGCTGAACCTGATGGCCCAGACAGGGAATGGATA CGCTGACCTGGGCTCCAACCACCATATCGCAATCTACCGGCTGCCCGATGGGAAGGCCGACTTCGAGATTGTGTC ACTGTTTGATGCTAGCAGAAGGCTGGCACAGAGAAATCCAATCGTGCAGAGGACACGAGCAGACGGAGCCAGCT TCGTCATGTCCCTGGCAGCCGGAGAGGCCATCATGATTCCCGAAGGCTCAAAGAAAGGGATCTGGATTGTGCAG GGAGTCTGGGCAAGCGGACAGGTGGTCCTGGAGAGGGACACCGATGCTGACCACTCTACAACTACCCGCCCTAT GCCAAACCCCATCCTGAAGGACGATGCCAAGAAAGTGAGTATCGATCCTATTGGCCGAGTCCGGCCATCAAATG ACTAAGAATTC
Length of Protein (without NLS and HA tag)	1037
PAM	**NNNCAT**
Spacer Length	30
Species: Staphylococcus aureus subsp. Aureus -- double check	
Sequencing Barcode	CTATGCGT
ID	21
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNNGUUUUAGUACUCUGGAAACAGAAUCUACUAAAACAAGGCAAAAUGCCGUG UUUAUCUCGUCAACUUGUUGGCGAGAUUUUUUU**
Cas9 Protein Sequence (wildtype, without NLS added)	MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTD HSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERL KKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMG HCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYR VTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSL KAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAR EKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIP RSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQK

FIG. 13X

	DFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKE WKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDD KGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKK DNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYE EAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKK YSTDILGNLYEVKSKKHPQIIKKG
Mammalian Codon Optimized Sequences (AgeI-HA-NLS-Cas9-NLS-taa-EcoRI)	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CATGAAAAGGAACTACATTCTGGGGCTGGACATCGGGATTACAAGCGTGGGGTATGGGATTATTGACTATGAAA CAAGGGACGTGATCGACGCAGGCGTCAGACTGTTCAAGGAGGCCAACGTGGAAAACAATGAGGGACGGAGAAG CAAGAGGGGAGCCAGGCGCCTGAAACGACGGAGAAGGCACAGAATCCAGAGGGTGAAGAAACTGCTGTTCGAT TACAACCTGCTGACCGACCATTCTGAGCTGAGTGGAATTAATCCTTATGAAGCCAGGGTGAAAGGCCTGAGTCAG AAGCTGTCAGAGGAAGAGTTTTCCGCAGCTCTGCTGCACCTGGCTAAGCGCCGAGGAGTGCATAACGTCAATGA GGTGGAAGAGGACACCGGCAACGAGCTGTCTACAAAGGAACAGATCTCACGCAATAGCAAAGCTCTGGAAGAG AAGTATGTCGCAGAGCTGCAGCTGGAACGGCTGAAGAAAGATGGCGAGGTGAGAGGGTCAATTAATAGGTTCAA GACAAGCGACTACGTCAAAGAAGCCAAGCAGCTGCTGAAAGTGCAGAAGGCTTACCACCAGCTGGATCAGAGCT TCATCGATACTTATATCGACCTGCTGGAGACTCGGAGAACCTACTATGAGGGACCAGGAGAAGGGAGCCCCTTC GGATGGAAAGACATCAAGGAATGGTACGAGATGCTGATGGGACATTGCACCTATTTTCCAGAAGAGCTGAGAAG CGTCAAGTACGCTTATAACGCAGATCTGTACAACGCCCTGAATGACCTGAACAACCTGGTCATCACCAGGGATGA AAACGAGAAACTGGAATACTATGAGAAGTTCCAGATCATCGAAAACGTGTTTAAGCAGAAGAAAAAGCCTACAC TGAAACAGATTGCTAAGGAGATCCTGGTCAACGAAGAGGACATCAAGGGCTACCGGGTGACAAGCACTGGAAA ACCAGAGTTCACCAATCTGAAAGTGTATCACGATATTAAGGACATCACAGCACGGAAAGAAATCATTGAGAACG CCGAACTGCTGGATCAGATTGCTAAGATCCTGACTATCTACCAGAGCTCCGAGGACATCCAGGAAGAGCTGACTA ACCTGAACAGCGAGCTGACCCAGGAAGAGATCGAACAGATTAGTAATCTGAAGGGGTACACCGGAACACACAA CCTGTCCCTGAAAGCTATCAATCTGATTCTGGATGAGCTGTGGCATACAAACGACAATCAGATTGCAATCTTTAA CCGGCTGAAGCTGGTCCCAAAAAAGGTGGACCTGAGTCAGCAGAAAGAGATCCCAACCACACTGGTGGACGATT TCATTCTGTCACCCGTGGTCAAGCGGAGCTTCATCCAGAGCATCAAAGTGATCAACGCCATCATCAAGAAGTACG GCCTGCCCAATGATATCATTATCGAGCTGGCTAGGGAGAAGAACAGCAAGGACGCACAGAAGATGATCAATGAG ATGCAGAAACGAAACCGGCAGACCAATGAACGCATTGAAGAGATTATCCGAACTACCGGGAAAGAGAACGCAA AGTACCTGATTGAAAAAATCAAGCTGCACGATATGCAGGAGGGAAAGTGTCTGTATTCTCTGGAGGCCATCCCCC TGGAGGACCTGCTGAACAATCCATTCAACTACGAGGTCGATCATATTATCCCCAGAAGCGTGTCCTTCGACAATT CCTTTAACAACAAGGTGCTGGTCAAGCAGGAAGAGAACTCTAAAAAGGGCAATAGGACTCCTTTCCAGTACCTG TCTAGTTCAGATTCCAAGATCTCTTACGAAACCTTTAAAAAGCACATTCTGAATCTGGCCAAAGGAAAGGGCCGC ATCAGCAAGACCAAAAAGGAGTACCTGCTGGAAGAGCGGGACATCAACAGATTCTCCGTCCAGAAGGATTTTAT TAACCGGAATCTGGTGGACACAAGATACGCTACTCGCGGCCTGATGAATCTGCTGCGATCCTATTTCCGGGTGAA CAATCTGGATGTGAAAGTCAAGTCCATCAACGGCGGGTTCACATCTTTTCTGAGGCGCAAATGGAAGTTTAAAAA

FIG. 13Y

	GGAGCGCAACAAAGGGTACAAGCACCATGCCGAAGATGCTCTGATTATCGCAAATGCCGACTTCATCTTTAAGG AGTGGAAAAAGCTGGACAAAGCCAAGAAAGTGATGGAGAACCAGATGTTCGAAGAGAAGCAGGCCGAATCTAT GCCCGAAATCGAGACAGAACAGGAGTACAAGGAGATTTTCATCACTCCTCACCAGATCAAGCATATCAAGGATT TCAAGGACTACAAGTACTCTCACCGGGTGGATAAAAAGCCCAACAGAGAGCTGATCAATGACACCCTGTATAGT ACAAGAAAAGACGATAAGGGGAATACCCTGATTGTGAACAATCTGAACGGACTGTACGACAAAGATAATGACA AGCTGAAAAAGCTGATCAACAAAAGTCCCGAGAAGCTGCTGATGTACCACCATGATCCTCAGACATATCAGAAA CTGAAGCTGATTATGGAGCAGTACGGCGACGAGAAGAACCCACTGTATAAGTACTATGAAGAGACTGGGAACTA CCTGACCAAGTATAGCAAAAAGGATAATGGCCCCGTGATCAAGAAGATCAAGTACTATGGGAACAAGCTGAATG CCCATCTGGACATCACAGACGATTACCCTAACAGTCGCAACAAGGTGGTCAAGCTGTCACTGAAGCCATACAGAT TCGATGTCTATCTGGACAACGGCGTGTATAAATTTGTGACTGTCAAGAATCTGGATGTCATCAAAAAGGAGAACT ACTATGAAGTGAATAGCAAGTGCTACGAAGAGGCTAAAAAGCTGAAAAAGATTAGCAACCAGGCAGAGTTCATC GCCTCCTTTTACAACAACGACCTGATTAAGATCAATGGCGAACTGTATAGGGTCATCGGGGTGAACAATGATCTG CTGAACCGCATTGAAGTGAATATGATTGACATCACTTACCGAGAGTATCTGGAAAACATGAATGATAAGCGCCCC CCTCGAATTATCAAAACAATTGCCTCTAAGACTCAGAGTATCAAAAAGTACTCAACCGACATTCTGGGAAACCTG TATGAGGTGAAGAGCAAAAAGCACCCTCAGATTATCAAAAAGGGCTAAGAATTC
Length of Protein (without NLS and HA tag)	1053
PAM	**NNGRRT**
Spacer Length	30
Species: Mycoplasma gallisepticum str. F	
Sequencing Barcode	CGCTATGT
ID	23
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNGUUUUAGCACUGUACAAGAAAUUGUCGUGCUAAAAUAAGGCGCUGUUAAU GCAGCUGCCGCAUCCGCCAGAGCAUUUAUGCUCUGGCUUUUUUU**
Cas9 Protein Sequence	MNNSIKSKPEVTIGLDLGVGSVGWAIVDNETNIIHHLGSRLFSQAKTAEDRRSFRGVRRLIRRRKYKLKRFVNLIWKYN SYFGFKNKEDILNNYQEQQKLHNTVLNLKSEALNAKIDPKALSWILHDYLKNRGHFYEDNRDFNVYPTKELAKYFDK

FIG. 13Z

(wildtype, without NLS added)	YGYYKGIIDSKEDNDNKLEEELTKYKFSNKHWLEEVKKVLSNQTGLPEKFKEEYESLFSYVRNYSEGPGSINSVSPYGI YHLDEKEGKVVQKYNNIWDKTIGKCNIFPDEYRAPKNSPIAMIFNEINELSTIRSYSIYLTGWFINQEFKKAYLNKLLDL LIKTNGEKPIDARQFKKLREETIAESIGKETLKDVENEEKLEKEDHKWKLKGLKLNTNGKIQYNDLSSLAKFVHKLKQ HLKLDFLLEDQYATLDKINFLQSLFVYLGKHLRYSNRVDSANLKEFSDSNKLFERILQKQKDGLFKLLFEQTDKDDEKIL AQTHSLSTKAMLLAITRMTNLDNDEDNQKNNDKGWNFEAIKNFDQKFIDITKKNNNLSLKQNKRYLDDRFINDAILSP GVKRILREATKVFNAILKQFSEEYDVTKVVIELARELSEEKELENTKNYKKLIKKNGDKISEGLKALGISEDEIKDILKSP TKSYKFLLWLQQDHIDPYSLKEIAFDDIFTKTEKFEIDHIIPYSISFDDSSSNKLLVLAESNQAKSNQTPYEFISSGNAGIK WEDYEAYCRKFKDGDSSLLDSTQRSKKFAKMMKTDTSSKYDIGFLARNLNDTRYATIVFRDALEDYANNHLVEDKP MFKVVCINGSVTSFLRKNFDDSSYAKKDRDKNIHHAVDASIISIFSNETKTLFNQLTQFADYKLFKNTDGSWKKIDPKT GVVTEVTDENWKQIRVRNQVSEIAKVIEKYIQDSNIERKARYSRKIENKTNISLFNDTVYSAKKVGYEDQIKRKNLKTL DIHESAKENKNSKVKRQFVYRKLVNVSLLNNDKLADLFAEKEDILMYRANPWVINLAEQIFNEYTENKKIKSQNVFEK YMLDLTKEFPEKFSEFLVKSMLRNKTAIIYDDKKNIVHRIKRLKMLSSELKENKLSNVIIRSKNQSGTKLSYQDTINSLA LMIMRSIDPTAKKQYIRVPLNTLNLHLGDHDFDLHNMDAYLKKPKFVKYLKANEIGDEYKPWRVLTSGTLLIHKKDK KLMYISSFQNLNDVIEIKNLIETEYKENDDSDSKKKKKANRFLMTLSTILNDYILLDAKDNFDILGLSKNRIDEILNSKLG LDKIVK
Mammalian Codon Optimized Sequences (AgeI-HA-NLS-Cas9-NLS-taa-EcoRI)	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CATGAACAATAGCATCAAATCTAAACCTGAAGTGACCATCGGGCTGGACCTGGGAGTGGGAAGCGTGGGGTGGG CAATCGTGGATAACGAAACAAACATCATTCACCATCTGGGCTCCAGGCTGTTTTCTCAGGCCAAGACTGCTGAGG ATCGGAGATCTTTCCGCGGGGTGAGGCGCCTGATCCGACGGAGAAAATACAAGCTGAAACGATTCGTCAATCTG ATTTGGAAGTACAACAGCTATTTCGGCTTCAAGAACAAAGAGGACATCCTGAACAATTATCAGGAGCAGCAGAA GCTGCACAATACCGTGCTGAACCTGAAATCAGAGGCACTGAATGCCAAGATCGATCCTAAAGCACTGAGCTGGA TTCTGCACGACTACCTGAAGAACAGAGGCCATTTTTATGAGGACAATAGGGATTTCAACGTGTACCCAACAAAGG AGCTGGCCAAGTACTTCGATAAGTACGGGTACTACAAGGGAATCATTGACAGCAAGGAGGACAATGATAACAAA CTGGAGGAAGAGCTGACAAAGTACAAATTCTCCAATAAGCACTGGCTGGAAGAGGTGAAGAAAGTCCTGTCTAA CCAGACTGGCCTGCCAGAAAAGTTTAAAGAAGAGTATGAGTCACTGTTCAGCTACGTGAGAAATTATTCAGAGG GCCCAGGGAGCATCAACTCTGTCAGTCCCTACGGGATCTACCATCTGGACGAAAAAGAGGGAAAGGTGGTCCAG AAGTACAACAACATCTGGGATAAGACAATCGGAAAGTGCAACATCTTCCCTGACGAGTATAGAGCTCCCAAGAA CAGTCCTATCGCAATGATTTTCAATGAAATCAACGAGCTGTCCACAATCAGGTCATACAGCATCTACCTGACTGG CTGGTTCATTAATCAGGAGTTCAAGAAAGCCTACCTGAACAAGCTGCTGGATCTGCTGATCAAAACCAACGGAG AGAAGCCAATTGACGCAAGGCAGTTCAAGAAACTGCGCGAAGAGACAATCGCCGAAAGCATTGGCAAAGAGAC ACTGAAGGATGTGGAGAATGAAGAGAAACTGGAAAAGGAGGACCACAAGTGGAAACTGAAGGGACTGAAGCTG AATACCAACGGCAAAATCCAGTACAACGATCTGAGCTCCCTGGCTAAGTTTGTGCACAAACTGAAGCAGCATCTG AAACTGGATTTCCTGCTGGAGGACCAGTATGCAACACTGGACAAGATCAATTTCCTGCAGTCCCTGTTTGTGTAC CTGGGCAAGCACCTGAGATATTCCAATAGGGTCGATTCTGCCAACCTGAAGGAATTTTCCGACTCTAACAAACTG TTCGAGCGCATCCTGCAGAAACAGAAGGATGGGCTGTTCAAGCTGTTTGAACAGACTGACAAAGACGATGAGAA

FIG. 13AA

	GATCCTGGCCCAGACACATAGTCTGTCAACTAAGGCCATGCTGCTGGCTATTACCCGGATGACAAATCTGGACAA CGATGAGGACAACCAGAAAAACAATGACAAGGGCTGGAATTTTGAGGCCATCAAAAACTTCGATCAGAAGTTTA TCGACATCACCAAGAAAAACAACAACCTGAGCCTGAAACAGAATAAGCGCTACCTGGACGATCGATTCATCAAC GATGCTATTCTGTCCCCTGGGGTGAAGCGAATCCTGCGGGAGGCAACCAAGGTCTTTAATGCCATTCTGAAACAG TTCTCTGAAGAGTACGACGTGACAAAGGTGGTCATCGAACTGGCTCGCGAGCTGAGCGAAGAGAAGGAACTGGA GAACACAAAGAACTACAAGAAACTGATCAAGAAAAACGGCGACAAGATTAGTGAGGGCCTGAAAGCACTGGGG ATCTCAGAAGATGAGATCAAAGACATTCTGAAGAGTCCCACTAAATCATACAAGTTTCTGCTGTGGCTGCAGCAG GACCACATCGATCCTTATAGCCTGAAGGAGATCGCCTTCGACGATATTTTTACCAAAACAGAAAAGTTCGAGATC GACCATATCATTCCCTACAGCATTTCCTTCGACGATTCTAGTTCAAACAAGCTGCTGGTGCTGGCTGAAAGTAATC AGGCAAAGTCAAACCAGACTCCTTATGAGTTCATCAGCTCCGGAAACGCAGGCATTAAGTGGGAAGATTACGAG GCCTATTGCCGCAAGTTCAAGGATGGGGACTCTAGTCTGCTGGACAGCACCCAGCGGTCCAAGAAATTCGCCAA AATGATGAAAACCGATACCTCAAGCAAGTACGACATCGGATTTCTGGCTCGAAATCTGAACGATACTCGGTACGC AACCATTGTGTTCCGGGACGCCCTGGAGGACTATGCTAATAACCACCTGGTCGAGGACAAACCCATGTTTAAGGT GGTCTGTATCAATGGGTCCGTGACCTCTTTCCTGCGGAAGAACTTTGACGATTCCTCTTACGCCAAGAAAGATAG AGACAAGAATATCCACCATGCTGTGGATGCAAGTATCATCTCAATTTTCAGCAACGAGACAAAGACTCTGTTCAA CCAGCTGACTCAGTTTGCTGACTATAAACTGTTCAAGAACACCGATGGCAGCTGGAAGAAAATCGACCCTAAGA CAGGGGTGGTCACTGAAGTGACCGACGAGAATTGGAAGCAGATTAGGGTGCGCAACCAGGTGAGCGAAATCGCC AAAGTCATTGAGAAGTACATCCAGGATAGCAACATCGAAAGAAAGGCTAGGTATTCCCGCAAAATCGAGAATAA GACTAACATTTCCCTGTTTAATGACACCGTGTACTCTGCCAAGAAAGTCGGCTATGAGGATCAGATCAAAAGAAA GAACCTGAAAACCCTGGACATTCACGAATCTGCTAAAGAGAATAAGAACAGTAAAGTGAAGCGGCAGTTTGTCT ACAGAAAGCTGGTGAATGTCAGCCTGCTGAATAACGATAAGCTGGCAGACCTGTTCGCCGAAAAAGAGGATATC CTGATGTATAGGGCCAATCCATGGGTCATCAACCTGGCTGAGCAGATTTTCAATGAATACACTGAGAACAAGAA AATCAAGTCCCAGAACGTGTTTGAAAAATATATGCTGGACCTGACCAAAGAGTTCCCCGAGAAGTTCAGCGAGTT TCTGGTGAAGTCCATGCTGAGAAACAAGACCGCCATCATCTACGACGATAAGAAAAACATTGTCCATCGAATCA AACGGCTGAAGATGCTGAGTTCAGAACTGAAAGAGAATAAGCTGTCTAACGTGATCATTAGGTCTAAGAATCAG AGTGGGACCAAACTGTCATACCAGGATACAATCAACAGCCTGGCCCTGATGATTATGCGCAGCATCGACCCTACT GCTAAGAAACAGTATATTCGAGTGCCACTGAATACCCTGAACCTGCACCTGGGAGATCATGACTTTGATCTGCAC AATATGGATGCTTACCTGAAGAAACCAAAATTCGTGAAGTATCTGAAAGCAAACGAAATCGGCGACGAGTACAA GCCCTGGAGGGTCCTGACATCTGGCACTCTGCTGATCCATAAGAAGGATAAGAAACTGATGTACATCAGCTCCTT CCAGAATCTGAACGACGTGATCGAAATTAAGAATCTGATCGAAACCGAGTATAAAGAGAACGACGATTCTGATA GTAAGAAAAAGAAAAAGGCAAACCGCTTTCTGATGACCCTGAGCACAATCCTGAATGACTACATTCTGCTGGAC GCCAAGGATAACTTCGACATCCTGGGGCTGTCTAAAAATCGGATCGATGAGATTCTGAACAGTAAGCTGGGACT GGACAAGATTGTGAAATAAGAATTC
Length of Protein	1269

FIG. 13BB

(without NLS and HA tag)	
PAM	**NNGAT**
Spacer Length	30
Species: Campylobacter lari CF89-12	
Sequencing Barcode	CTGTGGCG
ID	24
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNGUUUUAGUCUCUGAAAAGAGACUAAAAUAAGUGGUUUUUGGUCAUCCACG CAGGGUUACAAUCCCUUUAAAACCAUUAAAAUUCAAAUAAACUAGGUUGUAUCAACUUAGUUUUUUU**
Cas9 Protein Sequence (wildtype, without NLS added)	MRILGFDIGINSIGWAFVENDELKDCGVRIFTKAENPKNKESLALPRRNARSSRRRLKRRKARLIAIKRILAKELKLNYK DYVAADGELPKAYEGSLASVYELRYKALTQNLETKDLARVILHIAKHRGYMNKNEKKSNDAKKGKILSALKNNALK LENYQSVGEYFYKEFFQKYKKNTKNFIKIRNTKDNYNNCVLSSDLEKELKLILEKQKEFGYNYSEDFINEILKVAFFQR PLKDFSHLVGACTFFEEEKRACKNSYSAWEFVALTKIINEIKSLEKISGEIVPTQTINEVLNLILDKGSITYKKFRSCINLH ESISFKSLKYDKENAENAKLIDFRKLVEFKKALGVHSLSRQELDQISTHITLIKDNVKLKTVLEKYNLSNEQINNLLEIEF NDYINLSFKALGMILPLMREGKRYDEACEIANLKPKTVDEKKDFLPAFCDSIFAHELSNPVVNRAISEYRKVLNALLKK YGKVHKIHLELARDVGLSKKAREKIEKEQKENQAVNAWALKECENIGLKASAKNILKLKLWKEQKEICIYSGNKISIE HLKDEKALEVDHIYPYSRSFDDSFINKVLVFTKENQEKLNKTPFEAFGKNIEKWSKIQTLAQNLPYKKKNKILDENFKD KQQEDFISRNLNDTRYIATLIAKYTKEYLNFLLLSENENANLKSGEKGSKIHVQTISGMLTSVLRHTWGFDKKDRNNH LHHALDAIIVAYSTNSIIKAFSDFRKNQELLKARFYAKELTSDNYKHQVKFFEPFKSFREKILSKIDEIFVSKPPRKRARR ALHKDTFHSENKIIDKCSYNSKEGLQIALSCGRVRKIGTKYVENDTIVRVDIFKKQNKFYAIPIYAMDFALGILPNKIVIT GKDKNNNPKQWQTIDESYEFCFSLYKNDLILLQKKNMQEPEFAYYNDFSISTSSICVEKHDNKFENLTSNQKLLFSNAK EGSVKVESLGIQNLKVFEKYIITPLGDKIKADFQPRENISLKTSKKYGLR
Mammalian Codon Optimized Sequences (AgeI-HA-	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CATGAGGATTCTGGGGTTTGACATTGGCATTAACAGCATCGGGTGGGCTTTTGTGGAGAACGACGAACTGAAGG ACTGCGGAGTGCGGATCTTCACAAAGGCCGAGAACCCAAAAAATAAGGAAAGCCTGGCACTGCCCCGGAGAAAT GCACGCAGCTCCAGGCGCCGACTGAAACGGAGAAAGGCCCGGCTGATCGCTATTAAGAGAATCCTGGCCAAAGA GCTGAAGCTGAACTACAAGGACTATGTCGCAGCTGATGGAGAGCTGCCAAAGGCCTACGAAGGATCCCTGGCAT CTGTGTACGAGCTGCGGTATAAGGCCCTGACACAGAACCTGGAAACTAAAGATCTGGCCAGAGTGATCCTGCAC

FIG. 13CC

NLS-Cas9-NLS-taa-EcoRI)	ATTGCTAAGCATAGGGGGTACATGAACAAGAACGAGAAGAAATCAAACGACGCTAAGAAAGGAAAGATCCTGA GCGCTCTGAAAAACAATGCACTGAAGCTGGAGAACTACCAGAGCGTGGGCGAATACTTCTACAAGGAGTTCTTT CAGAAATACAAGAAAAACACAAAGAACTTCATCAAGATCCGCAACACTAAGGATAATTACAACAATTGCGTGCT GTCTAGTGACCTGGAAAAAGAGCTGAAGCTGATCCTGGAAAAACAGAAGGAGTTCGGCTACAACTACTCTGAAG ATTTCATCAACGAGATTCTGAAGGTCGCCTTCTTTCAGCGGCCCCTGAAGGACTTCAGTCACCTGGTGGGGGCCT GCACTTTCTTTGAGGAAGAGAAAAGGGCCTGTAAGAACAGCTACTCTGCCTGGGAGTTTGTGGCTCTGACCAAGA TCATTAACGAGATCAAGAGCCTGGAGAAGATCAGCGGCGAAATTGTGCCAACCCAGACAATCAACGAGGTCCTG AATCTGATCCTGGACAAGGGGTCTATCACCTACAAGAAATTCAGAAGTTGTATCAATCTGCATGAGAGTATCAGC TTCAAGAGCCTGAAGTATGATAAAGAAAACGCCGAGAATGCTAAACTGATCGACTTCCGCAAGCTGGTGGAGTT TAAGAAAGCCCTGGGAGTCCACAGCCTGTCCCGGCAGGAACTGGATCAGATCTCCACTCATATCACCCTGATTAA GGACAACGTGAAGCTGAAAACCGTCCTGGAGAAATACAACCTGAGTAATGAACAGATCAACAATCTGCTGGAAA TTGAGTTCAACGATTATATCAACCTGAGCTTCAAGGCCCTGGGAATGATTCTGCCACTGATGCGCGAGGGCAAAC GATACGACGAGGCCTGCGAGATCGCCAATCTGAAACCTAAGACCGTGGACGAGAAGAAAGATTTCCTGCCAGCA TTTTGTGATTCCATTTTCGCCCACGAGCTGTCTAACCCCGTGGTCAATAGGGCTATCAGCGAATACCGCAAGGTGC TGAACGCACTGCTGAAGAAATATGGAAAGGTCCACAAAATTCATCTGGAGCTGGCTCGCGACGTGGGCCTGTCC AAGAAAGCACGAGAGAAGATCGAAAAAGAGCAGAAGGAAAACCAGGCCGTGAATGCATGGGCCCTGAAGGAAT GCGAGAATATTGGCCTGAAGGCCAGCGCAAAGAACATCCTGAAACTGAAGCTGTGGAAAGAACAGAAGGAGAT CTGTATCTACTCCGGAAATAAGATCTCTATTGAGCACCTGAAAGATGAAAAGGCCCTGGAGGTGGACCATATCTA CCCCTATTCTAGGAGTTTCGACGATTCTTTTATCAACAAAGTGCTGGTGTTCACCAAGGAAAATCAGGAGAAACT GAACAAGACACCTTTCGAGGCCTTTGGCAAGAATATTGAAAAATGGAGCAAGATCCAGACCCTGGCTCAGAACC TGCCATACAAGAAAAAGAATAAGATTCTGGACGAGAACTTCAAAGATAAGCAGCAGGAGGACTTTATCTCTCGA AATCTGAACGACACCCGGTATATCGCTACACTGATTGCAAAATACACAAAGGAGTATCTGAACTTCCTGCTGCTG AGCGAAAATGAGAACGCCAATCTGAAGAGTGGCGAAAAAGGGTCAAAGATCCACGTGCAGACTATTAGCGGGA TGCTGACCTCCGTCCTGAGGCACACATGGGGGTTTGACAAAAAGGATCGCAACAATCATCTGCACCATGCACTGG ATGCCATCATTGTGGCCTACAGTACAAATTCAATCATTAAGGCTTTCAGCGATTTCCGGAAAAACCAGGAGCTGC TGAAGGCCAGATTCTACGCTAAAGAACTGACTTCCGATAACTATAAACATCAGGTCAAGTTCTTTGAGCCTTTCA AGAGTTTTAGAGAAAAAATCCTGTCAAAGATCGACGAGATTTTCGTGTCCAAACCACCTCGAAAGCGAGCTAGG CGCGCACTGCACAAGGATACCTTTCATTCTGAGAACAAGATCATTGACAAGTGCAGCTACAACTCCAAGGAAGG CCTGCAGATTGCCCTGAGCTGTGGAAGAGTGAGGAAAATCGGCACTAAGTATGTCGAGAATGATACCATCGTGA GGGTCGACATTTTCAAAAAGCAGAACAAGTTTTACGCTATCCCAATCTACGCAATGGATTTTGCCCTGGGGATCC TGCCCAATAAGATCGTGATTACTGGAAAAGATAAGAACAATAACCCCAAACAGTGGCAGACCATTGACGAATCA TACGAGTTCTGCTTTAGCCTGTATAAGAATGACCTGATCCTGCTGCAGAAAAAGAACATGCAGGAACCTGAGTTC GCCTACTATAACGATTTTTCAATCAGCACATCAAGCATTTGTGTGGAGAAACACGACAACAAGTTCGAAAATCTG ACTAGCAACCAGAAGCTGCTGTTTTCCAATGCAAAAGAGGGCTCTGTGAAGGTCGAAAGTCTGGGGATCCAGAA CCTGAAAGTGTTCGAGAAGTACATCATTACCCCCCTGGGAGATAAAATTAAGGCTGACTTTCAGCCTCGAGAAAA CATCAGCCTGAAAACCAGTAAAAAGTATGGCCTGAGGTAAGAATTC

FIG. 13DD

Length of Protein (without NLS and HA tag)	1003
PAM	**NNGGGYA**
Spacer Length	29
Species: Streptococcus pyogenes SF370	
Sequencing Barcode	CCAGTTAG
ID	25
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU**
Cas9 Protein Sequence (wildtype, without NLS added)	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNR ICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL AHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRE DLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKA QVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIK ELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKS DNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD ENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKM IAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQ TGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPID FLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLF VEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTST

FIG. 13EE

	KEVLDATLIHQSITGLYETRIDLSQLGGD
Mammalian Codon Optimized Sequences (AgeI-HA-NLS-Cas9-NLS-taa-EcoRI)	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACA AGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCC CTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGAC GGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCAC AGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGA CGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAGG CCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACC TGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAA AACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGA AAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCT GACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACG ACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCG ACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCA AGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTAC AAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTT CTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGG ACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCC ATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTC CGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGA AACCATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGA CCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGT ATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAA AAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCA AGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACC ACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATC GTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCGA CGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAAC GGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTT CATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCG ATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAG GTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGA ACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCT GGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACC

FIG. 13FF

	TGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACCAT ATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGG CAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCA AGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCC GGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGAT GAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGT CCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACC TGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTAC AAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTT CTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGAT CGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGA GCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCC AAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCC CCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAA GAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGG CTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGA AGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTC CTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGA ACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACG CTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATC ATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAG AGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACG GATCGACCTGTCTCAGCTGGGAGGCGACAGCCCCAAGAAGAAGAGAAAGGTGGAGGCCAGCTAAGAATTC
Length of Protein (without NLS and HA tag)	1368
PAM	**NGG**
Spacer Length	30
Species: Streptococcus thermophilus LMD9 (CRISPR 1)	

FIG. 13GG

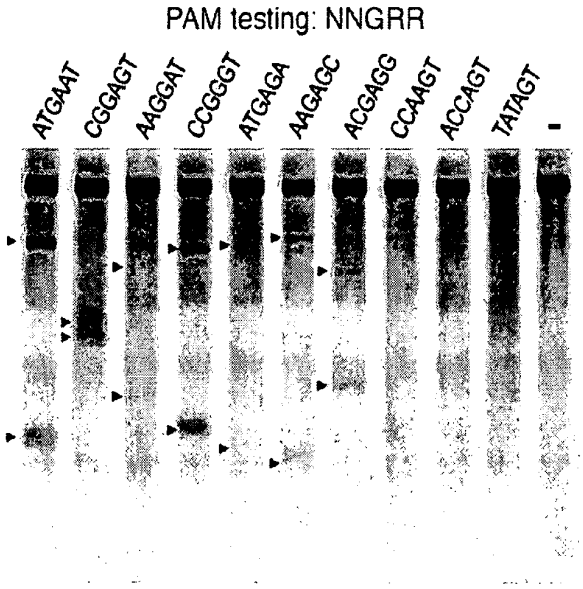
Sequencing Barcode	TCGGAATG
ID	26
Chimeric Guide RNA	**NNNNNNNNNNNNNNNNNNNGUUUUUGUACUCUCAAGAUUUAGAAACUUGCAGAAGCUACAAAGAUAAGG CUUCAUGCCGAAAUCAACACCCUGUCAUUUUAUGGCAGGGUGUUUU**
Cas9 Protein Sequence (wildtype, without NLS added)	MSDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTNRQGRRLARRKKHRRVRLNRLFEESGLITDFT KISINLNPYQLRVKGLTDELSNEELFIALKNMVKHRGISYLDDASDDGNSSVGDYAQIVKENSKQLETKTPGQIQLERY QTYGQLRGDFTVEKDGKKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLEILTGKRKYYHGPGNEKSRTDY GRYRTSGETLDNIFGILIGKCTFYPDEFRAAKASYTAQEFNLLNDLNNLTVPTETKKLSKEQKNQIINYVKNEKAMGPA KLFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETLDKLAYVLTLNTEREGIQEALEHEF ADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTE EIYNPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKAELP HSVFHGHKQLATKIRLWHQQGERCLYTGKTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKGQRTPY QALDSMDDAWSFRELKAFVRESKTLSNKKKEYLLTEEDISKFDVRKKFIERNLVDTRYASRVVLNALQEHFRAHKIDT KVSVVRGQFTSQLRRHWGIEKTRDTYHHHAVDALIIAASSQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKESV FKAPYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIY KKDKSKFLMYRHDPQTFEKVIEPILENYPNKQINEKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGN HIDITPKDSNNKVVLQSVSPWRADVYFNKTTGKYEILGLKYADLQFEKGTGTYKISQEKYNDIKKKEGVDSDSEFKFT LYKNDLLLVKDTETKEQQLFRFLSRTMPKQKHYVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLGKSNISIYKV RTDVLGNQHIIKNEGDKPKLM
Mammalian Codon Optimized Sequences (AgeI-HA-NLS-Cas9-NLS-taa-EcoRI)	ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTC CTCCGACCTGGTACTTGGACTGGATATTGGTATCGGTTCGGTGGGAGTCGGAATCCTCAACAAGGTCACGGGGGGA GATCATTCACAAGAACTCGCGGATCTTCCCCGCAGCTCAGGCTGAGAACAACTTGGTGCGGAGAACGAATAGGC AGGGCAGGCGACTGGCGAGGAGGAAGAAACACAGGAGAGTCCGATTGAACCGGCTGTTCGAGGAGTCCGGTTT GATCACCGACTTTACGAAAATCTCGATTAACCTTAATCCCTATCAGCTTCGGGTGAAAGGCCTGACAGACGAACT TTCGAATGAGGAACTTTTCATCGCGCTGAAAAACATGGTCAAGCACAGAGGGATTTCCTACCTCGATGACGCCTC GGATGACGGAAATTCCTCAGTAGGAGATTATGCACAGATCGTGAAAGAGAACTCAAAGCAACTGGAAACAAAG ACACCGGGGCAGATCCAACTTGAAAGATACCAGACATACGGACAGCTCAGAGGAGATTTTACGGTGGAGAAGGA CGGTAAAAAGCACAGACTCATTAACGTATTTCCCACGTCGGCGTACAGATCCGAAGCGCTCCGCATCCTTCAGAC TCAACAGGAGTTCAACCCGCAAATTACTGATGAGTTCATCAACCGCTATTTGGAAATCTTGACCGGAAAGCGCAA GTATTATCATGGGCCGGGTAATGAGAAATCCAGAACAGATTACGGCCGATACAGAACTTCGGGGGAAACCTTGG ATAACATCTTTGGTATTTTGATTGGAAAGTGCACCTTTTACCCGGACGAGTTTCGAGCGGCCAAGGCGTCATACA CAGCACAAGAGTTTAATCTCTTGAATGATTTGAACAACTTGACGGTCCCCACGGAGACAAAGAAGCTCTCCAAAG AGCAAAAGAACCAAATCATCAACTACGTCAAGAACGAGAAGGCTATGGGGCCAGCGAAGCTGTTCAAGTATATC

FIG. 13HH

	GCTAAACTTCTCAGCTGTGATGTGGCGGACATCAAAGGGTACCGAATCGACAAGTCGGGAAAAGCGGAAATTCA CACGTTTGAAGCATATCGAAAGATGAAAACGTTGGAAACACTGGACATTGAGCAGATGGACCGGGAAACGCTCG ACAAACTGGCATACGTGCTCACGTTGAATACTGAACGAGAGGGAATCCAAGAGGCCCTTGAACATGAGTTCGCC GATGGATCGTTCAGCCAGAAGCAGGTCGACGAACTTGTGCAATTCCGCAAGGCGAATAGCTCCATCTTCGGGAA GGGATGGCACAACTTTTCGGTCAAACTCATGATGGAGTTGATCCCAGAACTTTATGAGACTTCGGAGGAGCAAAT GACGATCTTGACGCGCTTGGGGAAACAGAAAACGACAAGCTCATCGAACAAAACTAAGTACATTGATGAGAAAT TGCTGACGGAAGAAATCTATAATCCGGTAGTAGCGAAATCGGTAAGACAAGCGATCAAAATCGTGAACGCGGCG ATCAAGGAATATGGTGACTTTGATAACATCGTAATTGAAATGGCTAGAGAGACGAACGAAGATGACGAGAAAAA GGCAATCCAGAAGATCCAGAAGGCCAACAAGGATGAAAAAGATGCAGCGATGCTTAAAGCGGCCAACCAATAC AATGGAAAGGCGGAGCTGCCCCATTCAGTGTTTCACGGTCATAAACAGTTGGCGACCAAGATCCGACTCTGGCAT CAGCAGGGTGAGCGGTGTCTCTACACCGGAAAGACTATCTCCATCCATGACTTGATTAACAATTCGAACCAGTTT GAAGTGGATCATATTCTGCCCCTGTCAATCACCTTTGACGACTCGCTTGCGAACAAGGTGCTCGTGTACGCAACG GCAAATCAGGAGAAAGGCCAGCGGACTCCGTATCAGGCGCTCGACTCAATGGACGATGCGTGGTCATTCCGGGA GCTGAAGGCGTTCGTACGCGAGAGCAAGACACTGAGCAACAAAAAGAAAGAGTATCTGCTGACAGAGGAGGAC ATCTCGAAATTCGATGTCAGGAAGAAGTTCATCGAGCGGAATCTTGTCGACACTCGCTACGCTTCCAGAGTAGTA CTGAACGCGCTCCAGGAACACTTTAGAGCGCACAAAATTGACACGAAGGTGTCAGTGGTGAGAGGGCAGTTCAC ATCCCAACTCCGCCGACATTGGGGCATCGAAAAGACGCGGGACACATATCACCATCATGCGGTGGACGCGCTGA TTATTGCCGCTTCGTCCCAGTTGAATCTCTGGAAAAAGCAGAAGAACACGCTGGTGTCGTATTCGGAGGATCAGC TTTTGGACATCGAAACCGGGGAGCTGATTTCCGACGATGAATACAAAGAATCGGTGTTTAAGGCACCATATCAGC ATTTCGTGGACACGCTGAAGAGCAAAGAGTTTGAGGACAGCATCCTCTTTTCGTACCAAGTGGACTCGAAGTTTA ATCGCAAGATTTCAGACGCCACAATCTACGCGACGAGGCAGGCGAAGGTGGGCAAAGATAAAGCAGATGAAAC CTACGTCCTTGGTAAAATCAAGGACATCTACACTCAGGACGGGTACGATGCGTTCATGAAAATCTACAAGAAGG ATAAGTCGAAGTTTCTCATGTACCGCCACGATCCACAGACTTTCGAAAAAGTCATTGAGCCTATTTTGGAGAACT ACCCTAACAAGCAAATCAACGAGAAAGGGAAAGAAGTCCCGTGCAACCCCTTTCTGAAGTACAAGGAAGAGCAC GGTTATATCCGCAAATACTCGAAGAAAGGAAATGGGCCTGAGATTAAGTCGCTTAAGTATTACGACTCAAAGTTG GGTAACCACATCGACATTACCCCGAAAGACTCCAACAACAAAGTCGTGTTGCAGTCCGTCTCGCCCTGGCGAGCA GATGTGTATTTTAATAAGACGACCGGCAAATATGAGATCCTTGGACTCAAATACGCAGACCTTCAATTCGAAAAG GGGACGGGCACTTATAAGATTTCACAAGAGAAGTACAACGACATCAAGAAAAAGGAAGGGGTCGATTCAGATTC GGAGTTCAAATTCACCCTCTACAAAAACGACCTCCTGCTTGTGAAGGACACAGAAACGAAGGAGCAGCAGCTCT TTCGGTTCCTCTCACGCACGATGCCCAAACAAAAACATTACGTCGAACTTAAACCTTACGATAAGCAAAAGTTTG AAGGGGGAGAGGCACTGATCAAAGTATTGGGTAACGTAGCCAATAGCGGACAGTGTAAGAAAGGGCTGGGAAA GTCCAATATCTCGATCTATAAAGTACGAACAGATGTATTGGGAAACCAGCATATCATCAAAAATGAGGGGGATA AACCCAAACTCGATTTCAGCCCCAAGAAGAAGAGAAAGGTGGAGGCCAGCTAAGAATTC
Length of Protein	1120

FIG. 13II

(without NLS and HA tag)	
PAM	**NNAGAA**
Spacer Length	30

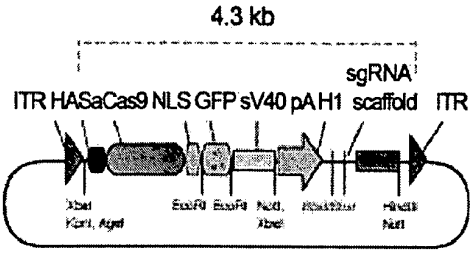


Spacer sequences for PAMs tested:

Spacer	PAM
GCCCGGGTGGAACTGGTAGCC	ATGAAAT
GTTGAAGATGAAGCCCAGAG	CGGAGT
GCTTCCGACGAGGTGGCCATC	AAGGAT
GCACCATCTCTCCGTGGTACC	CCGGGT
GGTGGAACTGGTAGCCATGA	ATGAGA
GCCATGAATGAGACCGACCCA	AAGAGC
GCATCCTCGTGGGCACTTCCG	ACGAGG
GCAGAGCGGAGTGCTGTTCTC	CCAAGT
GGTCGGTCTCATTCATGGCT	ACCAGT
GCAATAAAAGGTGCTATTGC	TATAGT

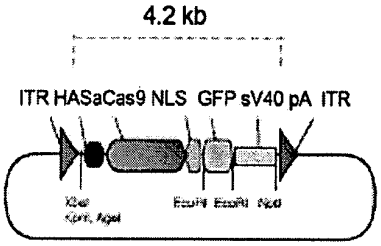
FIG. 14

a



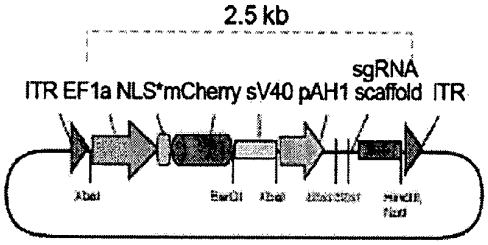
PC011:
HA-SaCas9-NLS-P2A-EGFP,
1 sgRNA cloning backbone

b



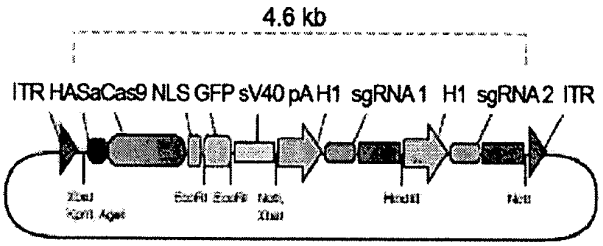
PC012:
HA-SaCas9-NLS-P2A-EGFP

c



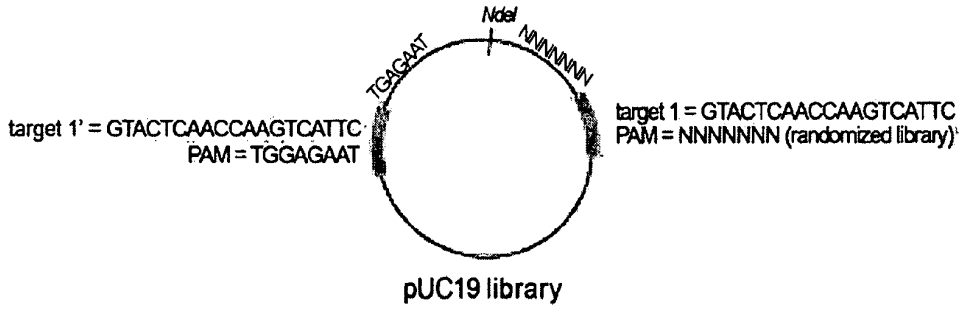
PC013:
*mCherry (modified to remove BbsI),
1 sgRNA cloning backbone

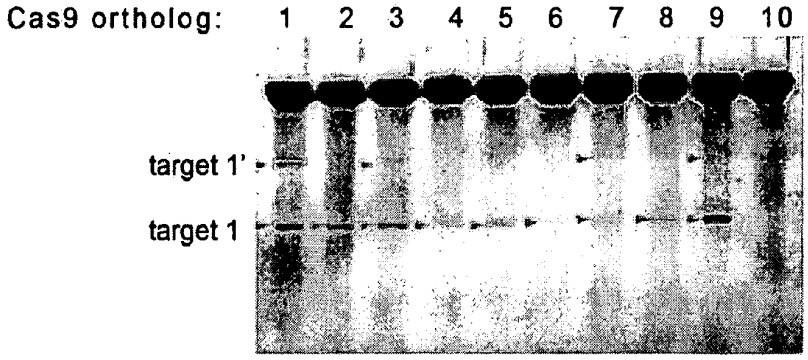
d

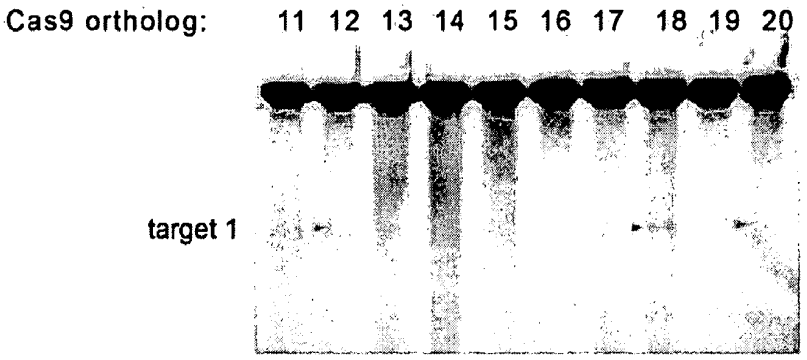


PC014:
HA-SaCas9-NLS-P2A-EGFP,
2 sgRNAs

FIG. 15


NdeI
TGAGAAT
NNNNNNN
target 1' = GTACTCAACCAAGTCATTC
PAM = TGGAGAAT
target 1 = GTACTCAACCAAGTCATTC
PAM = NNNNNNN (randomized library)
pUC19 library


Cas9 ortholog: 1 2 3 4 5 6 7 8 9 10
target 1'
target 1


Cas9 ortholog: 11 12 13 14 15 16 17 18 19 20
target 1

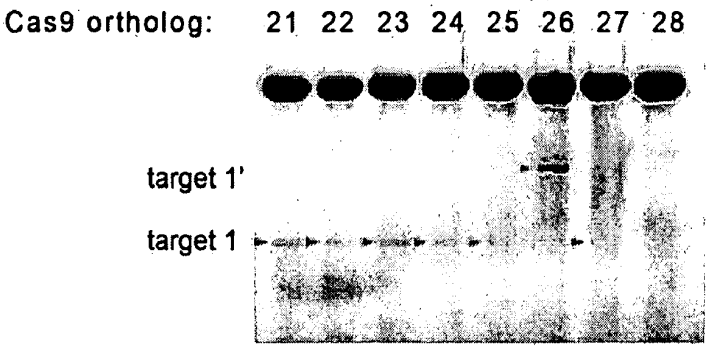

Cas9 ortholog: 21 22 23 24 25 26 27 28
target 1'
target 1

FIG. 16

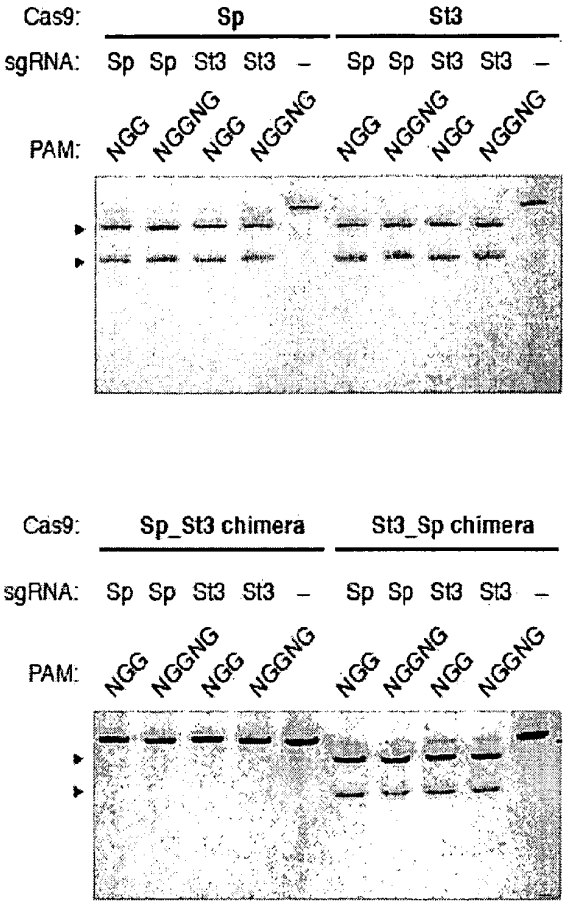

Cas9:
Sp
St3
sgRNA:
Sp Sp St3 St3 – Sp Sp St3 St3 –
PAM:
NGG NGGNG NGG NGGNG NGG NGGNG NGG NGGNG
Cas9:
Sp_St3 chimera
St3_Sp chimera
sgRNA:
Sp Sp St3 St3 – Sp Sp St3 St3 –
PAM:
NGG NGGNG NGG NGGNG NGG NGGNG NGG NGGNG

FIG. 17

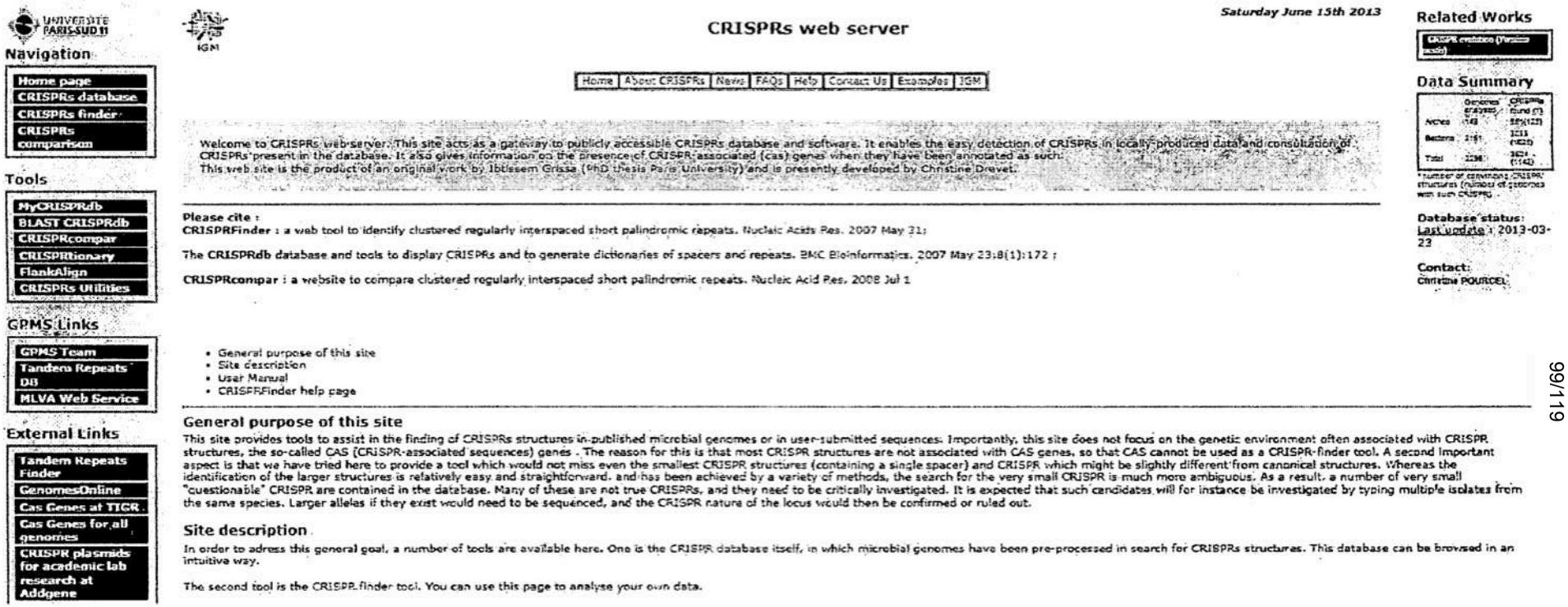

UNIVERSITE PARIS-SUD 11
IGM
Navigation
Home page
CRISPRs database
CRISPRs finder
CRISPRs comparison
Tools
MyCRISPRdb
BLAST CRISPRdb
CRISPRcompar
CRISPRtionary
FlankAlign
CRISPRs Utilities
GPMS Links
GPMS Team
Tandem Repeats DB
MLVA Web Service
External Links
Tandem Repeats Finder
GenomesOnline
Cas Genes at TIGR
Cas Genes for all genomes
CRISPR plasmids for academic lab research at Addgene
CRISPRs web server
Saturday June 15th 2013
Home
About CRISPRs
News
FAQs
Help
Contact Us
Examples
IGM
Welcome to CRISPRs web-server. This site acts as a gateway to publicly accessible CRISPRs database and software. It enables the easy detection of CRISPRs in locally-produced data and consultation of CRISPRs present in the database. It also gives information on the presence of CRISPR-associated (cas) genes when they have been annotated as such.
This web site is the product of an original work by Ibtissem Grissa (PhD thesis Paris University) and is presently developed by Christine Dreve.
Please cite :
CRISPRFinder : a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 2007 May 31;
The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinformatics. 2007 May 23;8(1):172 ;
CRISPRcompar : a website to compare clustered regularly interspaced short palindromic repeats. Nucleic Acid Res. 2008 Jul 1
General purpose of this site
Site description
User Manual
CRISPRFinder help page
General purpose of this site
This site provides tools to assist in the finding of CRISPRs structures in published microbial genomes or in user-submitted sequences. Importantly, this site does not focus on the genetic environment often associated with CRISPR structures, the so-called CAS (CRISPR-associated sequences) genes . The reason for this is that most CRISPR structures are not associated with CAS genes, so that CAS cannot be used as a CRISPR-finder tool. A second important aspect is that we have tried here to provide a tool which would not miss even the smallest CRISPR structures (containing a single spacer) and CRISPR which might be slightly different from canonical structures. Whereas the identification of the larger structures is relatively easy and straightforward, and has been achieved by a variety of methods, the search for the very small CRISPR is much more ambiguous. As a result, a number of very small "questionable" CRISPR are contained in the database. Many of these are not true CRISPRs, and they need to be critically investigated. It is expected that such candidates will for instance be investigated by typing multiple isolates from the same species. Larger alleles if they exist would need to be sequenced, and the CRISPR nature of the locus would then be confirmed or ruled out.
Site description
In order to adress this general goal, a number of tools are available here. One is the CRISPR database itself, in which microbial genomes have been pre-processed in search for CRISPRs structures. This database can be browsed in an intuitive way.
The second tool is the CRISPR-finder tool. You can use this page to analyse your own data.
Related Works
Data Summary
Archea
Bacteria
Total
Database status:
Last update : 2013-03-23
Contact:
Christine POURCEL

FIG. 18

FIG. 19A

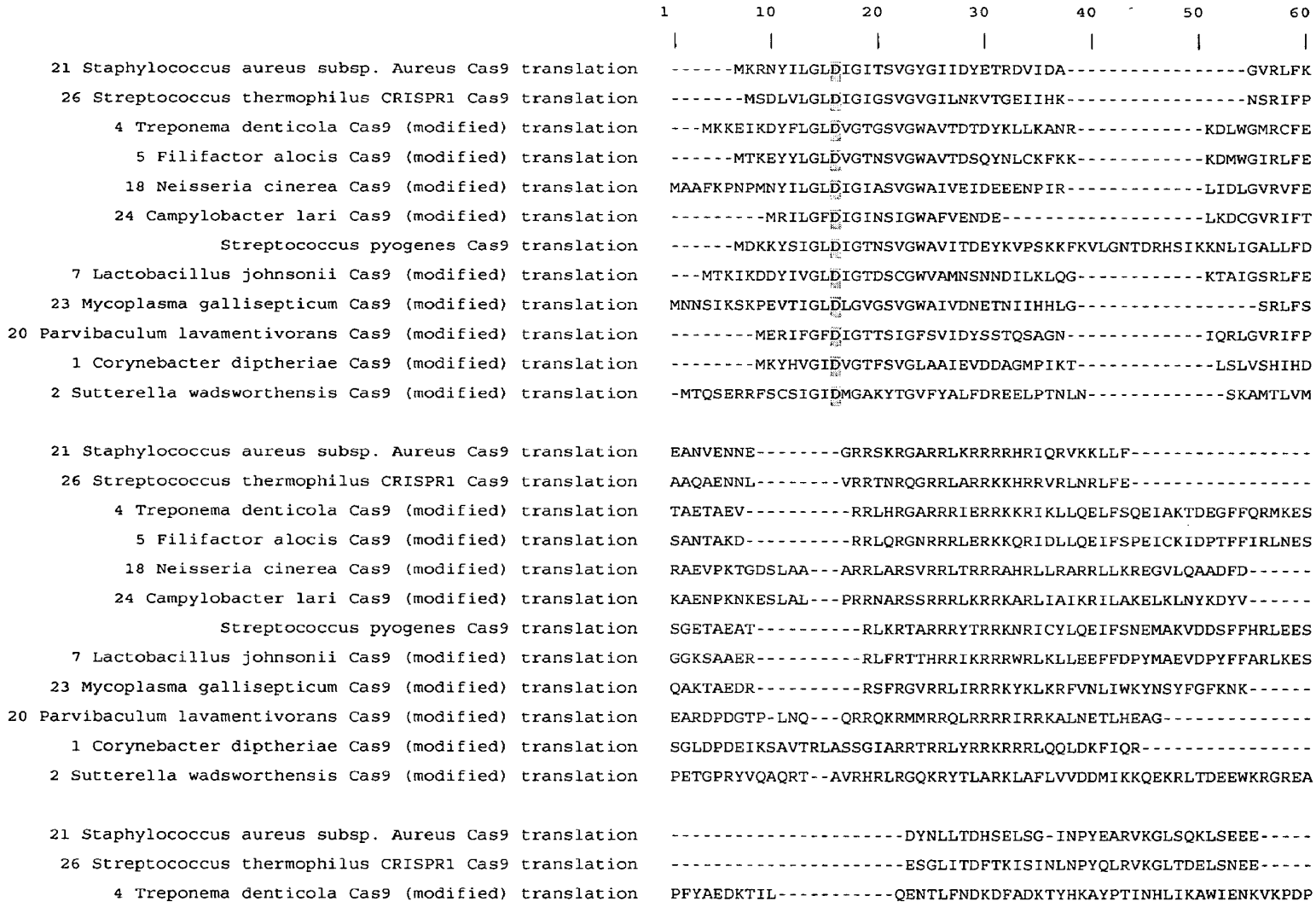

1 10 20 30 40 50 60
21 Staphylococcus aureus subsp. Aureus Cas9 translation ------MKRNYILGLDIGITSVGYGIIDYETRDVIDA----------GVRLFK
26 Streptococcus thermophilus CRISPR1 Cas9 translation -------MSDLVLGLDIGIGSVGVGILNKVTGEIIHK----------NSRIFP
4 Treponema denticola Cas9 (modified) translation ---MKKEIKDYFLGLDVGTGSVGWAVTDTDYKLLKANR----------KDLWGMRCFE
5 Filifactor alocis Cas9 (modified) translation ------MTKEYYLGLDVGTNSVGWAVTDSQYNLCKFKK----------KDMWGIRLFE
18 Neisseria cinerea Cas9 (modified) translation MAAFKPNPMNYILGLDIGIASVGWAIVEIDEEENPIR----------LIDLGVRVFE
24 Campylobacter lari Cas9 (modified) translation ---------MRILGFDIGINSIGWAFVENDE----------LKDCGVRIFT
Streptococcus pyogenes Cas9 translation ------MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFD
7 Lactobacillus johnsonii Cas9 (modified) translation ---MTKIKDDYIVGLDIGTDSCGWVAMNSNNDILKLQG----------KTAIGSRLFE
23 Mycoplasma gallisepticum Cas9 (modified) translation MNNSIKSKPEVTIGLDLGVGSVGWAIVDNETNIIHHLG----------SRLFS
20 Parvibaculum lavamentivorans Cas9 (modified) translation --------MERIFGFDIGTTSIGFSVIDYSSTQSAGN----------IQRLGVRIFP
1 Corynebacter diptheriae Cas9 (modified) translation --------MKYHVGIDVGTFSVGLAAIEVDDAGMPIKT----------LSLVSHIHD
2 Sutterella wadsworthensis Cas9 (modified) translation -MTQSERRFSCSIGIDMGAKYTGVFYALFDREELPTNLN----------SKAMTLVM
21 Staphylococcus aureus subsp. Aureus Cas9 translation EANVENNE--------GRRSKRGARRLKRRRRHRIQRVKKLLF----------
26 Streptococcus thermophilus CRISPR1 Cas9 translation AAQAENNL--------VRRTNRQGRRLARRKKHRRVRLNRLFE----------
4 Treponema denticola Cas9 (modified) translation TAETAEV----------RRLHRGARRRIERRKKRIKLLQELFSQEIAKTDEGFFQRMKES
5 Filifactor alocis Cas9 (modified) translation SANTAKD----------RRLQRGNRRRLERKKQRIDLLQEIFSPEICKIDPTFFIRLNES
18 Neisseria cinerea Cas9 (modified) translation RAEVPKTGDSLAA---ARRLARSVRRLTRRRAHRLLRARRLLKREGVLQAADFD------
24 Campylobacter lari Cas9 (modified) translation KAENPKNKESLAL---PRRNARSSRRRLKRRKARLIAIKRILAKELKLNYKDYV------
Streptococcus pyogenes Cas9 translation SGETAEAT----------RLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEES
7 Lactobacillus johnsonii Cas9 (modified) translation GGKSAAER----------RLFRTTHRRIKRRRWRLKLLEEFFDPYMAEVDPYFFARLKES
23 Mycoplasma gallisepticum Cas9 (modified) translation QAKTAEDR----------RSFRGVRRLIRRRKYKLKRFVNLIWKYNSYFGFKNK------
20 Parvibaculum lavamentivorans Cas9 (modified) translation EARDPDGTP-LNQ---QRRQKRMMRRQLRRRRIRRKALNETLHEAG----------
1 Corynebacter diptheriae Cas9 (modified) translation SGLDPDEIKSAVTRLASSGIARRTRRLYRRKRRRLQQLDKFIQR----------
2 Sutterella wadsworthensis Cas9 (modified) translation PETGPRYVQAQRT--AVRHRLRGQKRYTLARKLAFLVVDDMIKKQEKRLTDEEWKRGREA
21 Staphylococcus aureus subsp. Aureus Cas9 translation ----------DYNLLTDHSELSG-INPYEARVKGLSQKLSEEE-----
26 Streptococcus thermophilus CRISPR1 Cas9 translation ----------ESGLITDFTKISINLNPYQLRVKGLTDELSNEE-----
4 Treponema denticola Cas9 (modified) translation PFYAEDKTIL----------QENTLFNDKDFADKTYHKAYPTINHLIKAWIENKVKPDP

FIG. 19B

```
          5 Filifactor alocis Cas9 (modified) translation   RLHLEDKSND----------FKYPLFIEKDYSDIEYYKEFPTIFHLRKHLIESEEKQDI
         18 Neisseria cinerea Cas9 (modified) translation   ----------------------ENG-----LIKSLPNTPWQLRAAALDRKLTPLE-----
         24 Campylobacter lari Cas9 (modified) translation  ----------------------AADGELPKAYEGSLASVYELRYKALTQNLETKD-----
                    Streptococcus pyogenes Cas9 translation  FLVEEDKK------------HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL
          7 Lactobacillus johnsonii Cas9 (modified) translation  GLSPLDKRK-----------TVSSIVFPTSAEDKKFYDDYPTIYHLRYKLMTEDEKFDL
         23 Mycoplasma gallisepticum Cas9 (modified) translation  ----------------------EDILNNYQEQQKLHNTVLNLKSEALNAKIDPKA-----
     20 Parvibaculum lavamentivorans Cas9 (modified) translation  ----------------------FLPAYGSADWPVVMADEPYELRRRGLEEGLSAYE-----
          1 Corynebacter diptheriae Cas9 (modified) translation  ----------------------QGWPVIELEDYSDPLYPWKVRAELAASYIADEKERGE-
          2 Sutterella wadsworthensis Cas9 (modified) translation  LSGLLKRRGYSRPNADGEDLTPLENVRADVFAAHPAFSTYFSEVRSLAEQWEEFTANISN

         21 Staphylococcus aureus subsp. Aureus Cas9 translation  --FSAALLHLAKRRGVHNVN----------EVEEDTGNELSTKEQIS
         26 Streptococcus thermophilus CRISPR1 Cas9 translation  --LFIALKNMVKHRGISYLD----------DASDDGNSSVGDYAQIV
              4 Treponema denticola Cas9 (modified) translation  RLLYLACHNIIKKRGHFLFEGDFDSEN----QFDTSIQAL-FEYLREDMEVDIDADSQK-
               5 Filifactor alocis Cas9 (modified) translation  RLIYLALHNIIKTRGHFLIDGDLQSAK----QLRPILDTF-LLSLQEEQNLSVSLSENQK
              18 Neisseria cinerea Cas9 (modified) translation  --WSAVLLHLIKHRGYLSQR----------KNEGETADKELGALLKGV
              24 Campylobacter lari Cas9 (modified) translation  --LARVILHIAKHRGYMN----------KNEKKSNDAKKGKILSAL
                    Streptococcus pyogenes Cas9 translation  RLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQT-YNQLFEENPINASGVDAKA
          7 Lactobacillus johnsonii Cas9 (modified) translation  REVYLAIHHIIKYRGNFLYNTSVKDFKASKIDVKSSIEKLNELYENLGLDLNVEFNISNT
         23 Mycoplasma gallisepticum Cas9 (modified) translation  --LSWILHDYLKNRGHFYEDN----------RDFNVYPTKELAKYFDKY
     20 Parvibaculum lavamentivorans Cas9 (modified) translation  --FGRAIYHLAQHRHFKGRE----------LEESDTPDPDVDDEKEAA
          1 Corynebacter diptheriae Cas9 (modified) translation  -KLSVALRHIARHRGWRNPYAKVSS----------LYLPDGPSDAFKAIREEI
          2 Sutterella wadsworthensis Cas9 (modified) translation  VEKFLGDPNIPADKEFIEFAVAEG----------LIDKTEKKAYQSALSTL

         21 Staphylococcus aureus subsp. Aureus Cas9 translation  R-NSKALEEKYVA----------
         26 Streptococcus thermophilus CRISPR1 Cas9 translation  KENSKQLETKTPG----------
              4 Treponema denticola Cas9 (modified) translation  --VKEILKDSSLKNSEKQSRLNKILGLK--PSDKQKKAITNLIS-------GNKINFADL
               5 Filifactor alocis Cas9 (modified) translation  DEYEEILKNRSIAKSEKVKKLKNLFEISDELEKEEKKAQSAVIENFCKFIVGNKGDVCKF
              18 Neisseria cinerea Cas9 (modified) translation  ADNTHALQTGD----------
              24 Campylobacter lari Cas9 (modified) translation  KNNALKLENYQSVG----------
                    Streptococcus pyogenes Cas9 translation  ILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN----------
          7 Lactobacillus johnsonii Cas9 (modified) translation  AEIEKVLKDKQIFKRDKVKKIAELFAIKTDNKEQSKRIKDISKQVANAVLGYKTRFDTIA
         23 Mycoplasma gallisepticum Cas9 (modified) translation  GYYKGIIDSKEDND----------
```

FIG. 19C

```
20 Parvibaculum lavamentivorans Cas9 (modified) translation   NERAATLKALKNEQ----------------------------------------------
      1 Corynebacter diptheriae Cas9 (modified) translation   KRASGQPVPETATVG---------------------------------------------
     2 Sutterella wadsworthensis Cas9 (modified) translation   RANANVLTGLRQMG----------------------------------------------

     21 Staphylococcus aureus subsp. Aureus Cas9 translation   ----------------------------------------------------------
     26 Streptococcus thermophilus CRISPR1 Cas9 translation   ----------------------------------------------------------
          4 Treponema denticola Cas9 (modified) translation   YDNPDLKDAEKNSISFSKDDFD-ALSDDLASILGDSFELLLKAKAVYNCSVLSKVIGD--
            5 Filifactor alocis Cas9 (modified) translation   LR-VSKEELEIDSFSFSEGKYEDDIVKNLEEKVPEKVYLFEQMKAMYDWNILVDILET--
           18 Neisseria cinerea Cas9 (modified) translation   ----------------------------------------------------------
          24 Campylobacter lari Cas9 (modified) translation   ----------------------------------------------------------
                    Streptococcus pyogenes Cas9 translation   -----FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNT
      7 Lactobacillus johnsonii Cas9 (modified) translation   LKEISKDELSDWNFKLSDIDADSKFEALMGNLDENEQAILLTIKELFNEVTLNGIVED--
     23 Mycoplasma gallisepticum Cas9 (modified) translation   ----------------------------------------------------------
20 Parvibaculum lavamentivorans Cas9 (modified) translation   ----------------------------------------------------------
      1 Corynebacter diptheriae Cas9 (modified) translation   ----------------------------------------------------------
     2 Sutterella wadsworthensis Cas9 (modified) translation   ----------------------------------------------------------

     21 Staphylococcus aureus subsp. Aureus Cas9 translation   -----ELQLERLKKDGEVRG-----------------------------------------
     26 Streptococcus thermophilus CRISPR1 Cas9 translation   -----QIQLERYQTYGQLRGDFTVEK-------------------------------
          4 Treponema denticola Cas9 (modified) translation   ---EQYLSFAKVKIYEKHKTDLTKLKNVIKKHFPK-DYKKVFGYNKNEKNNNNYSGYVGV
            5 Filifactor alocis Cas9 (modified) translation   ---EEYISFAKVKQYEKHKTNLRLLRDIILKYCTKDEYNRMFN---DEKEAGSYTAYVGK
           18 Neisseria cinerea Cas9 (modified) translation   ------FRTPAELALNKFEKESGHIR---------------------------------
          24 Campylobacter lari Cas9 (modified) translation   ----EYFYKEFFQKYKKNTKNFIKIR-------------------------------
                    Streptococcus pyogenes Cas9 translation   EITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD----------------
      7 Lactobacillus johnsonii Cas9 (modified) translation   ---GNTLSESMINKYNDHRDDLKLLKEVIENHIDRKKAKELALAYDLYVNN-----RHGQ
     23 Mycoplasma gallisepticum Cas9 (modified) translation   ----NKLEEELTKYKFSNKHWLEEVKKVLSNQTG-------------------------
20 Parvibaculum lavamentivorans Cas9 (modified) translation   ----TTLGAWLARRPPSDRKR--------------------------------------
      1 Corynebacter diptheriae Cas9 (modified) translation   -----QMVTLCELGTLKLRGE-------------------------------------
     2 Sutterella wadsworthensis Cas9 (modified) translation   ----HKPRSEYFKAIEADLKKDSRLAKINEAFGGAERLARLLG----------------

     21 Staphylococcus aureus subsp. Aureus Cas9 translation   ----------------------------------------------------------S
     26 Streptococcus thermophilus CRISPR1 Cas9 translation   ----------------------------------------------------DGKKHRL
```

FIG. 19D

```
         4 Treponema denticola Cas9 (modified) translation   CKTKSKKLIINNSVNQEDFYKFLKTILSAKSEIKEVNDILT----EIETGTFLPKQISKS
           5 Filifactor alocis Cas9 (modified) translation   LKKNNKKYWIEKKRNPEEFYKSLGKLLDKIEPLKEDLEVLTMMIEECKNHTLLPIQKNKD
          18 Neisseria cinerea Cas9 (modified) translation   ----------------------------------------------------NQRGDY
          24 Campylobacter lari Cas9 (modified) translation   ----------------------------------------------------NTKDNY
                       Streptococcus pyogenes Cas9 translation   QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN-------REDLLRKQRTFD
           7 Lactobacillus johnsonii Cas9 (modified) translation   LLQAKKKLGKIKPRSKEDFYKVVNKNLDDSRASKEIKKKIELDS-------FMPKQRTNA
      23 Mycoplasma gallisepticum Cas9 (modified) translation   ----------------------------------------------------LPEKFKEEY
20 Parvibaculum lavamentivorans Cas9 (modified) translation   ------------------------------------------------------------
           1 Corynebacter diptheriae Cas9 (modified) translation   ------------------------------------------------------------
           2 Sutterella wadsworthensis Cas9 (modified) translation   ----------------------------------------NLSNLQLRAERWYFNAP

      21 Staphylococcus aureus subsp. Aureus Cas9 translation   INRFKTSDYVKEAKQLLKVQKAYH-QLDQS---------FIDTYIDLLETRRTYYEGPGE
          26 Streptococcus thermophilus CRISPR1 Cas9 translation   INVFPTSAYRSEALRILQTQQEFNPQITDE---------FINRYLEILTGKRKYYHGPGN
         4 Treponema denticola Cas9 (modified) translation   NAEIPYQLRKMELEKILSNAEKHFSFLKQKD---EKGLSHSEKIIMLLTFKIPYYIGPIN
           5 Filifactor alocis Cas9 (modified) translation   NGVIPHQVHEVELKKILENAKKYYSFLTETD---KDGYSVVQKIESIFRFRIPYYVGPLS
          18 Neisseria cinerea Cas9 (modified) translation   SHTFNRKDLQAELNLLFEKQKEFGNPHVSD---------GLKEGIETLLMTQRPAL----
          24 Campylobacter lari Cas9 (modified) translation   NNCVLSSDLEKELKLILEKQKEFGYNYSED---------FINEILKVAFFQR--------
                       Streptococcus pyogenes Cas9 translation   NGSIPHQIHLGELHAILRRQEDFYPFLKDN----------REKIEKILTFRIPYYVGPLA
           7 Lactobacillus johnsonii Cas9 (modified) translation   NGVIPYQLQQLELDKIIENQSKYYPFLKEINPVSSHLKEAPYKLDELIRFRVPYYVGPLI
      23 Mycoplasma gallisepticum Cas9 (modified) translation   ESLFSYVRNYSEGPGSINSVSPYGIYHLDEK-----EGKVVQKYNNIWDKTIGKCNIFPD
20 Parvibaculum lavamentivorans Cas9 (modified) translation   GIHAHRNVVAEEFERLWEVQSKFHPALKSEE--------MRARISDTIFAQRPVFWR---
           1 Corynebacter diptheriae Cas9 (modified) translation   GGVLSARLQQSDYAREIQEICRMQEIGQELY----------RKIIDVVFAAES-------
           2 Sutterella wadsworthensis Cas9 (modified) translation   DIMKDRGWEPDRFKKTLVRAFKFFHPAKDQNKQHLELIKQIENSEDIIETLCTLDPNRTI

      21 Staphylococcus aureus subsp. Aureus Cas9 translation   ---GSPFG-------------------------------WKDIKEWYEMLMGHCTYFPEE
          26 Streptococcus thermophilus CRISPR1 Cas9 translation   EKSRTDYGR-------------------------YRTSGETLDNIFGILIGKCTFYPDE
         4 Treponema denticola Cas9 (modified) translation   DNHKKFFPDRCW----VVKKEKSPSGKTTPWNFFDHIDKEKTAEAFITSRTNFCTYLVGE
           5 Filifactor alocis Cas9 (modified) translation   TRHQEKGSN-VW----MVRKP-GREDRIYPWNMEEIIDFEKSNENFITRMTNKCTYLIGE
          18 Neisseria cinerea Cas9 (modified) translation   ----------------------------------------SGDAVQKMLGHCTFEPTE
          24 Campylobacter lari Cas9 (modified) translation   ----------------------------------------PLKDFSHLVGACTFFEEE
                       Streptococcus pyogenes Cas9 translation   RGNSRFAWMTR-----------KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE
           7 Lactobacillus johnsonii Cas9 (modified) translation   SPNESTKDIQTKKNQNFAWMIRKEEGRITPWNFDQKVDRIESANKFIKRMTTKDTYLFGE
```

FIG. 19E

```
    23 Mycoplasma gallisepticum Cas9 (modified) translation   EYRAPKNSP-------------------------IAMIFNEINELSTIRSYSIYLTGW
20 Parvibaculum lavamentivorans Cas9 (modified) translation   ---------------------------------------------KNTLGECRFMPGE
     1 Corynebacter diptheriae Cas9 (modified) translation   ------------------------------------------PKGSASSRVGKDPLQPGK
     2 Sutterella wadsworthensis Cas9 (modified) translation   PPYEDQNNRRPP-----LDQTLLLSPEKLTRQYGEIWKTWSARLTSAEPTLAPAAEILER

    21 Staphylococcus aureus subsp. Aureus Cas9 translation   LRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKK---PTLKQIA
    26 Streptococcus thermophilus CRISPR1 Cas9 translation   FRAAKASYTAQEFNLLNDLNNLTVP-TETKKLSKEQKNQIINYVKNEKAMGPAKLFKYIA
     4 Treponema denticola Cas9 (modified) translation   SVLPKSSLLYSEYTVLNEINNLQIIIDGKNICDIKLKQKIYEDLFKKYKK---ITQKQIS
     5 Filifactor alocis Cas9 (modified) translation   DVLPKHSLLYSKYMVLNELN--NVKVRGKKLP-TSLKQKVFEDLFENKSK---VTGKNLL
    18 Neisseria cinerea Cas9 (modified) translation   PKAAKNTYTAERFVWLTKLN--NLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQARK
    24 Campylobacter lari Cas9 (modified) translation   KRACKNSYSAWEFVALTKIIN-EIKSLEKISGEIVPTQTINEVLNLILDKGSITYKKFRS
            Streptococcus pyogenes Cas9 translation   KVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK
     7 Lactobacillus johnsonii Cas9 (modified) translation   DVLPANSLLYQKFTVLNELNN---IRINGKRISVDLKQEIYENLFKKHTT----VTVKKL
    23 Mycoplasma gallisepticum Cas9 (modified) translation   FINQEFKKAYLNKLLDLLIKTNGEKPIDARQFKKLREETIAESIGKETLKDVENEEKLEK
20 Parvibaculum lavamentivorans Cas9 (modified) translation   PLCPKGSWLSQQRRMLEKLNNLAIAGGNARPLDAEERDAILSKLQQQASMSWPGVRSALK
     1 Corynebacter diptheriae Cas9 (modified) translation   NRALKASDAFQRYRIAALIGNLRVRVDGEKRILSVEEKNLVFDHLVNLTPKKEPEWVTIA
     2 Sutterella wadsworthensis Cas9 (modified) translation   STDRKSRVAVNGHEPLPTLAYQLSYALQRAFDRSKALDPYALRALAAGSKSNKLTSARTA

    21 Staphylococcus aureus subsp. Aureus Cas9 translation   KEILVNEEDIKG---YRVT--STGKPEFTNLKVYHDIKD--ITA--RKEIIENAELLDQI
    26 Streptococcus thermophilus CRISPR1 Cas9 translation   KLLSCDVADIKG---YRID--KSGKAEIHTFEAYRKMKT--LET--LDIEQMDRETLDKL
     4 Treponema denticola Cas9 (modified) translation   TFIKHEGICNKT---DEVIILGIDKECTSSLKSYIELKN-IFGK--QVDEISTKNMLEEI
     5 Filifactor alocis Cas9 (modified) translation   EYLQIQDKDIQI---DDLS--GFDKDFKTSLKSYLDFKKQIFGE--EIEKESIQNMIEDI
    18 Neisseria cinerea Cas9 (modified) translation   LLDLDDTAFFKG---LRYGKDNAEASTLMEMKAYHAISR-ALEK--EGLKDKKSPLNLSP
    24 Campylobacter lari Cas9 (modified) translation   CINLHESISFKS---LKYDKENAENAKLIDFRKLVEFKK-ALG----------VHSLSR
            Streptococcus pyogenes Cas9 translation   EDYFKKIECFDS---VEIS--GVEDRFNASLGTYHDLLKIIKDKD-FLDNEENEDILEDI
     7 Lactobacillus johnsonii Cas9 (modified) translation   ENYLKENHNLVK---VEIKGLADEKKFNSGLTTYNRFKNLNIFDN-QIDDLKYRNDFEKI
    23 Mycoplasma gallisepticum Cas9 (modified) translation   EDHKWKLKGLKLNTNGKIQYNDLSSLAKFVHKLKQHLKLDFLLED-QYATLDKINFLQSL
20 Parvibaculum lavamentivorans Cas9 (modified) translation   ALYKQRGEPGAEKS-LKFNLELGGESKLLGNALEAKLADMFGPDWPAHPRKQEIRHAVHE
     1 Corynebacter diptheriae Cas9 (modified) translation   EILGIDRGQLIG------TATMTDDGERAGARPPTHDTNRSIVNSRIAPLVDWWKTASAL
     2 Sutterella wadsworthensis Cas9 (modified) translation   LENCIGGQNVKT--FLDCARRYYREADDAKVGLWFDNADGLLERSDLHPPMKKKILPLLV

    21 Staphylococcus aureus subsp. Aureus Cas9 translation   AKILTIYQ--SSEDIQEELTN--LNSELTQEEIEQISNLKGYT----------------
```

FIG. 19F

```
26 Streptococcus thermophilus CRISPR1 Cas9 translation   AYVLTLNT--EREGIQEALEHEFADGSFSQKQVDELVQFRKANS----------
 4 Treponema denticola Cas9 (modified) translation       IRWATIYDEGEGKTILKTKIKAEYGKYCSDEQIKKILNLKFSGWGRLSRK----------
 5 Filifactor alocis Cas9 (modified) translation         IKWITIY--GNDKEMLKRVIRANYSNQLTEEQMKKITGFQYSGWGNFSKM----------
18 Neisseria cinerea Cas9 (modified) translation         ELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKF----------
24 Campylobacter lari Cas9 (modified) translation        QELDQISTHITLIKDNVKLKTVLEKYNLSNEQINNLLEIEFNDY----------
          Streptococcus pyogenes Cas9 translation        VLTLTLFE----DREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRK----------
 7 Lactobacillus johnsonii Cas9 (modified) translation   IEWSTIFE-----DKSIYKEKLRSIDWLNEKQINALSNIRLQGWG----------
23 Mycoplasma gallisepticum Cas9 (modified) translation  FVYLGKHLRYSNRVDSANLKEFSDSNKLFERILQKQKDGLFKLFEQTDKDD---------
20 Parvibaculum lavamentivorans Cas9 (modified) translation RLWAADYGETPDKKRVIILSEKDRKAHREAAANSFVADFGITGEQAAQLQA---------
 1 Corynebacter diptheriae Cas9 (modified) translation   EQHAMVKALSNAEVDDFDSPEGAKVQAFFADLDDDVHAKLDSLHLP----------
 2 Sutterella wadsworthensis Cas9 (modified) translation ANILQTDETTGQKFLDEIWRKQIKGRETVASRCARIETVRKSFGGGFNIAYNTAQYREVN

21 Staphylococcus aureus subsp. Aureus Cas9 translation   ------------GTHNLSLKAINLILDELWHTND------------NQIAIFN-------
26 Streptococcus thermophilus CRISPR1 Cas9 translation   -------SIFGKGWHNFSVKLMMELIPELYETSE------------EQMTILT-------
 4 Treponema denticola Cas9 (modified) translation       -------FLETVTSEMPGFSEPVNIITAMRETQN------------NLMELLSSEFTFTE
 5 Filifactor alocis Cas9 (modified) translation         -------FLKGISGSDVSTGETFDIITAMWETDN------------NLMQILSKKFTFMD
18 Neisseria cinerea Cas9 (modified) translation         -----------VQISLKALRRIVPLMEQGNRYDE------------ACTEIYG-------
24 Campylobacter lari Cas9 (modified) translation        -----------INLSFKALGMILPLMREGKRYDE------------ACEIANLK------
          Streptococcus pyogenes Cas9 translation        -------LINGIRDKQSGKTILDFLKSDGFANRN------------FMQLIHDDS----L
 7 Lactobacillus johnsonii Cas9 (modified) translation   -------RLSKKLLAQLHDHNGQTIIEQLWDSQN------------NFMQIVTQAD----
23 Mycoplasma gallisepticum Cas9 (modified) translation  -------EKILAQTHSLSTKAMLLAITRMTNLDNDEDNQKNNDKGWNFEAIKNFDQKFID
20 Parvibaculum lavamentivorans Cas9 (modified) translation -------LKLPTGWEPYSIPALNLFLAELEKGERFG-----ALVNGPDWEGWRRTNFPHR
 1 Corynebacter diptheriae Cas9 (modified) translation   -----------VGRAAYSEDTLVRLTRRMLSDGVD------------------------
 2 Sutterella wadsworthensis Cas9 (modified) translation KLPRNAQDKELLTIRDRVAETADFIAANLGLSDEQKRKFANPFSLAQFYTLIETEVSGFS

21 Staphylococcus aureus subsp. Aureus Cas9 translation   RLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYG---LPNDIII
26 Streptococcus thermophilus CRISPR1 Cas9 translation   RLGKQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIKIVNAAIKEYG---DFDNIVI
 4 Treponema denticola Cas9 (modified) translation       NIKKINSGFEDAEKQFSYDGLVKPLFLSPSVKKMLWQTLKLVKEISHITQA--PPKKIFI
 5 Filifactor alocis Cas9 (modified) translation         NVEDFNSGKVGKIDKITYDSTVKEMFLSPENKRAVWQTIQVAEEIKKVMGC--EPKKIFI
18 Neisseria cinerea Cas9 (modified) translation         ----DHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYG---SPARIHI
24 Campylobacter lari Cas9 (modified) translation        ---PKTVDEKKDFLPAFCDSIFAHELSNPVVNRAISEYRKVLNALLKKYG---KVHKIHL
          Streptococcus pyogenes Cas9 translation        TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRH-KPENIVI
```

FIG. 19G

```
  7 Lactobacillus johnsonii Cas9 (modified) translation   FKDAIAKANQNLLVATSVEDILNNAYTSPANKKAIRQVIKVVDDIVKAASGK-VPKQIAI
 23 Mycoplasma gallisepticum Cas9 (modified) translation   ITKKNNNLSLKQNKRYLDDRFINDAILSPGVKRILREATKVFNAILKQFSEEYDVTKVVI
20 Parvibaculum lavamentivorans Cas9 (modified) translation   NQPTGEILDKLPSPASKEERERISQLRNPTVVRTQNELRKVVNNLIGLYG---KPDRIRI
  1 Corynebacter diptheriae Cas9 (modified) translation   LYTARLQEFGIEPSWTPPTPRIGEPVGNPAVDRVLKTVSRWLESATKTWG---APERVII
  2 Sutterella wadsworthensis Cas9 (modified) translation   ATTLAVHLENAWRMTIKDAVINGETVRAAQCSRLPAETARPFDGLVRRLVDR-QAWEIAK

 21 Staphylococcus aureus subsp. Aureus Cas9 translation   ELAR----EKNSKD-AQKMINEMQKRNRQTNERIEEIIRTTGKEN-----------AKY
 26 Streptococcus thermophilus CRISPR1 Cas9 translation   EMAR----ETNEDD-EKKAIQKIQKANK--DEKDAAMLKAANQYNGKAELPHSVFHGHKQ
  4 Treponema denticola Cas9 (modified) translation   EMAKGA--ELEPAR-TKTRLKILQDLYNNCKNDADAFSSEIKDLSGKIENE----DNLRL
  5 Filifactor alocis Cas9 (modified) translation   EMARGG--EKVKKR-TKSRKAQLLELYAACEEDCRELIKEIEDR----------DERDF
 18 Neisseria cinerea Cas9 (modified) translation   ETAREV--GKSFKD-RKEIEKRQEENRKDREKSAAKFREYFPNFVG---------EPKSK
 24 Campylobacter lari Cas9 (modified) translation   ELARDV--GLSKKA-REKIEKEQKEN----QAVNAWALKECENIGL---------KASAK
      Streptococcus pyogenes Cas9 translation   EMAREN--QTTQKG-QKNSRERMKRIEEGIKELGSQILKEHPVEN-----------TQL
  7 Lactobacillus johnsonii Cas9 (modified) translation   EFTRDA--DENPKR-SQTRGSKLQKVYKDLSTELASKTIAEELNEAIK--------DKKL
 23 Mycoplasma gallisepticum Cas9 (modified) translation   ELARELSEEKELEN-TKNYKKLIKKNGDKISEGLKALGISEDEIKD--------ILKSPT
20 Parvibaculum lavamentivorans Cas9 (modified) translation   EVGRDVG--KSKRE-REEIQSGIRRNEKQRKKATEDLIKNGIANPS-----------RD
  1 Corynebacter diptheriae Cas9 (modified) translation   EHVREG--FVTEKR-AREMDGDMRRRAARNAKLFQEMQEKLNVQGK----------PSRA
  2 Sutterella wadsworthensis Cas9 (modified) translation   RVSTDIQSKVDFSNGIVDVSIFVEENKFEFSASVADLKKNKRVKDKMLS--EAEKLETRW

 21 Staphylococcus aureus subsp. Aureus Cas9 translation   LIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNS----FNNKVLV
 26 Streptococcus thermophilus CRISPR1 Cas9 translation   LATKIRLWHQQGERCLYTGKTISIHDLINNSNQFEVDHILPLSITFDDS----LANKVLV
  4 Treponema denticola Cas9 (modified) translation   RSDKLYLYYTQLGKCMYCGKPIEIGHVFD-TSNYDIDHIYPQSKIKDDS----ISNRVLV
  5 Filifactor alocis Cas9 (modified) translation   NSMKLFLYYTQFGKCMYSGDDIDINELIRGNSKWDRDHIYPQSKIKDDS----IDNLVLV
 18 Neisseria cinerea Cas9 (modified) translation   DILKLRLYEQQHGKCLYSGKEINLGRLNE-KGYVEIDHALPFSRTWDDS----FNNKVLA
 24 Campylobacter lari Cas9 (modified) translation   NILKLKLWKEQKEICIYSGNKISIEHLKD-EKALEVDHIYPYSRSFDDS----FINKVLV
      Streptococcus pyogenes Cas9 translation   QNEKLYLYYLQNGRDMYVDQELDINRLSD----YDVDHIVPQSFLKDDS----IDNKVLT
  7 Lactobacillus johnsonii Cas9 (modified) translation   VQDKYYLYFMQLGRDAYTGEPINIDEIQK----YDIDHILPQSFIKDDA----LDNRVLV
 23 Mycoplasma gallisepticum Cas9 (modified) translation   KSYKFLLWLQQDHIDPYSLKEIAFDDIFTKTEKFEIDHIIPYSISFDDS----SSNKLLV
20 Parvibaculum lavamentivorans Cas9 (modified) translation   DVEKWILWKEGQERCPYTGDQIGFNALPR-EGRYEVEHIWPRSRSFDNS----PRNKTLC
  1 Corynebacter diptheriae Cas9 (modified) translation   DLWRYQSVQRQNCQCAYCGSPITFS-------NSEMDHIVPRAGQGSTN----TRENLVA
  2 Sutterella wadsworthensis Cas9 (modified) translation   LIKNERIKKASRGTCPYTGDRLAEGG--------EIDHILPRSLIKDARGIVFNAEPNLI
```

FIG. 19H

```
21 Staphylococcus aureus subsp. Aureus Cas9 translation   KQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLA-KGKGRISKTKKEYLLEERDINR
26 Streptococcus thermophilus CRISPR1 Cas9 translation    YATANQEKGQRTPYQALDSMDDAWSFRELKAFVR-----ESKTLSNKKKEYLLTEEDISK
 4 Treponema denticola Cas9 (modified) translation        CSSCNKNKEDKYPLKSEIQSKQRGFWNFLQ---------RNNFISLEKLNRLTRATPISD
 5 Filifactor alocis Cas9 (modified) translation          NKTYNAKKSNE-LLSEDIQKKMHSFWLSLL---------NKKLITKSKYDRLTRKGDFTD
18 Neisseria cinerea Cas9 (modified) translation          LGSENQNKGNQTPYEYFNGKDNSREWQEFKAR------VETSRFPRSKKQRILLQK--FD
24 Campylobacter lari Cas9 (modified) translation         FTKENQEKLNKTPFEAFG--KNIEKWSKIQTL------AQN--LPYKKKNKILDEN--FK
           Streptococcus pyogenes Cas9 translation        RSDKNRGKSDNVPSEEVVKKMKNYWRQLLN----------AKLITQRKFDNLTKAERGGL
 7 Lactobacillus johnsonii Cas9 (modified) translation    SRAVNNGKSDNVPVKLFGNEMAANLGMTIRKMWEEW--KNIGLISKTKYNNLLTDPDHIN
23 Mycoplasma gallisepticum Cas9 (modified) translation   LAESNQAKSNQTPYEFISSGNAGIKWEDYEAYCRKFKDGDSSLLDSTQRSKKFAKMMKTD
20 Parvibaculum lavamentivorans Cas9 (modified) translation RKDVNIEKGNRMPFEAFGHDEDRWSAIQIR---------LQGMVSAKGGTGMSPGKVKRF
 1 Corynebacter diptheriae Cas9 (modified) translation    VCHRCNQSKGNTPFAIWAKNTSIEGVSVKEAVERTRHWVTDTGMRSTDFKKFTKAVVERF
 2 Sutterella wadsworthensis Cas9 (modified) translation  YASSRGNQLKKNQRYSLSDLKANYRNEIFKTSNIAAITAEIEDVVTKLQQTHRLKFFDLL

21 Staphylococcus aureus subsp. Aureus Cas9 translation   FSVQKD--FINRNLVDTRYATRGLMNLLRSYFRVNNLD----------------------
26 Streptococcus thermophilus CRISPR1 Cas9 translation    FDVRKK--FIERNLVDTRYASRVVLNALQEHFRAHKID----------------------
 4 Treponema denticola Cas9 (modified) translation        DETAK---FIARQLVETRQATKVAAKVLEKMFPETKIV----------------------
 5 Filifactor alocis Cas9 (modified) translation          EELSG---FIARQLVETRQSTKAIADIFKQIY-SSEVV----------------------
18 Neisseria cinerea Cas9 (modified) translation          E---DG--FKERNLNDTRYINRFLCQFVADHMLLTGKG----------------------
24 Campylobacter lari Cas9 (modified) translation         DKQQED--FISRNLNDTRYIATLIAKYTKEYLNFLLLSENENANL---------------
           Streptococcus pyogenes Cas9 translation        SELDKAG-FIKRQLVETRQITKHVAQILDSRMNTKYDENDK-------------------
 7 Lactobacillus johnsonii Cas9 (modified) translation    KYKSAG--FIRRQLVETSQIIKLVSTILQSRYPNTEII----------------------
23 Mycoplasma gallisepticum Cas9 (modified) translation   TSSKYDIGFLARNLNDTRYATIVFRDALEDYANNHLVEDKP------------------
20 Parvibaculum lavamentivorans Cas9 (modified) translation LAKTMPEDFAARQLNDTRYAAKQILAQLKRLWPDMGPEAP--------------------
 1 Corynebacter diptheriae Cas9 (modified) translation    QRATMDEEIDARSMESVAWMANELRSRVAQHFASHGTT----------------------
 2 Sutterella wadsworthensis Cas9 (modified) translation  NEHEQDCVRHALFLDDGSEARDAVLELLATQRRTRVNGTQIWMIKNLANKIREELQNWCK

21 Staphylococcus aureus subsp. Aureus Cas9 translation   ---------VKVKSINGGFTSFLRR---KWKFKK------ERNKGYKHHAEDALIIANAD
26 Streptococcus thermophilus CRISPR1 Cas9 translation    ---------TKVSVVRGQFTSQLRR---HWGIEK------TRDT-YHHHAVDALIIAASS
 4 Treponema denticola Cas9 (modified) translation        ---------YSKAETVSMFRNK-------FDIVK------CREINDFHHAHDAYLNIVVG
 5 Filifactor alocis Cas9 (modified) translation          ---------YVKSSLVSDFRKKP------LNYLK------SRRVNDYHHAKDAYLNIVVG
18 Neisseria cinerea Cas9 (modified) translation          --------KRRVFASNGQITNLLRG---FWGLRK------VRAENDRHHALDAVVVACST
24 Campylobacter lari Cas9 (modified) translation         -KSGEKGSKIHVQTISGMLTSVLRH---TWGFDK------KDRNNHLHHALDAIIVAYST
```

FIG. 19I

```
                Streptococcus pyogenes Cas9 translation   -----LIREVKVITLKSKLVSDFRK---DFQFYK------VREINNYHHAHDAYLNAVVG
  7 Lactobacillus johnsonii Cas9 (modified) translation   ------------TVKAKYNHYLRE---KFDLYK------SREVNDYHHAIDAYLSAICG
 23 Mycoplasma gallisepticum Cas9 (modified) translation   --------MFKVVCINGSVTSFLRKNFDDSSYAK------KDRDKNIHHAVDASIISIFS
20 Parvibaculum lavamentivorans Cas9 (modified) translation   ---------VKVEAVTGQVTAQLRKLWTLNNILADDG--EKTRADHRHHAIDALTVACTH
  1 Corynebacter diptheriae Cas9 (modified) translation   -----------VRVYRGSLTAEARRASGISGKLKFFDGVGKSRLDRRHHAIDAAVIAFTS
  2 Sutterella wadsworthensis Cas9 (modified) translation   TTNNRLHFQAAATNVSDAKNLRLKLAQNQPDFEK------PDIQPIASHSIDALCSFAVG

 21 Staphylococcus aureus subsp. Aureus Cas9 translation   FIFKEWKKLDKAKKVM------------------ENQMFEEKQAESMPEIET-EQEYKE
 26 Streptococcus thermophilus CRISPR1 Cas9 translation   QLNLWKKQKNTLVSYS------------------EDQLLDIETGELISDDEYKESVFKA
  4 Treponema denticola Cas9 (modified) translation   NVYNTKFTNNPWNFIK------------------EKRDNPKIADTYNYYKVFDYDVKRN
  5 Filifactor alocis Cas9 (modified) translation   NVYNKKFTSNPIQWMK------------------KNRD-----TNYSLNKVFEHDVVIN
 18 Neisseria cinerea Cas9 (modified) translation   IAMQQKIT-RFVRYKE------------------MNAFDGKTIDKETGEVLHQKAHFPQ
 24 Campylobacter lari Cas9 (modified) translation   NSIIKAFS-DFRKNQE------------------LLKAR--------------------
                Streptococcus pyogenes Cas9 translation   TALIKKYPKLESEFVYGDYK-----VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTE
  7 Lactobacillus johnsonii Cas9 (modified) translation   NLLYQNYPNLRPFFVYG------------------QYKKFSSDPDKEKAIFNKTRKFSFI
 23 Mycoplasma gallisepticum Cas9 (modified) translation   NETKTLFNQLTQFADYKLFK-----NTDGSWKKIDPKTGVVTEVTDENWKQIRVRNQVSE
20 Parvibaculum lavamentivorans Cas9 (modified) translation   PGMTNKLSRYWQLRDD------------------------------------------
  1 Corynebacter diptheriae Cas9 (modified) translation   DYVAETLAVRSNLKQSQ------------------AHRQEAPQWREFTGKDAEHRAAWRV
  2 Sutterella wadsworthensis Cas9 (modified) translation   SADAERDQNGFDYLDGKTVLGLYPQSCEVIHLQAKPQEEKSHFDSVAIFKEGIYAEQFLP

 21 Staphylococcus aureus subsp. Aureus Cas9 translation   IFITPHQIKHIKDFKD-YKYSHRVDKKPNRELINDTLYSTR-----KDDKGNTLIVNNLN
 26 Streptococcus thermophilus CRISPR1 Cas9 translation   PYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQAKVGKDKADETYVLGKIK
  4 Treponema denticola Cas9 (modified) translation   NITAWEK--------------GKTIITVKDMLKRNTPIYTRQAACKKGELFNQTIMKKGL
  5 Filifactor alocis Cas9 (modified) translation   GEVIWEKCTYHEDTNT---YDGGTLDRIRKIVERDNILYTEYAYCEKGELFNATIQNKNG
 18 Neisseria cinerea Cas9 (modified) translation   PWEFFAQEVMIRVFGK--PDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHKYVTPLFISR
 24 Campylobacter lari Cas9 (modified) translation   ---FYAKELTSDNY------------------KHQVKFFEPFKSFREKILSKIDEIFVSK
                Streptococcus pyogenes Cas9 translation   ITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES
  7 Lactobacillus johnsonii Cas9 (modified) translation   SQLLKNKSENSKEIAK----------KLKRAYQFKYMLVSRETETRDQEMFKMTVYPRFS
 23 Mycoplasma gallisepticum Cas9 (modified) translation   IAKVIEKYIQDSNIERKARYSRKIENKTNISLFNDTVYSAKKVGYEDQIKRKNLKTLDIH
20 Parvibaculum lavamentivorans Cas9 (modified) translation   ----------------------------PRAEKPALTPPWDTIRADAEKAVSEIVVSH
  1 Corynebacter diptheriae Cas9 (modified) translation   WCQKMEKLSALLTEDLR---------DDRVVVMSNVRLRLGNGSAHKETIGKLSKVKLSS
  2 Sutterella wadsworthensis Cas9 (modified) translation   IFTLNEKIWIGYETLNAKGERCGAIEVSGKQPKELLEMLAPFFNKPVGDLSAHATYRILK
```

FIG. 19J

```
 21 Staphylococcus aureus subsp. Aureus Cas9 translation   GLYDKD-NDKLKKLINKSPEKLLMYHHDP--------------------QTYQK-LKLI
 26 Streptococcus thermophilus CRISPR1 Cas9 translation    DIYTQDGYDAFMKIYKKDKSKFLMYRHDP--------------------QTFEKVIEPI
  4 Treponema denticola Cas9 (modified) translation        GQHPLKKEGPFSNISKYGG-----YNKVS--------------------AAYYTLIEYE
  5 Filifactor alocis Cas9 (modified) translation          NSTVSLKKG--LDVKKYGG-----YFSAN--------------------TSYFSLIEFE
 18 Neisseria cinerea Cas9 (modified) translation          APNRKMSGQGHMETVKSAKRLDEGISVLR--------------------VPLTQLKLKD
 24 Campylobacter lari Cas9 (modified) translation         PPRKRARRALHKDTFHSENKIIDKCSYN------------------------------
          Streptococcus pyogenes Cas9 translation          ILPKRNSDKLIARKKDWDPKKYGGFDSPT--------------------VAYSVLVVAK
  7 Lactobacillus johnsonii Cas9 (modified) translation    HDTVKAPRNLIPKKMGMSPDIYGGYTNNS--------------------DAYMVIVRID
 23 Mycoplasma gallisepticum Cas9 (modified) translation   ESAKENKNSKVKRQFVYRKLVNVSLLNNDKLADLFAEKEDILMYRANPWVINLAEQIFNE
20 Parvibaculum lavamentivorans Cas9 (modified) translation RVRKKVSGPLHKETTYGDTGTDIKTKSGT------------------------------
  1 Corynebacter diptheriae Cas9 (modified) translation    QLSVSDIDKASSEALWCALTREPGFDPKE------------------------------
  2 Sutterella wadsworthensis Cas9 (modified) translation  KPAYEFLAKAALQPLSAEEKRLAALLDALRYCTSRKSLMSLFMAANGKSLKKREDVLKPK

 21 Staphylococcus aureus subsp. Aureus Cas9 translation   MEQYGD----EK------NPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDI
 26 Streptococcus thermophilus CRISPR1 Cas9 translation    LENYPNKQINEKGKEVPCNPFLKYKEEHG-YIRKYSKKGNGPEIKSLKYYDSKLGNHIDI
  4 Treponema denticola Cas9 (modified) translation        EK-GNKIRSLETIP----LYLVKDIQKDQDVLKSYLTDLLGKKEFKILVPKIKINSLLKI
  5 Filifactor alocis Cas9 (modified) translation          DKKGDRARHIIGVP----IYIANMLEHSPSAFLEYCE-QKGYQNVRILVEKIKKNSLLII
 18 Neisseria cinerea Cas9 (modified) translation          LEKMVNREREPKLYEALKARLEAHKDDPAKAFAEPFYKYDKAGNRTQQVKAVRVEQVQKT
 24 Campylobacter lari Cas9 (modified) translation         ------------------------------------------SKEGLQIALSCGRVRKI
          Streptococcus pyogenes Cas9 translation          VEKGKSKK----------LKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKL
  7 Lactobacillus johnsonii Cas9 (modified) translation    KKKGTEYKILG-------IPTRELVNLKKAEKEDHYKSYLKEILTPRILYNKNGKRDKKI
 23 Mycoplasma gallisepticum Cas9 (modified) translation   YTENKKIKSQNVFEKYMLDLTKEFPEKFSEFLVKSMLRNKTAIIYDDKKNIVHRIKRLKM
20 Parvibaculum lavamentivorans Cas9 (modified) translation ------------------------------YRQFVTRKKIESLSKGELDEIRDPRIKEI
  1 Corynebacter diptheriae Cas9 (modified) translation    ------------------GLPANPERHIRVNGTHVYAGDNIGLFPVSAGSIALRGGYAE
  2 Sutterella wadsworthensis Cas9 (modified) translation  LFQLKVELKGEKSFKLNGSLTLPVKQDWLRICDSPELADAFGKPCSADELTSKLARIWKR

 21 Staphylococcus aureus subsp. Aureus Cas9 translation   TDDYPNSRNKVVKLSLKPYRFDVYLDN--GVYKFVTVKNLDVIKKE--------------
 26 Streptococcus thermophilus CRISPR1 Cas9 translation    TPK--DSNNKVVLQSVSPWRADVYFNKTTGKYEILGLKYADLQFEKGT------------
  4 Treponema denticola Cas9 (modified) translation        NGFPCHITGKTNDSFLLRPAVQFCCSNNEVLYFKKIIRFSEIRSQREKIGKTISPYEDLS
  5 Filifactor alocis Cas9 (modified) translation          NGYPLRIRGENEVDTSFKRAIQLKLDQKNYELVRNIEKFLEKYVEK--------------
 18 Neisseria cinerea Cas9 (modified) translation          GVWVHNHNGIADNATIVRVDVFEKGGKYYLVPIYSWQVAKGILPDR--------------
```

FIG. 19K

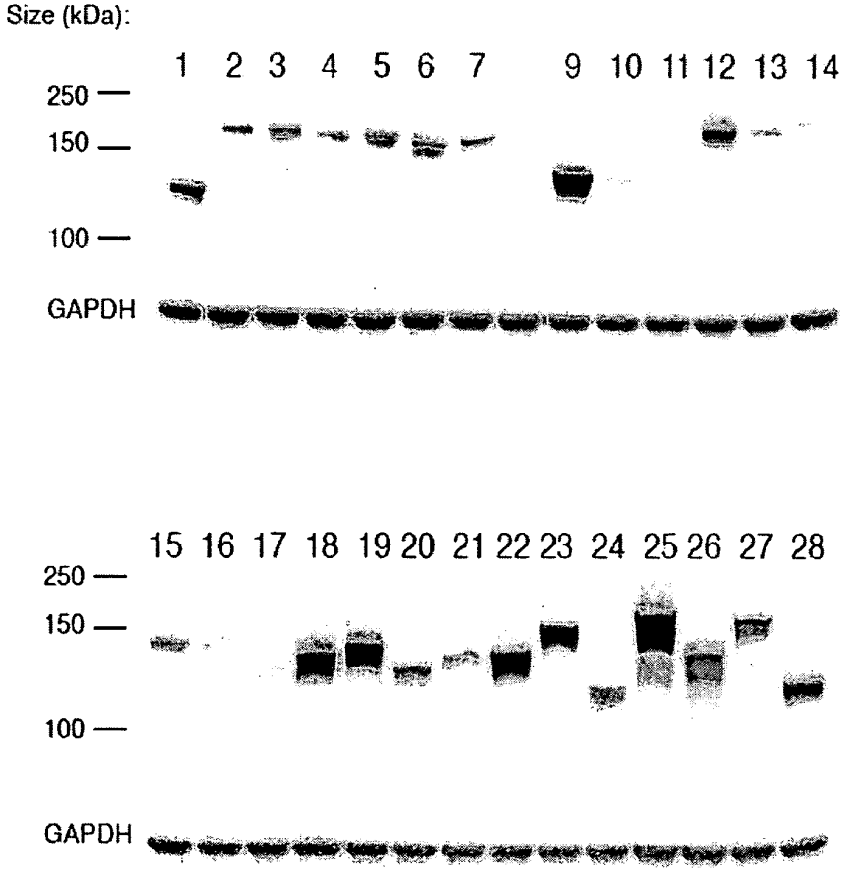
```
        24 Campylobacter lari Cas9 (modified) translation   G------TKYVENDTIVRVDIFKKQNKFYAIPIYAMDFALGILPNK--------------
                  Streptococcus pyogenes Cas9 translation   PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEK--------------
         7 Lactobacillus johnsonii Cas9 (modified) translation   TSFEIVKSKIPYKQVIQDGDKKFMLGSSTYVYNAKQLTLSTESMKA--------------
        23 Mycoplasma gallisepticum Cas9 (modified) translation   LSSELKENKLSNVIIRSKNQSGTKLSYQDTINSLALMIMRSIDPTAKKQYIRVPLN----
    20 Parvibaculum lavamentivorans Cas9 (modified) translation   VAAHVAGRGGDPKKAFPPYPCVSPGGPEIRKVRLTSKQQLNLMAQTG-------------
         1 Corynebacter diptheriae Cas9 (modified) translation   LGSSFHHARVYKITSGKKPAFAMLRVYTIDLLPYRNQDLFSVELKP--------------
        2 Sutterella wadsworthensis Cas9 (modified) translation   PVMRDLAHAPVRREFSLPAIDNPSGGFRIRRTNLFGNELYQVHAINAKKYRGFASAG---

        21 Staphylococcus aureus subsp. Aureus Cas9 translation   -----NYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKING------ELYRVIG-VN
         26 Streptococcus thermophilus CRISPR1 Cas9 translation   -----GTYKISQEKYNDIKKKEGVDSDSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRTM
             4 Treponema denticola Cas9 (modified) translation   FRSYIKENLWKKTKNDEIGEKEFYDLLQKKNLEIYDMLLTKHKDTIYKKRPNSATIDILV
               5 Filifactor alocis Cas9 (modified) translation   -----KGNYPIDENRDHIT--------HEKMNQLYEVLLSKMKK--FNKKGMADPSDRIE
              18 Neisseria cinerea Cas9 (modified) translation   ------AVVQGKDEED----WTVMDDSFEFKFVLYANDLIKLTAK--KNEFLGYFVSLNR
            24 Campylobacter lari Cas9 (modified) translation   ------IVITGKDKNNNPKQWQTIDESYEFCFSLYKNDLILLQKKNMQEPEFAYYNDFSI
                      Streptococcus pyogenes Cas9 translation   ---------LKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYN
         7 Lactobacillus johnsonii Cas9 (modified) translation   ---------ITNNFDKDSDENDALIKAYDEILDKVDKYLPLFDIN--------KFREKLH
        23 Mycoplasma gallisepticum Cas9 (modified) translation   -----TLNLHLGDHDFDLHNMDAYLKKPKFVKYLKANEIGDEYKPWRVLTSGTLLIHKKD
    20 Parvibaculum lavamentivorans Cas9 (modified) translation   ------NGYADLGSNHHIAIYRLPDGKADFEIVSLFDASRRLAQR---NPIVQRTRADGA
         1 Corynebacter diptheriae Cas9 (modified) translation   -------QTMSMRQAEKKLRDALATGNAEYLGWLVVDDELVVDTSKIATDQVKAVEAELG
        2 Sutterella wadsworthensis Cas9 (modified) translation   ----SNVDWSKGILFNELQHENLTECGGRFITSADVTPMSEWRKVVAEDNLSIWIAPGTE

        21 Staphylococcus aureus subsp. Aureus Cas9 translation   NDLLNRIEVNMIDITYREYLENM---------NDKR-----PPRIIKTIASKTQSIKKYS
         26 Streptococcus thermophilus CRISPR1 Cas9 translation   PKQKHYVELKPYDKQKFEGGEAL---------IKVLGNVANSGQCKKGLGKSNISIYKVR
             4 Treponema denticola Cas9 (modified) translation   KGKEKFKSLI-IENQFEVILEIL---------KLFSATRN-VSDLQHIGGSKYSGVAKIG
               5 Filifactor alocis Cas9 (modified) translation   KSKPKFIKLEDLIDKINVINKML---------NLLRCDNDTKADLSLIELPKNAGSFVVK
              18 Neisseria cinerea Cas9 (modified) translation   ATGAIDIRTHD----------------------TDSTKGKNGIFQSVGVKTALSFQKYQ
            24 Campylobacter lari Cas9 (modified) translation   STSSICVEKHDNKFENLTSNQKL----------LFSNAKEGSVKVESLGIQNLKVFEKYI
                      Streptococcus pyogenes Cas9 translation   KHRDKPIREQAENIIHLFTLTNL---------GAPAAFKYFDTTIDRK---RYTSTKEVL
         7 Lactobacillus johnsonii Cas9 (modified) translation   SGREKFIKLSLEDKKDTILKVLE---------GLHDN--AVMTKIPTIGLSTPLGFMQFP
        23 Mycoplasma gallisepticum Cas9 (modified) translation   KKLMYISSFQNLNDVIEIKNLIE---------TEYKENDDSDSKKKKKANRFLMTLSTIL
    20 Parvibaculum lavamentivorans Cas9 (modified) translation   SFVMSLAAGEAIMIPEGSKKGIW---------IVQGVWASGQVVLERDTDADHSTTTRPM
         1 Corynebacter diptheriae Cas9 (modified) translation   TIRRWRVDGFFSPSKLRLRPLQM---------SKEGIKKESAPELSKIIDRPGWLPAVNK
```

FIG. 19L

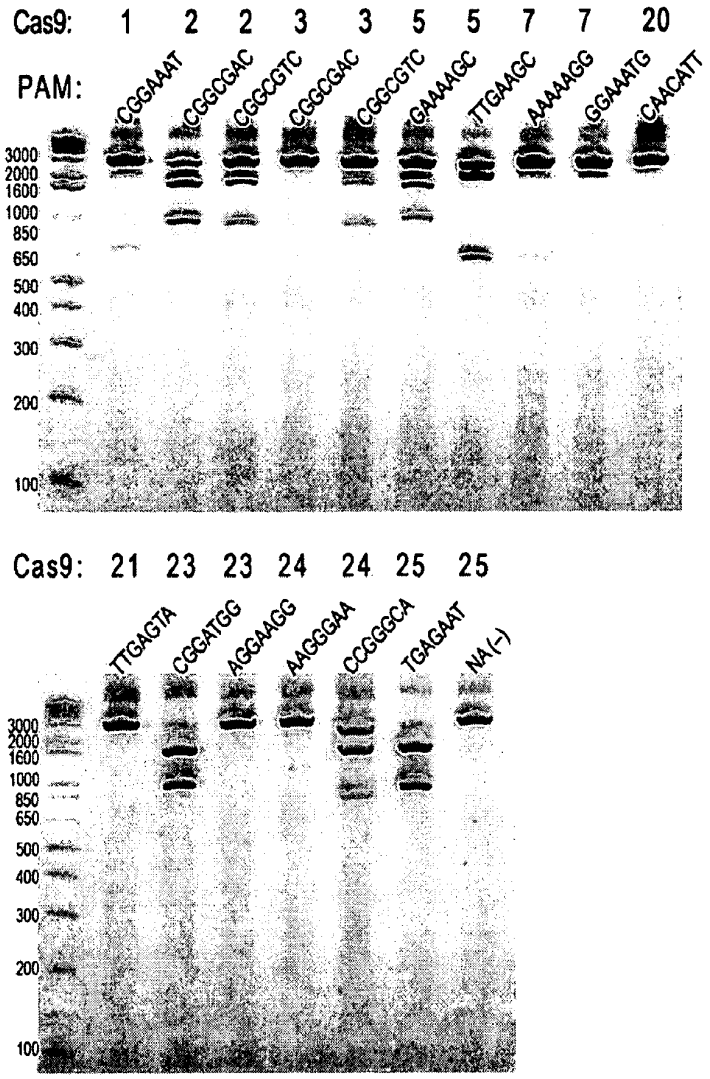
2	Sutterella wadsworthensis Cas9 (modified)	translation	GRRYVRVETTFIQASHWFEQSVENWAITSPLSLPASFKVDKPAEFQKAVGTELSELLGQP
21	Staphylococcus aureus subsp. Aureus Cas9	translation	TDILGNLYEVKSKKHPQIIKKG----------
26	Streptococcus thermophilus CRISPR1 Cas9	translation	TDVLGNQHIIKNEGDKPKLDF----------
4	Treponema denticola Cas9 (modified)	translation	NKISSLDNCILIYQSITGIFEKRIDLLKV---------
5	Filifactor alocis Cas9 (modified)	translation	KNTIGKSKIILVNQSVTGLYENRREL----------
18	Neisseria cinerea Cas9 (modified)	translation	IDELGKEIRP---------CRLKKRPPVR---------
24	Campylobacter lari Cas9 (modified)	translation	ITPLGDKIKADFQPRENISLKTSKKYGLR---------
	Streptococcus pyogenes Cas9	translation	DATLIHQSITGLYETRIDLSQLGGD----------
7	Lactobacillus johnsonii Cas9 (modified)	translation	NGVILSENAKLIYQSPTGLFKKSVKISDL---------
23	Mycoplasma gallisepticum Cas9 (modified)	translation	NDYILLDAKDNFDILGLSKNRIDEILNSKLGLDKIVK-
20	Parvibaculum lavamentivorans Cas9 (modified)	translation	PNPILKDDAKKVSIDPIGRVRPSND----------
1	Corynebacter diptheriae Cas9 (modified)	translation	LFSDGNVTVVRRDSLGRVRLESTAHLPVTWKVQ-----
2	Sutterella wadsworthensis Cas9 (modified)	translation	RSEIFIENVGNAKHIRFWYIVVSSNKKMNESYNNVSKS



1	Corynebacter_diphtheriae	119.24
2	Sutterella_wadsworthensis	156.42
3	Legionella_pneumophila_Paris	150.92
4	Treponema_denticola	153.45
5	Filifactor_alocis	150.15
6	Staphylococcus_pseudintermedius	146.74
7	Lactobacillus_johnsonii	151.25
*8	Eubacterium_ventriosum	121.77
9	Streptococcus_pasteurianus	124.3
10	Lactobacillus_farciminis	123.86
11	Mycoplasma_mobile	135.96
12	Bacteroides_coprophilus	165.99
13	Flaviivola_taffensis	160.38
14	Flavobacterium_columnare	168.85
15	Sphaerochaeta_globus_Buddy	129.69
16	Azospirillum_B510	128.48
17	Gluconacetobacter_diazotrophicus	115.5
18	Neisseria_cinerea	119.02
19	Roseburia_intestinalis	125.4
20	Parvibaculum_lavamentivorans	114.07
21	Staphylococcus aureus subsp. Aureus –	115.83
22	Nitratifractor salsuginis DSM 16511	124.52
23	Mycoplasma gallisepticum str. F	139.59
24	Campylobacter lari CF89-12	110.33
25	Streptococcus pyogenes	158
26	Streptococcus thermophilus CRISPR 1	127
27	Streptococcus thermophilus CRISPR 3	153
28	Campylobacter jejuni	110

* No DNA added for #8

FIG. 20



Cas9 #	Species	Putative PAM	pUC19 target
1	Corynebacter_diphtheriae		AAAAGGGAATAAGGGCGACA
2	Sutterella_wadsworthensis	NGGNNNN	CATTCTGAGAATAGTGTATG
2	Sutterella_wadsworthensis		GCGACCGAGTTGCTCTTGCC
3	Legionella_pneumophila_Paris	NGGNNNN	CATTCTGAGAATAGTGTATG
3	Legionella_pneumophila_Paris		GCGACCGAGTTGCTCTTGCC
5	Filifactor_alocis	NNNAAGC	TTGAGTACTCACCAGTCACA
5	Filifactor_alocis		CTCTTCCTTTTTCAATATTA
7	Lactobacillus_johnsonii	NNAAANN	TAAATGCTTCAATAATATTG
7	Lactobacillus_johnsonii		AAAGGGAATAAGGGCGACAC
20	Parvibaculum_lavamentivorans	NNNCATN	AAAGGAAGAGTATGAGTATT
21	Staphylococcus aureus subsp. Aureus		CTATTCTCAGAATGACTTGG
23	Mycoplasma gallisepticum str. F	NNGANgg	AGTCACAGAAAAGCATCTTA
23	Mycoplasma gallisepticum str. F		GTTTCTGGGTGAGCAAAAAC
24	Campylobacter lari CF89-12	NNGGGNA	GAAGGCAAAATGCCGCAAAA
24	Campylobacter lari CF89-12		GGTATTATCCCGTATTGACG
25	SpCas9		GGGACTCAACCAAGTCATTC
25	SpCas9		

FIG. 21

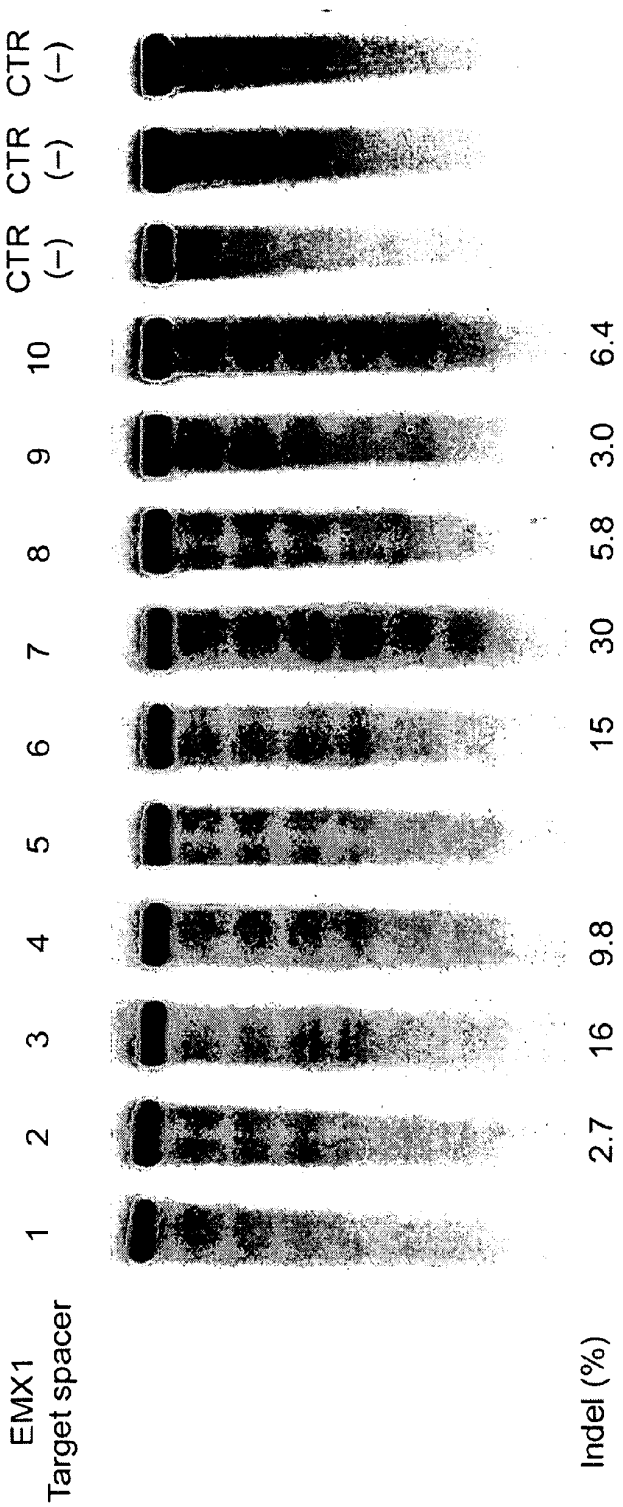

EMX1
Target spacer
1
2
3
4
5
6
7
8
9
10
CTR (–)
CTR (–)
CTR (–)
Indel (%)
2.7
16
9.8
15
30
5.8
3.0
6.4

FIG. 22

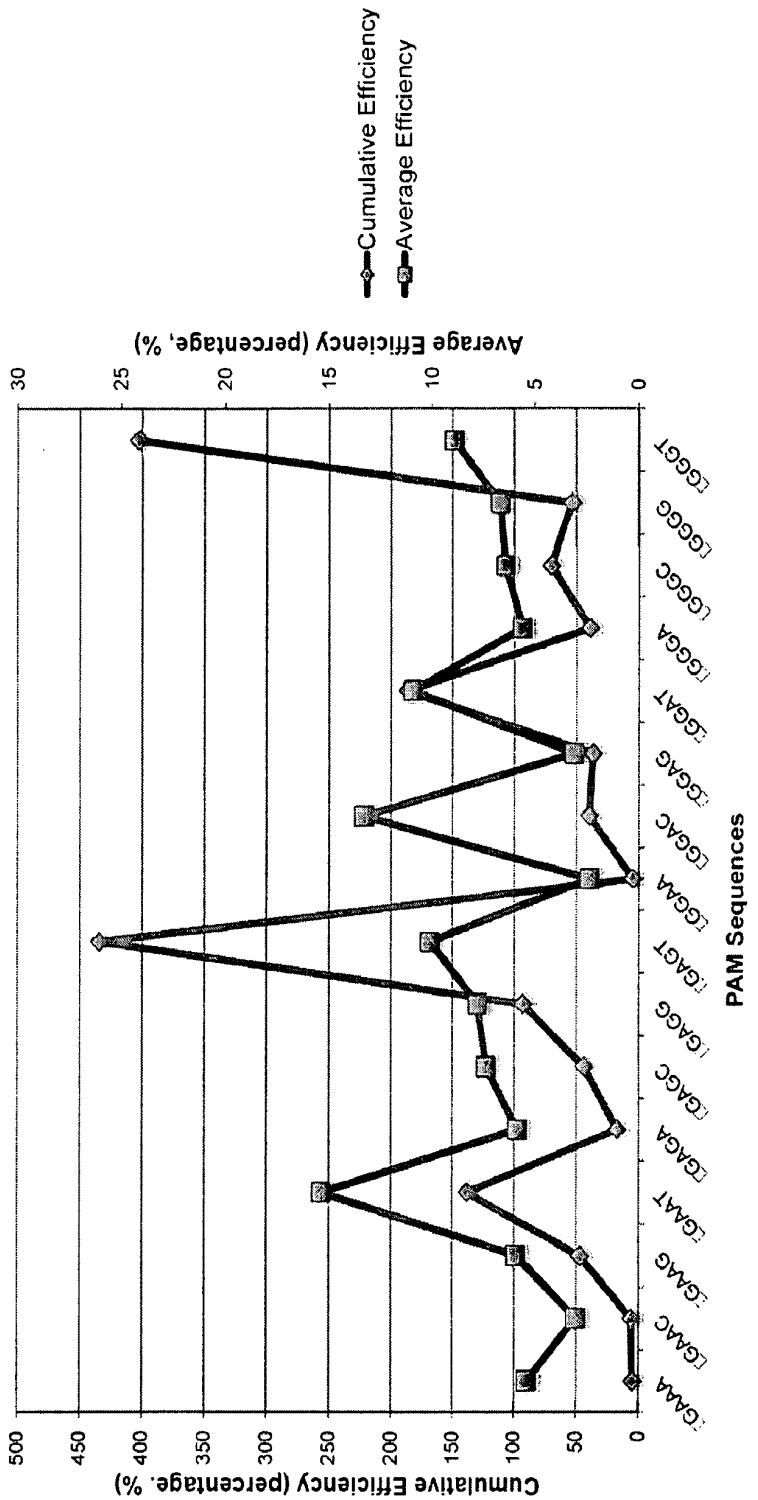

Cumulative Efficiency
Average Efficiency
Cumulative Efficiency (percentage, %)
Average Efficiency (percentage, %)
PAM Sequences
TGAA
TGAC
TGAG
TGAT
TGCA
TGCC
TGCG
TGCT
TGGA
TGGC
TGGG
TGGT

FIG. 23

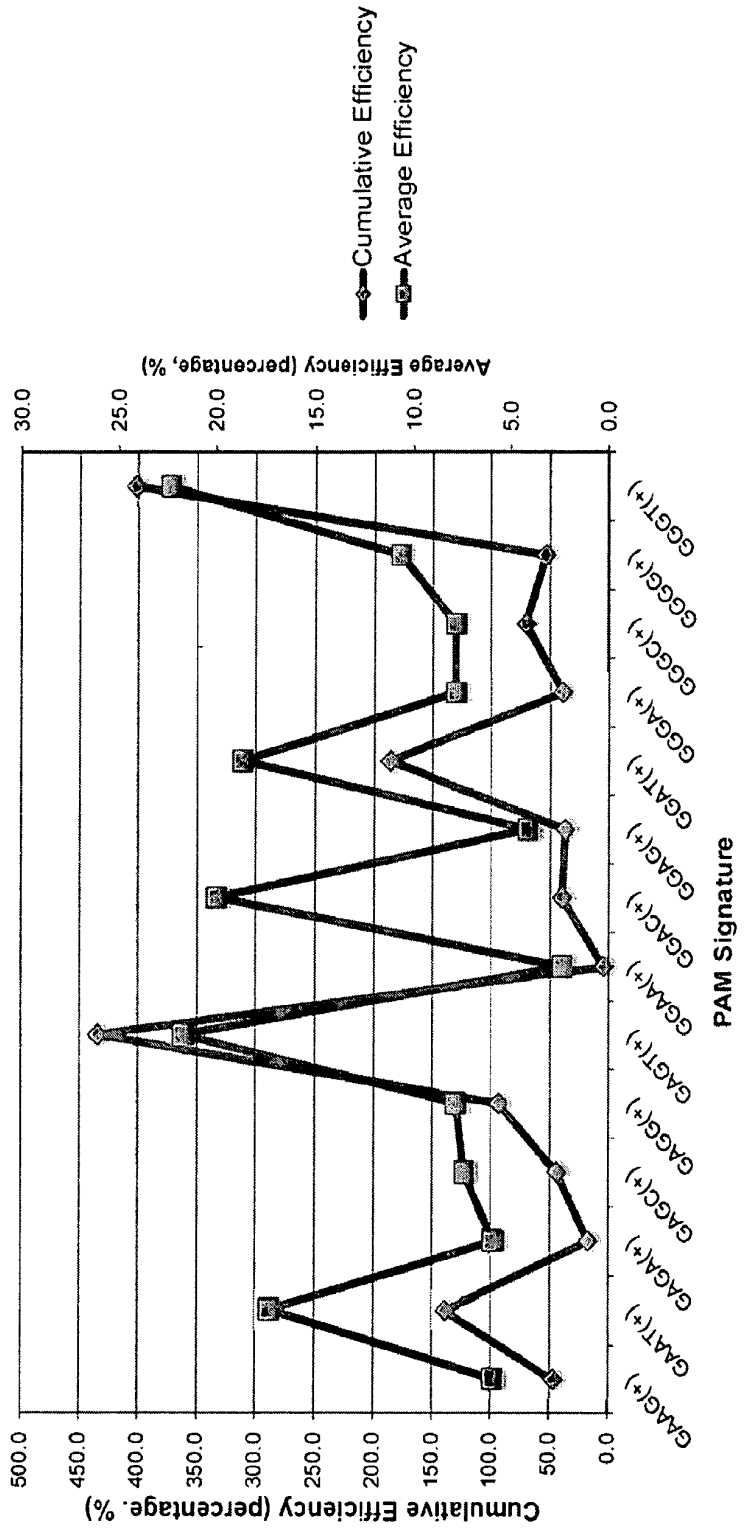

Cumulative Efficiency
Average Efficiency
Cumulative Efficiency (percentage, %)
Average Efficiency (percentage, %)
PAM Signature
GAAG(+)
GAAT(+)
GAGA(+)
GAGC(+)
GAGG(+)
GAGT(+)
GGAA(+)
GGAC(+)
GGAG(+)
GGAT(+)
GGGA(+)
GGGC(+)
GGGG(+)
GGGT(+)

FIG. 24

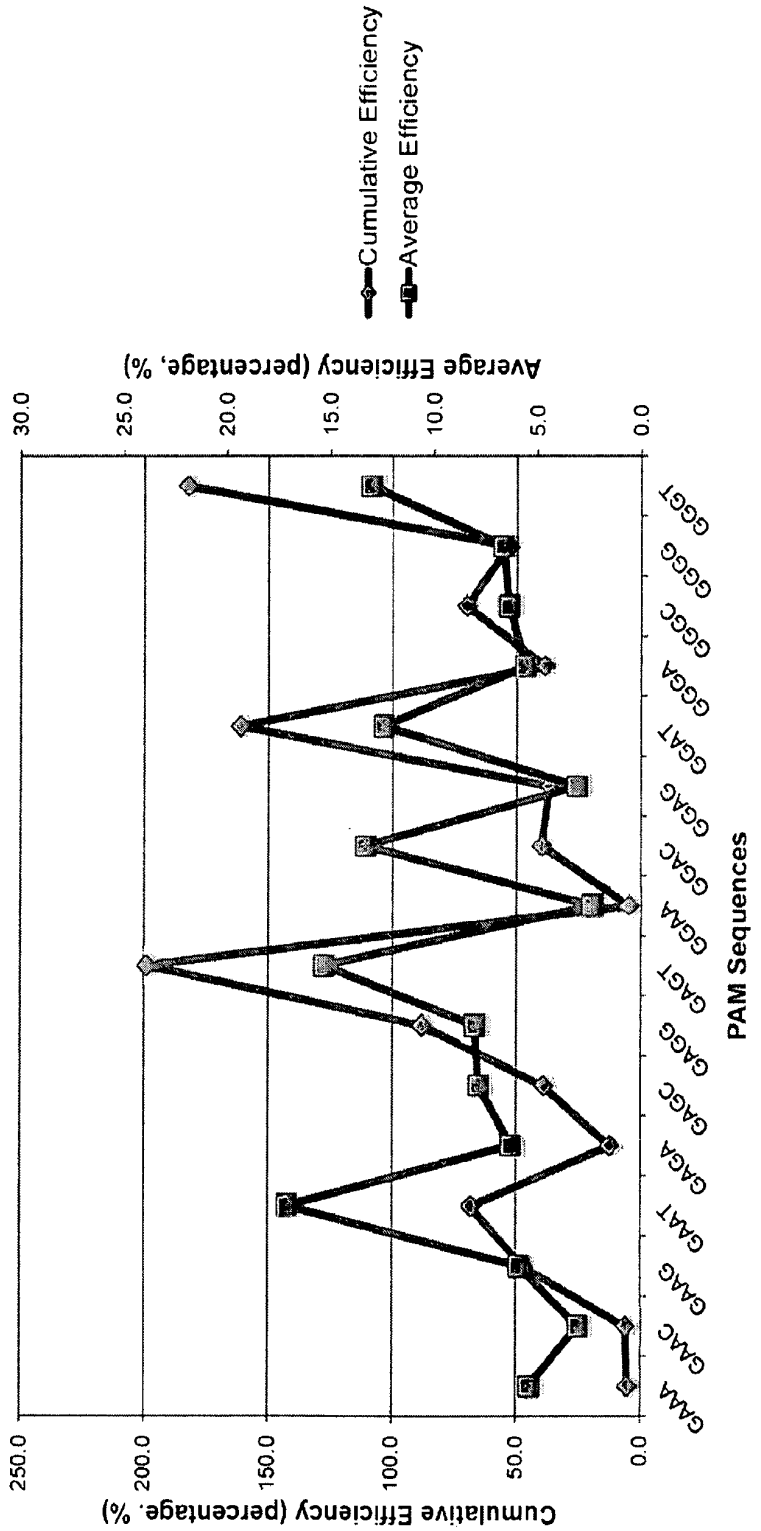

Cumulative Efficiency
Average Efficiency
Cumulative Efficiency (percentage, %)
Average Efficiency (percentage, %)
PAM Sequences
GAA
GAC
GAG
GAT
GCA
GCC
GCG
GCT
GGA
GGC
GGG
GGT
0.0
50.0
100.0
150.0
200.0
250.0
0.0
5.0
10.0
15.0
20.0
25.0
30.0

FIG. 25

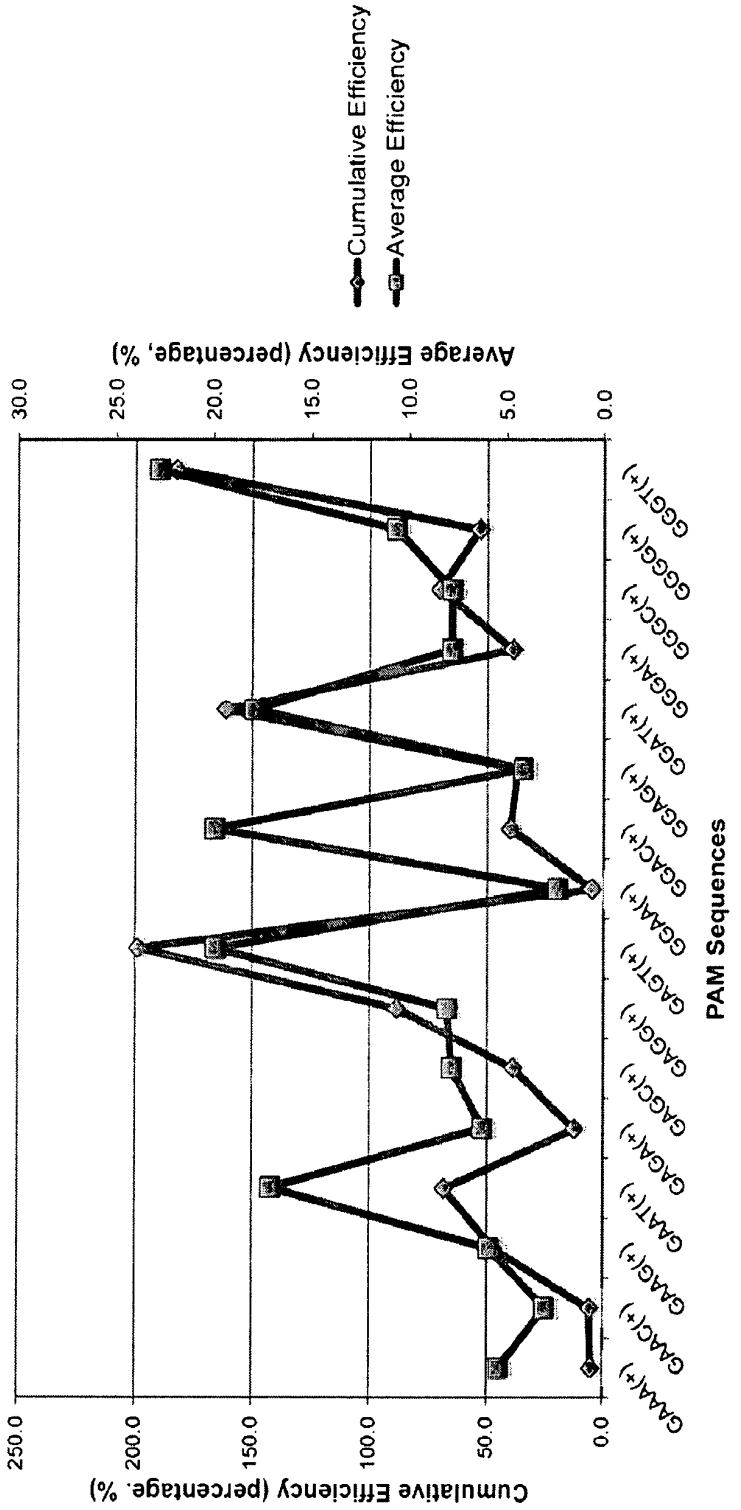

Cumulative Efficiency
Average Efficiency
Cumulative Efficiency (percentage, %)
Average Efficiency (percentage, %)
PAM Sequences
GAAA(+)
GAAC(+)
GAAG(+)
GAAT(+)
GACA(+)
GACC(+)
GACG(+)
GACT(+)
GAGA(+)
GAGC(+)
GAGG(+)
GAGT(+)
GATA(+)
GATC(+)
GATG(+)
GATT(+)
GGAA(+)
GGAC(+)
GGAG(+)
GGAT(+)
GGCA(+)
GGCC(+)
GGCG(+)
GGCT(+)

FIG. 26

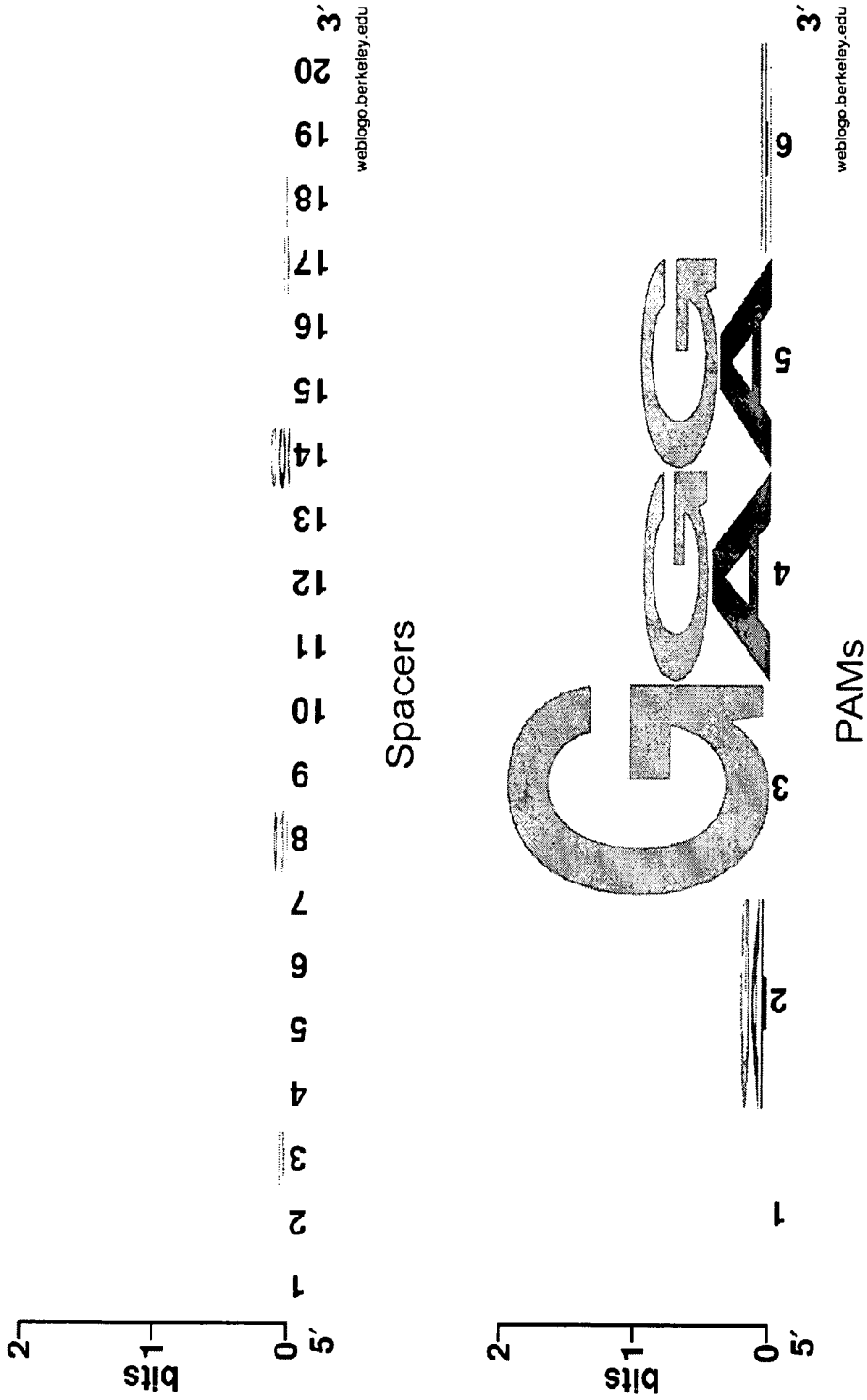

Spacers
PAMs
bits
2
1
0
5′
3′
weblogo.berkeley.edu
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
1 2 3 4 5 6

FIG. 27