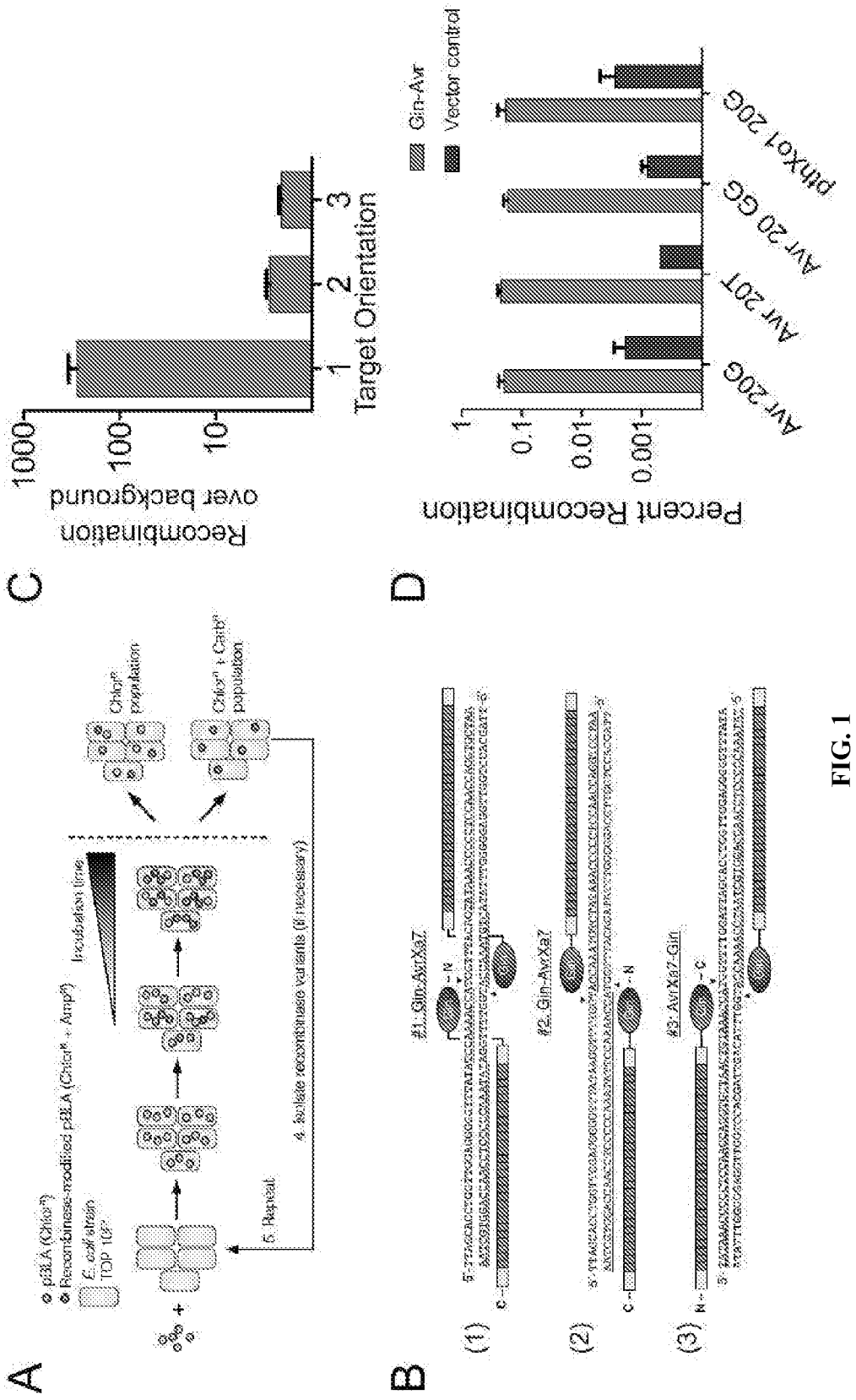

A
pBLA (Chlor^R)
Recombinase-modified pBLA (Chlor^R + Amp^R)
E. coli strain TOP 10^F'
Incubation time
Chlor^R population
Chlor^R + Carb^R population
4. Isolate recombinase variants (if necessary)
5. Repeat
B
#1. Gin-AvrXa7
#2. Gin-AvrXa7
#3. AvrXa7-Gin
C
Recombination over background
Target Orientation
D
Gin-Avr
Vector control
Percent Recombination
AVR 20G
AVR 20T
AVR 20 GG
pthXo1 20G

FIG. 1

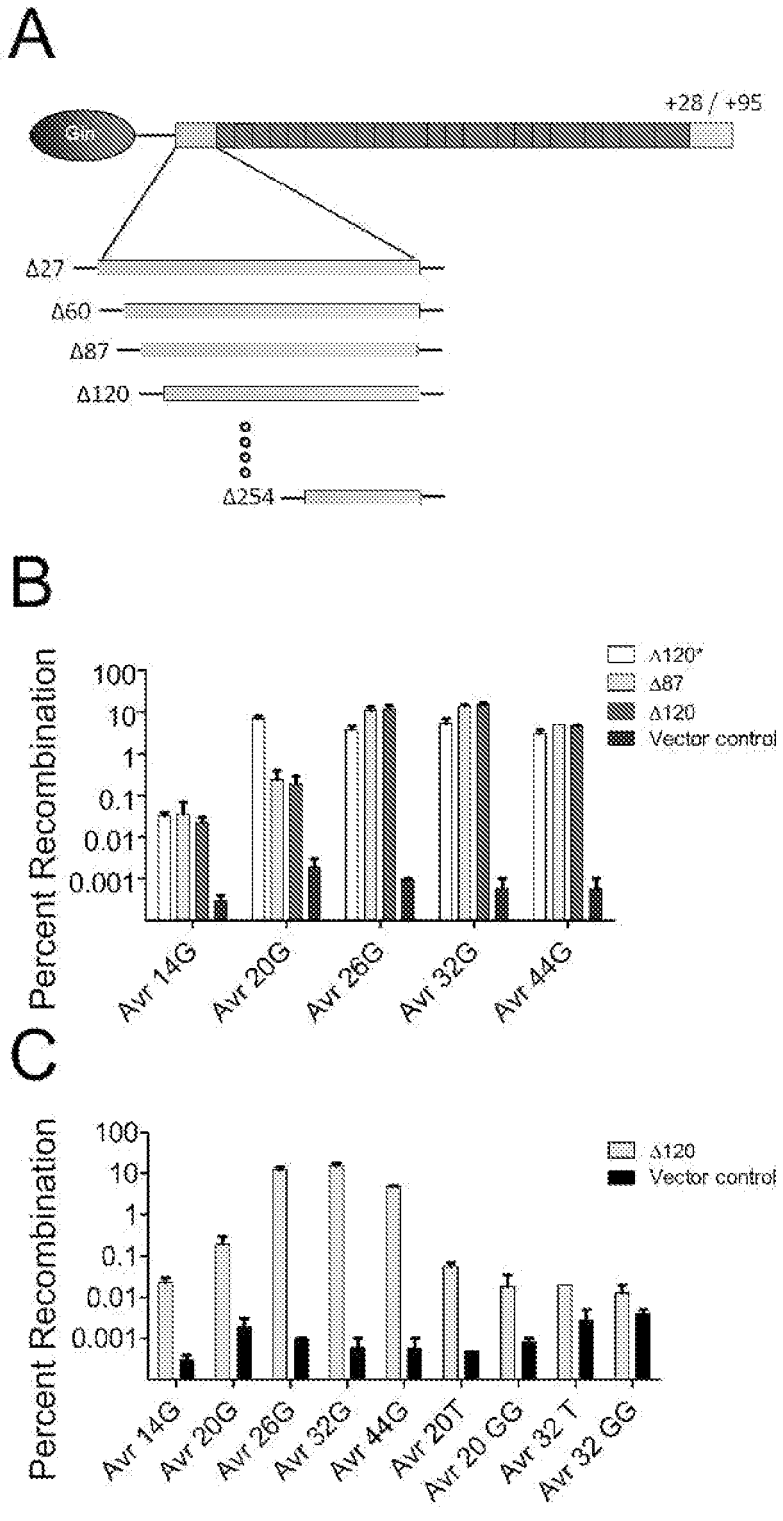

A
+28 / +95
Δ27
Δ60
Δ87
Δ120
Δ254
B
Percent Recombination
100
10
1
0.1
0.01
0.001
Δ120*
Δ87
Δ120
Vector control
Avr 14G
Avr 20G
Avr 26G
Avr 32G
Avr 44G
C
Percent Recombination
100
10
1
0.1
0.01
0.001
Δ120
Vector control
Avr 14G
Avr 20G
Avr 26G
Avr 32G
Avr 44G
Avr 20T
Avr 20 GG
Avr 32 T
Avr 32 GG

FIG. 2

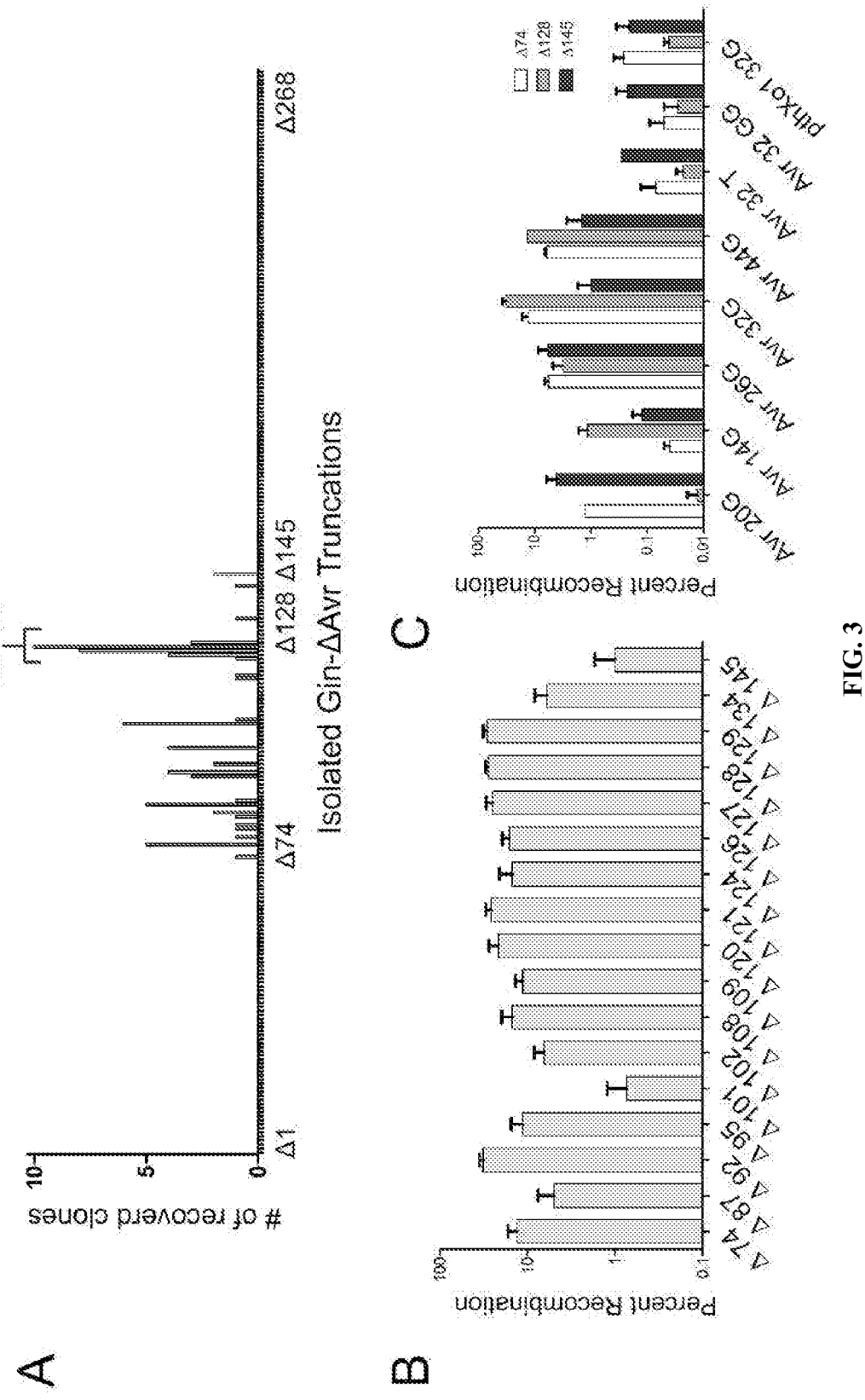

A
26/73
# of recovered clones
10
5
0
Δ1
Δ74
Δ128 Δ145
Δ268
Isolated Gin-ΔAvr Truncations
B
Percent Recombination
100
10
1
0.1
Δ74
Δ87
Δ92
Δ95
Δ101
Δ102
Δ108
Δ109
Δ120
Δ121
Δ124
Δ126
Δ127
Δ128
Δ129
Δ134
Δ145
C
Percent Recombination
100
10
1
0.1
0.01
Δ74
Δ128
Δ145
AVR 20G
AVR 14G
AVR 26G
AVR 32G
AVR 44G
AVR 32 T
AVR 32 GG
pthXo1 32G

FIG. 3

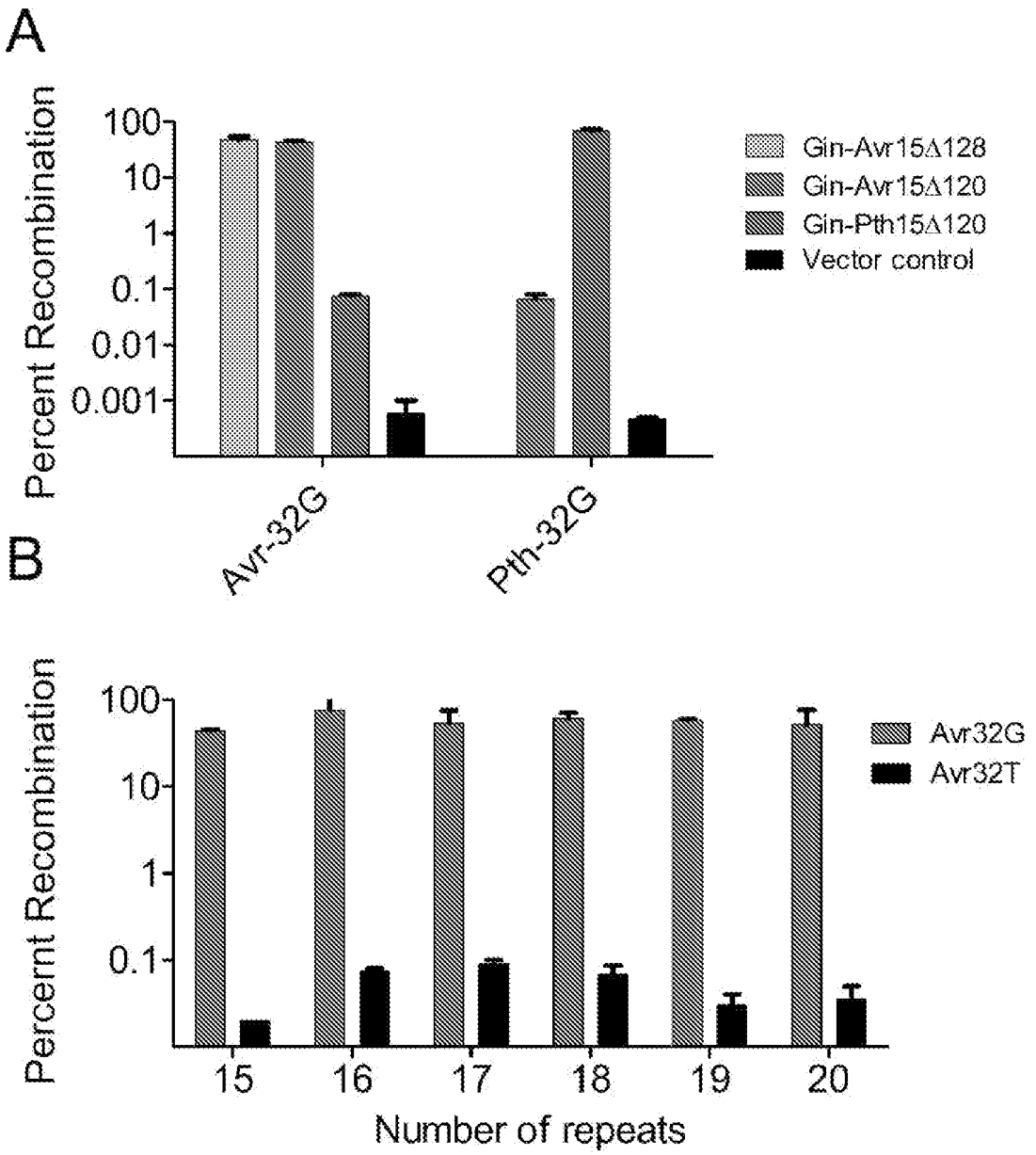

A
Percent Recombination
100
10
1
0.1
0.01
0.001
Gin-Avr15Δ128
Gin-Avr15Δ120
Gin-Pth15Δ120
Vector control
Avr-32G
Pth-32G
B
Percernt Recombination
100
10
1
0.1
Avr32G
Avr32T
15
16
17
18
19
20
Number of repeats

FIG. 4

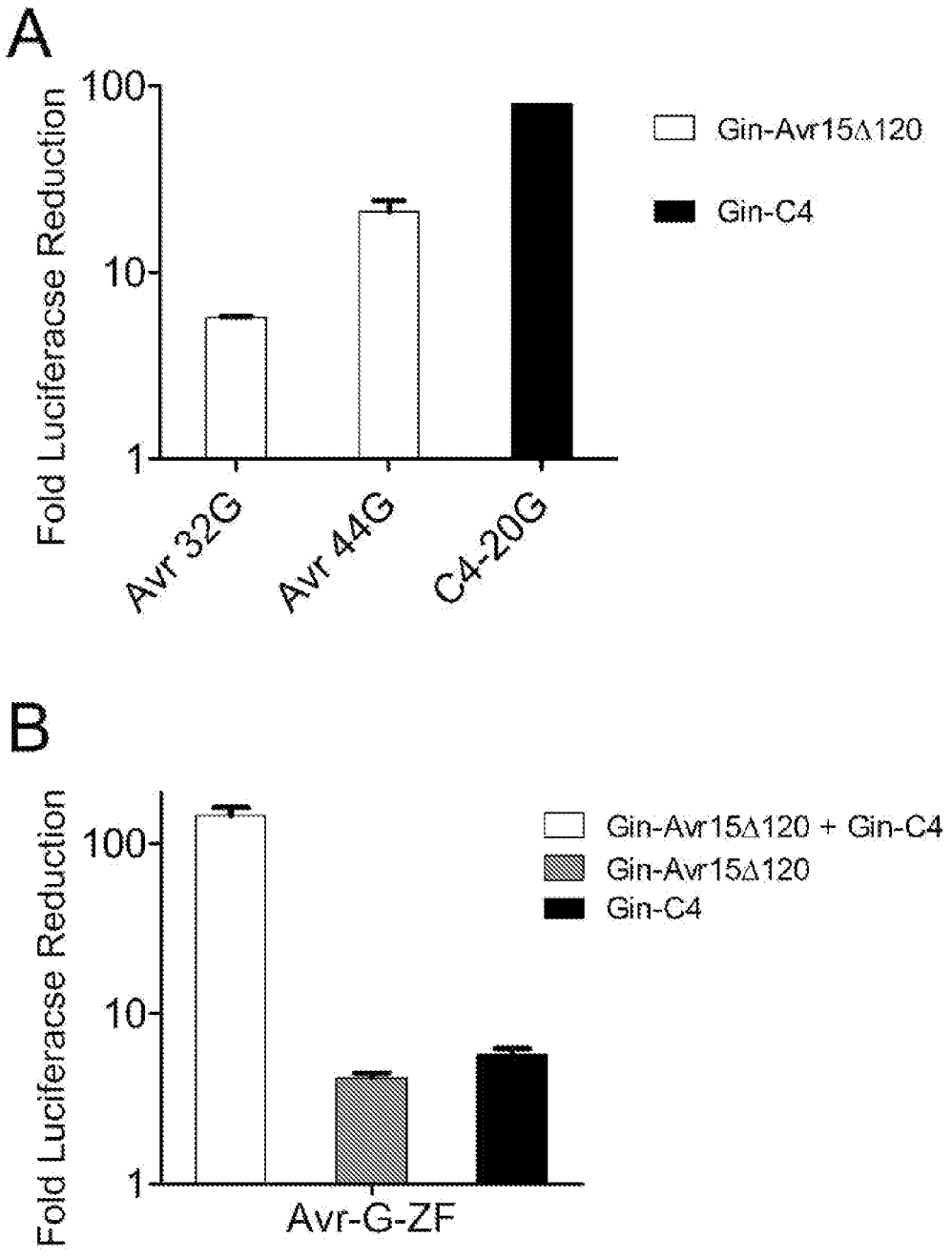

A
Fold Luciferacse Reduction
100
10
1
Avr 32G
Avr 44G
C4-20G
Gin-Avr15Δ120
Gin-C4
B
Fold Luciferacse Reduction
100
10
1
Avr-G-ZF
Gin-Avr15Δ120 + Gin-C4
Gin-Avr15Δ120
Gin-C4

FIG. 5

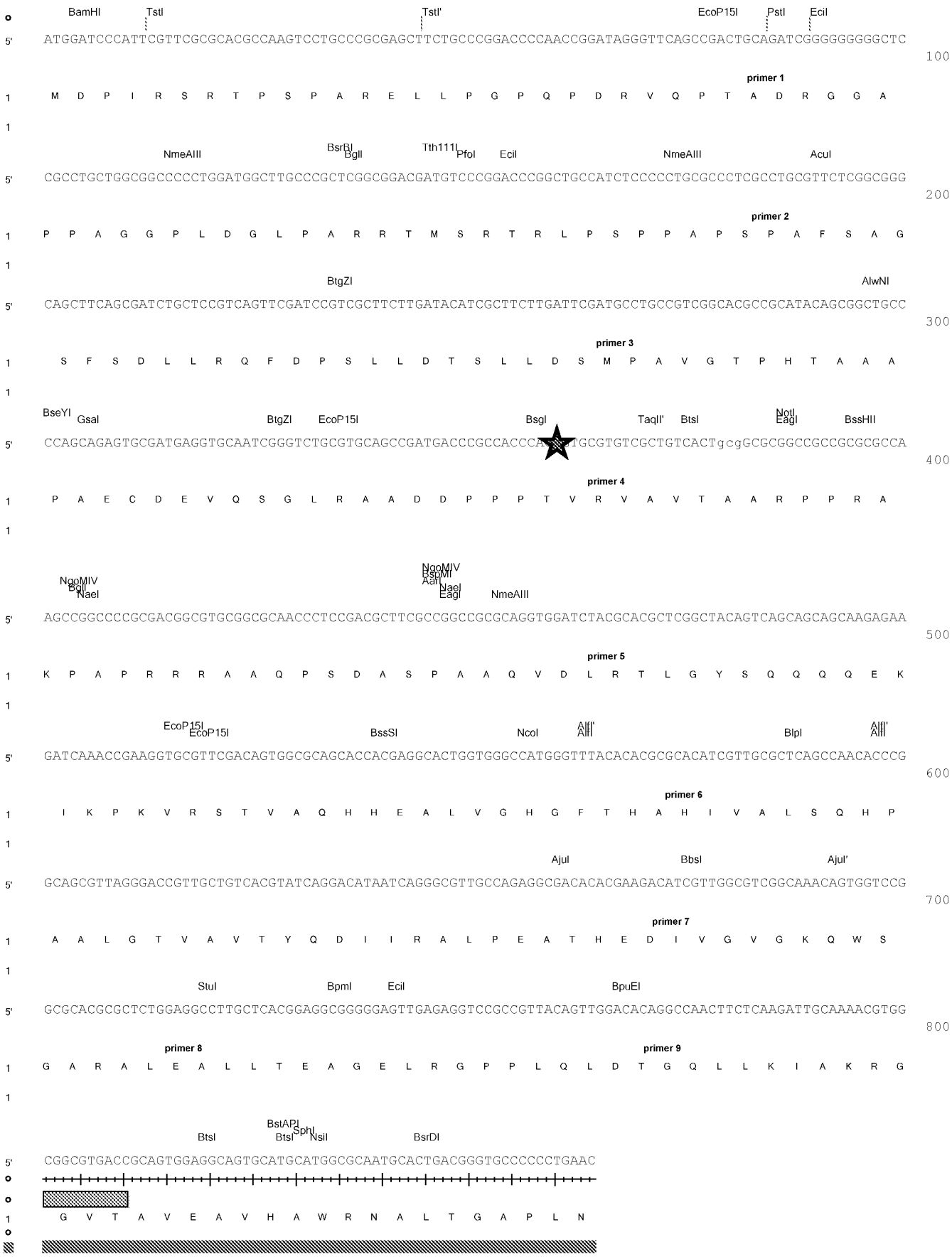

BamHI
TstI
TstI'
EcoP15I
PstI
EciI
ATGGATCCCATTCGTTCGCGCACGCCAAGTCCTGCCCGCGAGCTTCTGCCCGGACCCCAACCGGATAGGGTTCAGCCGACTGCAGATCGGGGGGGGGCTC
100
primer 1
M D P I R S R T P S P A R E L L P G P Q P D R V Q P T A D R G G A
NmeAIII
BsrBI
BglI
Tth111I
PfoI
EciI
NmeAIII
AcuI
CGCCTGCTGGCGGCCCCCTGGATGGCTTGCCCGCTCGGCGGACGATGTCCCGGACCCGGCTGCCATCTCCCCCTGCGCCCTCGCCTGCGTTCTCGGCGGG
200
primer 2
P P A G G P L D G L P A R R T M S R T R L P S P P A P S P A F S A G
BtgZI
AlwNI
CAGCTTCAGCGATCTGCTCCGTCAGTTCGATCCGTCGCTTCTTGATACATCGCTTCTTGATTCGATGCCTGCCGTCGGCACGCCGCATACAGCGGCTGCC
300
primer 3
S F S D L L R Q F D P S L L D T S L L D S M P A V G T P H T A A A
BseYI
GsaI
BtgZI
EcoP15I
BsgI
TaqII'
BtsI
NotI
EagI
BssHII
CCAGCAGAGTGCGATGAGGTGCAATCGGGTCTGCGTGCAGCCGATGACCCGCCACCCA[star]TGCGTGTCGCTGTCACTgcgGCGCGGCCGCCGCGCGCCA
400
primer 4
P A E C D E V Q S G L R A A D D P P P T V R V A V T A A R P P R A
NgoMIV
BglI
NaeI
NgoMIV
BsrDI
AarI
NaeI
EagI
NmeAIII
AGCCGGCCCCGCGACGGCGTGCGGCGCAACCCTCCGACGCTTCGCCGGCCGCGCAGGTGGATCTACGCACGCTCGGCTACAGTCAGCAGCAGCAAGAGAA
500
primer 5
K P A P R R R A A Q P S D A S P A A Q V D L R T L G Y S Q Q Q Q E K
EcoP15I
EcoP15I
BssSI
NcoI
AlfI'
AflI
BlpI
AlfI'
AflI
GATCAAACCGAAGGTGCGTTCGACAGTGGCGCAGCACCACGAGGCACTGGTGGGCCATGGGTTTACACACGCGCACATCGTTGCGCTCAGCCAACACCCG
600
primer 6
I K P K V R S T V A Q H H E A L V G H G F T H A H I V A L S Q H P
AjuI
BbsI
AjuI'
GCAGCGTTAGGGACCGTTGCTGTCACGTATCAGGACATAATCAGGGCGTTGCCAGAGGCGACACACGAAGACATCGTTGGCGTCGGCAAACAGTGGTCCG
700
primer 7
A A L G T V A V T Y Q D I I R A L P E A T H E D I V G V G K Q W S
StuI
BpmI
EciI
BpuEI
GCGCACGCGCTCTGGAGGCCTTGCTCACGGAGGCGGGGGAGTTGAGAGGTCCGCCGTTACAGTTGGACACAGGCCAACTTCTCAAGATTGCAAAACGTGG
800
primer 8
primer 9
G A R A L E A L L T E A G E L R G P P L Q L D T G Q L L K I A K R G
BstAPI
BtsI
BtsI
SphI
NsiI
BsrDI
CGGCGTGACCGCAGTGGAGGCAGTGCATGCATGGCGCAATGCACTGACGGGTGCCCCCCTGAAC
G V T A V E A V H A W R N A L T G A P L N

FIG. 6

AvrXa7	NI	HG	NI	NI	NS	HD	NN	HD	HD	NS	N*	N*	HD	HD	NS	NS	NN	NN	NI	NG	NN	NI	N*	NS	N*
Target Sequence	A	T	A	A	A	C	C	C	C	C	T	C	C	A	A	C	C	A	G	G	T	G	C	T	A	A
Synthetic	NI	NG	NI	NI	NI	HD	HD	HD	HD	HD	NG	HD	HD	NI	NI	HD	HD	NI	NN	NN	NG	NN	HD	NG	NI	NI

FIG. 7

TALE and TALER sequences

AvrXa7 protein:

MDPIRSRTPSPARELLPGPQPDRVQPTADRGGAPPAGGPLDGLPARRTMSRTRLPSPPAPSPA
FSAGSFSDLLRQFDPSLLDTSLLDSMPAVGTPHTAAAPAECDEVQSGLRAADDPPPTVRVAV
TAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGH
GFTHAHIVALSQHPAALGTVAVTYQDIIRALPEATHEDIVGVGKQWSGARALEALLTEAGEL
RGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNIGGKQALET
VQRLLPVLCQDHGLTPDQVVAIASHGGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIG
GKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPAQVV
AIASNSGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETLQRLLPVLCQDHG
LTPDQVVAIANNNGGKQALETLQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLP
VLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPAQVVAIASHDGGKQAL
ETVQRLLPVLCQDHGLTPDQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPDQVVAIASN
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGKQALETVQRLLPVQRLLPVLCQDHG
LTQDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLP
VLCQDHGLTPDQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNSGGKQALE
TVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIANN
NGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQDHGLTLDQ
VVAIASNGGSKQALETVQRLLPVLCQDHGLTPDQVVAIANNNGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTLDQVVAIASNGGKQALETVQRL
LPVLCQDHGLTPNQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPNQVVAIASNGGKQAL
ESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPELIRRINRRIPERTSHRV
PDLAHVVRVLGFFQSHSHPAQAFDDAMTQFEMSRHGLVQLFRRVGVTEFEARYGTLPPASQ
RWDRILQASGMKRAKPSPTSAQTPDQASLHAFADSLERDLDAPSPMHEGDQTRASSRKRSR
SDRAVTGPSTQQSFEVRVPEQQDALHLPLSWRVKRPRTRIGGGLPDPGTPIAADLAASSTVM
WEQDAAPFAGAADDFPAFNEEELAWLMELLPQSGSVGGTI

FIG. 8

AvrXa7 gene:

ATGGATCCCATTCGTTCGCGCACGCCAAGTCCTGCCCGCGAGCTTCTGCCCGGACCCCA
ACCGGATAGGGTTCAGCCGACTGCAGATCGGGGGGGGGCTCCGCCTGCTGGCGGCCCC
CTGGATGGCTTGCCCGCTCGGCGGACGATGTCCCGGACCCGGCTGCCATCTCCCCCTGC
GCCCTCGCCTGCGTTCTCGGCGGGCAGCTTCAGCGATCTGCTCCGTCAGTTCGATCCGTC
GCTTCTTGATACATCGCTTCTTGATTCGATGCCTGCCGTCGGCACGCCGCATACAGCGGC
TGCCCCAGCAGAGTGCGATGAGGTGCAATCGGGTCTGCGTGCAGCCGATGACCCGCCAC
CCACCGTGCGTGTCGCTGTCACTgcgGCGCGGCCGCCGCGCGCCAAGCCGGCCCCGCGAC
GGCGTGCGGCGCAACCCTCCGACGCTTCGCCGGCCGCGCAGGTGGATCTACGCACGCTC
GGCTACAGTCAGCAGCAGCAAGAGAAGATCAAACCGAAGGTGCGTTCGACAGTGGCGC
AGCACCACGAGGCACTGGTGGGCCATGGGTTTACACACGCGCACATCGTTGCGCTCAGC
CAACACCCGGCAGCGTTAGGGACCGTTGCTGTCACGTATCAGGACATAATCAGGGCGTT
GCCAGAGGCGACACACGAAGACATCGTTGGCGTCGGCAAACAGTGGTCCGGCGCACGC
GCTCTGGAGGCCTTGCTCACGGAGGCGGGGGAGTTGAGAGGTCCGCCGTTACAGTTGGA
CACAGGCCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGACCGCAGTGGAGGCAGTG
CATGCATGGCGCAATGCACTGACGGGTGCCCCCCTGAACCTGACCCCGGACCAAGTGGT
GGCCATCGCCAGCAATATTGGCGGCAAGCAGGCGCTGGAGACGGTACAGCGGCTGTTG
CCGGTGCTGTGCCAGGACCATGGCCTGACCCCGGACCAGGTCGTGGCCATCGCCAGCCA
TGGCGGCGGCAAGCAGGCGCTGGAGACGGTGC

FIG. 9

AGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACCCCGGACCAGGTGGTGGCC
ATCGCCAGCAATATTGGCGGCAAGCAGGCGCTAGAGACGGTGCAGCGGCTGTTGCCGG
TGCTGTGCCAGGCCCATGGCCTGACCCCGGACCAGGTCGTGGCCATCGCCAGCAATATT
GGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACC
ATGGCCTGACCCCGGCCCAGGTGGTGGCCATCGCCAGCAATAGTGGCGGCAAGCAGGC
GCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACCCCGG
ACCAAGTCGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGCTGCA
GCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACCCCGGACCAGGTCGTGGCCA
TCGCCAACAATAACGGCGGCAAGCAGGCGCTGGAGACGCTGCAGCGGCTGTTGCCGGT
GCTGTGCCAGGACCATGGCCTGACCCCGGACCAAGTGGTGGCCATCGCCAGCCACGATG
GCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCA
TGGCCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCG
CTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACCCCGGC
CCAAGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAG
CGGCTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACCCCGGACCAGGTGGTGGCCAT
CGCCAGCAATAGCGGCGGCAAGCAGGCGCTGGAGACGGTACAGCGGCTGTTGCCGGTG
CTGTGCCAGGACCATGGACTGACCCCGGACCAGGTCGTGGCCATCGCCAGCAATGGCG
GCAAGCAGGCGCTGGAGACGGTACAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGG
CCTGACCCCGGACCAGGTCGTGGCCATCGCCAGCAATGGCGGCAAGCAGGCGCTGGAG
ACGGTGCAGCGGCTGTTGCCGGTACAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGG
CCTGACCCAGGACCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTG
GAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACCCCGGACCA
AGTGGTGGCCATCGCCAGCCACGATGGCGGCAAACAGGCGCTGGAGACGGTGCAGCGG
CTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACCCCGGACCAGGTGGTGGCCATCGC
CAGCAATAGTGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTG
TGCCAGGACCATGGCCTGACCCCGGACCAAGTGGTGGCCATCGCCAGCAATAGTGGCG
GCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGG
CCTGACCCCGGACCAGGTGGTGGCCATCGCCAGCAATAACGGCGGCAAGCAGGCGCTG
GAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACCCCGGACCA
GGTCGTGGCCATCGCCAACAATAACGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGG
CTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACCCCGGCGCAGGTGGTGGCCATCGC
CAGCAATATTGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGT
GCCAGGACCATGGCCTGACCCTGGACCAGGTGGTGGCCATTGCCAGCAATGGCGGCAG
CAAACAGGCGCTAGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGC
CTGACCCCGGACCAAGTGGTGGCCATCGCCAACAATAACGGCGGCAAGCAGGCGCTGG
AGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACCCCGGACCAG
GTCGTGGCCATCGCCAGCAATATTGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGC
TGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACCCTGGACCAGGTGGTGGCCATCGCC
AGCAATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCC
AGGACCATGGCCTGACCCCGAACCAGGTGGTGGCCATCGCCAGCAATAGTGGCGGCAA
GCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGA
CCCCGAACCAGGTGGTGGCCATCGCCAGCAATGGCGGCAAGCAGGCGCTGGAGAGCAT
TGTTGCCCAGTTATCTCGCCCTGATCCGGCGTTGGCCGCGTTGACCAACGACCACCTCGT
CGCCTTGGCCTGCCTCGGCGGACGTCCTGCCCTGGATGCAGTGAAAAAGGGATTGCCGC
ACGCGCCGGAATTGATCAGAAGAATCAATCGCCGCATTCCCGAACGCACGTCCCATCGC
GTTCCCGACCTCGCGCACGTGGTTCGCGTGCTTGGTTTTTTCCAGAGCCACTCCCACCCA
GCGCAAGCATTCGATGACGCCATGACGCAGTTCGAGATGAGCAGGCACGGCTTGGTAC
AGCTCTTTCGCAGAGTGGGCGTCACCGAATTCGAAGCCCGCTACGGAACGCTCCCCCCA
GCCTCGCAGCGTTGGGACCGTATCCTCCAGGCATCAGGG

FIG. 9 (cont.)

ATGAAAAGGGCCAAACCGTCCCCTACTTCAGCTCAAACACCGGATCAGGCGTCTTTGCA
TGCATTCGCCGATTCGCTGGAGCGTGACCTTGATGCGCCCAGCCCAATGCACGAGGGAG
ATCAGACGCGGGCAAGCAGCCGTAAACGGTCCCGATCGGATCGTGCTGTCACCGGCCCC
TCCACACAGCAATCTTTCGAGGTGCGCGTTCCCGAACAGCAAGATGCGCTGCATTTGCC
CCTCAGCTGGAGGGTAAAACGCCCGCGTACCAGGATCGGGGGCGGCCTCCCGGATCCT
GGTACGCCCATCGCTGCCGACCTGGCAGCGTCCAGCACCGTGATGTGGGAACAAGATGC
GGCCCCCTTCGCAGGGGCAGCGGATGATTTCCCGGCATTCAACGAAGAGGAGCTCGCAT
GGTTGATGGAGCTATTGCCTCAGTCAGGCTCAGTCGGAGGGACGATCTGA

FIG. 9 (cont.)

Gin-AvrΔ74

MLIGYVRVSTNDQNTDLQRNALVCAGCEQIFEDKLSGTRTDRPGLKRALKRLQKGDTLVV
WKLDRLGRSMKHLISLVGELRERGINFRSLTDSIDTSSPMGRFFFYVMGALAEMERELIIERT
MAGLAAARNKGRIGGRPRKSGSGSPRQFDPSLLDTSLLDSMPAVGTPHTAAAPAECDEVQS
GLRAADDPPPTVRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKP
KVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVTYQDIIRALPEATHEDIVGVGKQ
WSGARALEALLTEAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTP
DQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHGGGKQALETVQRLLPVLC
QDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQ
RLLPVLCQDHGLTPAQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGK
QALETLQRLLPVLCQDHGLTPDQVVAIANNNGGKQALETLQRLLPVLCQDHGLTPDQVVAI
ASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGL
TPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNSGGKQALETVQRLLPV
LCQDHGLTPDQVVAIASNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGKQALETV
QRLLPVQRLLPVLCQDHGLTQDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAI
ASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNSGGKQALETVQRLLPVLCQDHGL
TPDQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPV
LCQDHGLTPDQVVAIANNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALET
VQRLLPVLCQDHGLTLDQVVAIASNGGSKQALETVQRLLPVLCQDHGLTPDQVVAIANNN
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTLDQV
VAIASNGGKQALETVQRLLPVLCQDHGLTPNQVVAIASNSGGKQALETVQRLLPVLCQDHG
LTPNQVVAIASNGGKQALESIVAQLSRPDPALAALTNDHLVALACLG

FIG. 10

Gin-AvrΔ87

MLIGYVRVSTNDQNTDLQRNALVCAGCEQIFEDKLSGTRTDRPGLKRALKRLQKGDTLVV
WKLDRLGRSMKHLISLVGELRERGINFRSLTDSIDTSSPMGRFFFYVMGALAEMERELIIERT
MAGLAAARNKGRIGGRPRKSGSGSPDSMPAVGTPHTAAAPAECDEVQSGLRAADDPPPTV
RVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHHE
ALVGHGFTHAHIVALSQHPAALGTVAVTYQDIIRALPEATHEDIVGVGKQWSGARALEALL
TEAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNIGG
KQALETVQRLLPVLCQDHGLTPDQVVAIASHGGGKQALETVQRLLPVLCQDHGLTPDQVV
AIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGL
TPAQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETLQRLLPV
LCQDHGLTPDQVVAIANNNGGKQALETLQRLLPVLCQDHGLTPDQVVAIASHDGGKQALE
TVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPAQVVAIASHD
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGKQALETVQRLLPVQRLLPV
LCQDHGLTQDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALE
TVQRLLPVLCQDHGLTPDQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNS
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIANNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQD
HGLTLDQVVAIASNGGSKQALETVQRLLPVLCQDHGLTPDQVVAIANNNGGKQALETVQR
LLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTLDQVVAIASNGGKQA
LETVQRLLPVLCQDHGLTPNQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPNQVVAIAS
NGGKQALESIVAQLSRPDPALAALTNDHLVALACLG

FIG. 11

Gin-AvrΔ120

MLIGYVRVSTNDQNTDLQRNALVCAGCEQIFEDKLSGTRTDRPGLKRALKRLQKGDTLVV
WKLDRLGRSMKHLISLVGELRERGINFRSLTDSIDTSSPMGRFFFYVMGALAEMERELIIERT
MAGLAAARNKGRIGGRPRKSGSGSTVRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDL
RTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVTYQDIIRA
LPEATHEDIVGVGKQWSGARALEALLTEAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHA
WRNALTGAPLNLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHGGGK
QALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAI
ASNIGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNSGGKQALETVQRLLPVLCQDHGLT
PDQVVAIASHDGGKQALETLQRLLPVLCQDHGLTPDQVVAIANNNGGKQALETLQRLLPVL
CQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALET
VQRLLPVLCQDHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNSG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGKQALETVQRLLPVLCQDHGLTPDQVV
AIASNGGKQALETVQRLLPVQRLLPVLCQDHGLTQDQVVAIASHDGGKQALETVQRLLPVL
CQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNSGGKQALET
VQRLLPVLCQDHGLTPDQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNG
GKQALETVQRLLPVLCQDHGLTPDQVVAIANNNGGKQALETVQRLLPVLCQDHGLTPAQV
VAIASNIGGKQALETVQRLLPVLCQDHGLTLDQVVAIASNGGSKQALETVQRLLPVLCQDH
GLTPDQVVAIANNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLL
PVLCQDHGLTLDQVVAIASNGGKQALETVQRLLPVLCQDHGLTPNQVVAIASNSGGKQALE
TVQRLLPVLCQDHGLTPNQVVAIASNGGKQALESIVAQLSRPDPALAALTNDHLVALACLG

FIG. 12

Gin-AvrΔ120*

MLIGYVRVSTNDQNTDLQRNALVCAGCEQIFEDKLSGTRTDRPGLKRALKRLQKGDTLVV
WKLDRLGRSMKHLISLVGELRERGINFRSLTDSIDTSSPMGRFFFYVMGALAEMERELIIERT
MAGLAAARNKGRIGGRPRKSGSGSTVRVAVTAARPPHAVAGPAAQVDLRTLGYSQQQQE
KIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVTYQDIIRALPEATHEDIVGV
GKQWSGARALEALLTEAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
LTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHGGGKQALETVQRLLP
VLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALE
TVQRLLPVLCQDHGLTPAQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHD
GGKQALETLQRLLPVLCQDHGLTPDQVVAIANNNGGKQALETLQRLLPVLCQDHGLTPDQ
VVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNSGGKQALETVQ
RLLPVLCQDHGLTPDQVVAIASNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGKQA
LETVQRLLPVQRLLPVLCQDHGLTQDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNSGGKQALETVQRLLPVLCQD
HGLTPDQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRL
LPVLCQDHGLTPDQVVAIANNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQA
LETVQRLLPVLCQDHGLTLDQVVAIASNGGSKQALETVQRLLPVLCQDHGLTPDQVVAIAN
NNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTLD
QVVAIASNGGKQALETVQRLLPVLCQDHGLTPNQVVAIASNSGGKQALETVQRLLPVLCQD
HGLTPNQVVAIASNGGKQALESIVAQLSRPDPALAALTNDHLVALACLG

FIG. 13

Gin-AvrΔ147

MLIGYVRVSTNDQNTDLQRNALVCAGCEQIFEDKLSGTRTDRPGLKRALKRLQKGDTLVV
WKLDRLGRSMKHLISLVGELRERGINFRSLTDSIDTSSPMGRFFFYVMGALAEMERELIIERT
MAGLAAARNKGRIGGRPRKSGSGSPASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHHE
ALVGHGFTHAHIVALSQHPAALGTVAVTYQDIIRALPEATHEDIVGVGKQWSGARALEALL
TEAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNIGG
KQALETVQRLLPVLCQDHGLTPDQVVAIASHGGGKQALETVQRLLPVLCQDHGLTPDQVV
AIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGL
TPAQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETLQRLLPV
LCQDHGLTPDQVVAIANNNGGKQALETLQRLLPVLCQDHGLTPDQVVAIASHDGGKQALE
TVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPAQVVAIASHD
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGKQALETVQRLLPVQRLLPV
LCQDHGLTQDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALE
TVQRLLPVLCQDHGLTPDQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNS
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIANNNGGKQALETVQRLLPVLCQDHGLTPAQVVAIASNIGGKQALETVQRLLPVLCQD
HGLTLDQVVAIASNGGSKQALETVQRLLPVLCQDHGLTPDQVVAIANNNGGKQALETVQR
LLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTLDQVVAIASNGGKQA
LETVQRLLPVLCQDHGLTPNQVVAIASNSGGKQALETVQRLLPVLCQDHGLTPNQVVAIAS
NGGKQALESIVAQLSRPDPALAALTNDHLVALACLG

FIG. 14

GinAvr15Δ128-synthetic protein

MLIGYVRVSTNDQNTDLQRNALVCAGCEQIFEDKLSGTRTDRPGLKRALKRLQKGDTLVV
WKLDRLGRSMKHLISLVGELRERGINFRSLTDSIDTSSPMGRFFFYVMGALAEMERELIIERT
MAGLAAARNKGRIGGRPRKSGSGSPALRPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQ
QQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVTYQHIITALPEATHED
IVGVGKQWSGARALEALLTDAGELRGPPLQLDTGQLVKIAKRGGVTAMEAVHASRNALTG
APLNLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGGKQALETVQ
RLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQ
ALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIAS
HDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTP
DQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVL
CQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNGGGKQALET
VQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVV
AIASNIGGKQALESIVAQLSRPDPALAALTNDHLVALACLGPKKKRKV

FIG. 15

Gin-Avr15Δ128-synthetic DNA:

ATGCTGATTGGCTATGTAAGGGTATCAACAAATGACCAGAATACAGACCTGCAACGAA
ACGCTCTTGTTTGTGCAGGATGTGAACAAATATTTGAAGATAAATTAAGCGGAACAAGG
ACAGACCGACCGGGATTAAAACGCGCTTTAAAGCGCCTTCAAAAAGGTGACACACTGG
TTGTCTGGAAACTGGATCGCCTCGGGCGAAGCATGAAACATTTGATTTCTCTCGTAGGG
GAATTACGAGAGCGAGGGATTAATTTTCGCAGTCTTACTGACAGTATTGATACGTCATC
TCCAATGGGGCGTTTTTTCTTCTACGTTATGGGTGCCCTGGCTGAAATGGAACGAGAACT
AATTATCGAGCGAACGATGGCTGGACTTGCTGCCGCCAGAAATAAAGGCCGTATTGGAG
GTCGCCCGCGTAAATCGGGGTCTGGATCCCCCGCGCGGCCGCCGCGCGCCAAGCCGGCC
CCGCGACGGCGTGCTGCGCAACCCTCCGACGCTTCGCCGGCCGCGCAGGTGGATCTACG
CACGCTCGGCTACAGTCAGCAGCAGCAAGAGAAGATCAAACCGAAGGTGCGTTCGACA
GTGGCGCAGCACCACGAGGCACTGGTGGGCCATGGGTTTACACACGCGCACATCGTTGC
GCTCAGCCAACACCCGGCAGCGTTAGGGACCGTCGCTGTCACGTATCAGCACATAATCA
CGGCGTTGCCAGAGGCGACACACGAAGACATCGTTGGCGTCGGCAAACAGTGGTCCGG
CGCACGCGCCCTGGAGGCCTTGCTCACGGATGCGGGGGAGTTGAGAGGTCCGCCGTTAC
AGTTGGACACAGGCCAACTTGTGAAGATTGCAAAACGTGGCGGCGTGACCGCAATGGA
GGCAGTGCATGCATCGCGCAATGCACTGACGGGTGCCCCCCTGGAGCTGACTCCGGACC
AAGTGGTGGCTATCGCCAGCAACATTGGCGGCAAGCAAGCGCTCGAAACGGTGCAGCG
GCTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACTCCGGACCAAGTGGTGGCTATCG
CCAGCAACGGTGGCGGCAAGCAAGCGCTCGAAACGGTGCAGCGGCTGTTGCCGGTGCT
GTGCCAGGACCATGGCCTGACTCCGGACCAAGTGGTGGCTATCGCCAGCAACATTGGCG
GCAAGCAAGCGCTCGAAACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGG
CCTGACTCCGGACCAAGTGGTGGCTATCGCCAGCAACATTGGCGGCAAGCAAGCGCTCG
AAACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACTCCGGACCAA
GTGGTGGCTATCGCCAGCAACATTGGCGGCAAGCAAGCGCTCGAAACGGTGCAGCGGC
TGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACTCCGGACCAAGTGGTGGCTATCGCC
AGCCACGATGGCGGCAAGCAAGCGCTCGAAACGGTGCAGCGGCTGTTGCCGGTGCTGT
GCCAGGACCATGGCCTGACTCCGGACCAAGTGGTGGCTATCGCCAGCCACGATGGCGG
CAAGCAAGCGCTCGAAACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGC
CTGACTCCGGACCAAGTGGTGGCTATCGCCAGCCACGATGGCGGCAAGCAAGCGCTCG
AAACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACTCCGGACCAA
GTGGTGGCTATCGCCAGCCACGATGGCGGCAAGCAAGCGCTCGAAACGGTGCAGCGGC
TGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACTCCGGACCAAGTGGTGGCTATCGCC
AGCCACGATGGCGGCAAGCAAGCGCTCGAAACGGTGCAGCGGCTGTTGCCGGTGCTGT
GCCAGGACCATGGCCTGACCCCGGACCAAGTGGTGGCTATCGCCAGCAACGGTGGCGG
CAAGCAAGCGCTCGAAACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGC
CTGACTCCGGACCAAGTGGTGGCTATCGCCAGCCACGATGGCGGCAAGCAAGCGCTCG
AAACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACTCCGGACCAA
GTGGTGGCTATCGCCAGCCACGATGGCGGCAAGCAAGCGCTCGAAACGGTGCAGCGGC
TGTTGCCGGTGCTGTGCCAGGACCATGGCCTGACCCCGGACCAAGTGGTGGCTATCGCC
AGCAACATTGGCGGCAAGCAAGCGCTCGAAACGGTGCAGCGGCTGTTGCCGGTGCTGT
GCCAGGACCATGGCCTGACCCCGGACCAAGTGGTGGCTATCGCCAGCAACATTGGCGG
CAAGCAAGCGCTCGAAAGCATTGTGGCCCAGCTGAGCCGGCCTGATCCGGCGTTGGCCG
CGTTGACCAACGACCACCTCGTCGCCTTGGCCTGCCTCGGCCCCAAGAAGAAGCGCAAG
GTGTAG

FIG. 16

GinAvr15Δ128-synthetic protein:

MLIGYVRVSTNDQNTDLQRNALVCAGCEQIFEDKLSGTRTDRPGLKR
ALKRLQKGDTLVVWKLDRLGRSMKHLISLVGELRERGINFRSLTDSI
DTSSPMGRFFFYVMGALAEMERELIIERTMAGLAAARNKGRIGGRPR
*KSGSGS*PALRPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQE
KIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVTYQHI
ITALPE
ATHEDIVGVGKQWSGARALEALLTDAGELRGPPLQ (-1)
LDTGQLVKIAKRGGVTAMEAVHASRNALTGAPLN (0)T
LTPDQVVAIAS**NI**GGKQALETVQRLLPVLCQDHG (1)A
LTPDQVVAIAS**NG**GGKQALETVQRLLPVLCQDHG (2)T
LTPDQVVAIAS**NI**GGKQALETVQRLLPVLCQDHG (3)A
LTPDQVVAIAS**NI**GGKQALETVQRLLPVLCQDHG (4)A
LTPDQVVAIAS**NI**GGKQALETVQRLLPVLCQDHG (5)A
LTPDQVVAIAS**HD**GGKQALETVQRLLPVLCQDHG (6)C
LTPDQVVAIAS**HD**GGKQALETVQRLLPVLCQDHG (7)C
LTPDQVVAIAS**HD**GGKQALETVQRLLPVLCQDHG (8)C
LTPDQVVAIAS**HD**GGKQALETVQRLLPVLCQDHG (9)C
LTPDQVVAIAS**HD**GGKQALETVQRLLPVLCQDHG (10)C
LTPDQVVAIAS**NG**GGKQALETVQRLLPVLCQDHG (11)T
LTPDQVVAIAS**HD**GGKQALETVQRLLPVLCQDHG (12)C
LTPDQVVAIAS**HD**GGKQALETVQRLLPVLCQDHG (13)C
LTPDQVVAIAS**NI**GGKQALETVQRLLPVLCQDHG (14)A
LTPDQVVAIAS**NI**GGKQALE (15)A
SIVAQLSRPDPALAALTNDHLVALACLGPKKKRKV

Key:
Gin Recombinase (italics)
TALE N-terminal framework (double underline)
RVD (DNA binding domains) (emboldened)
TALE C-terminal framework (dotted underline)

FIG. 17

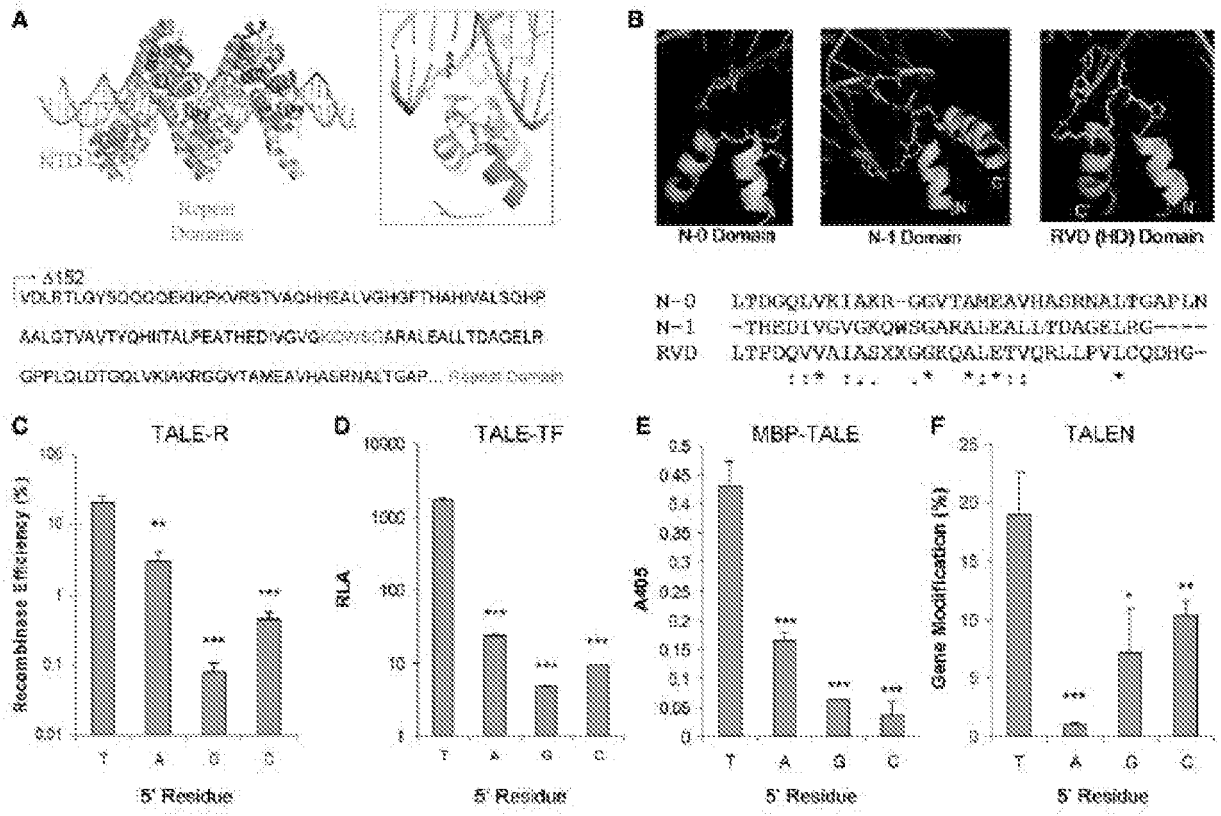
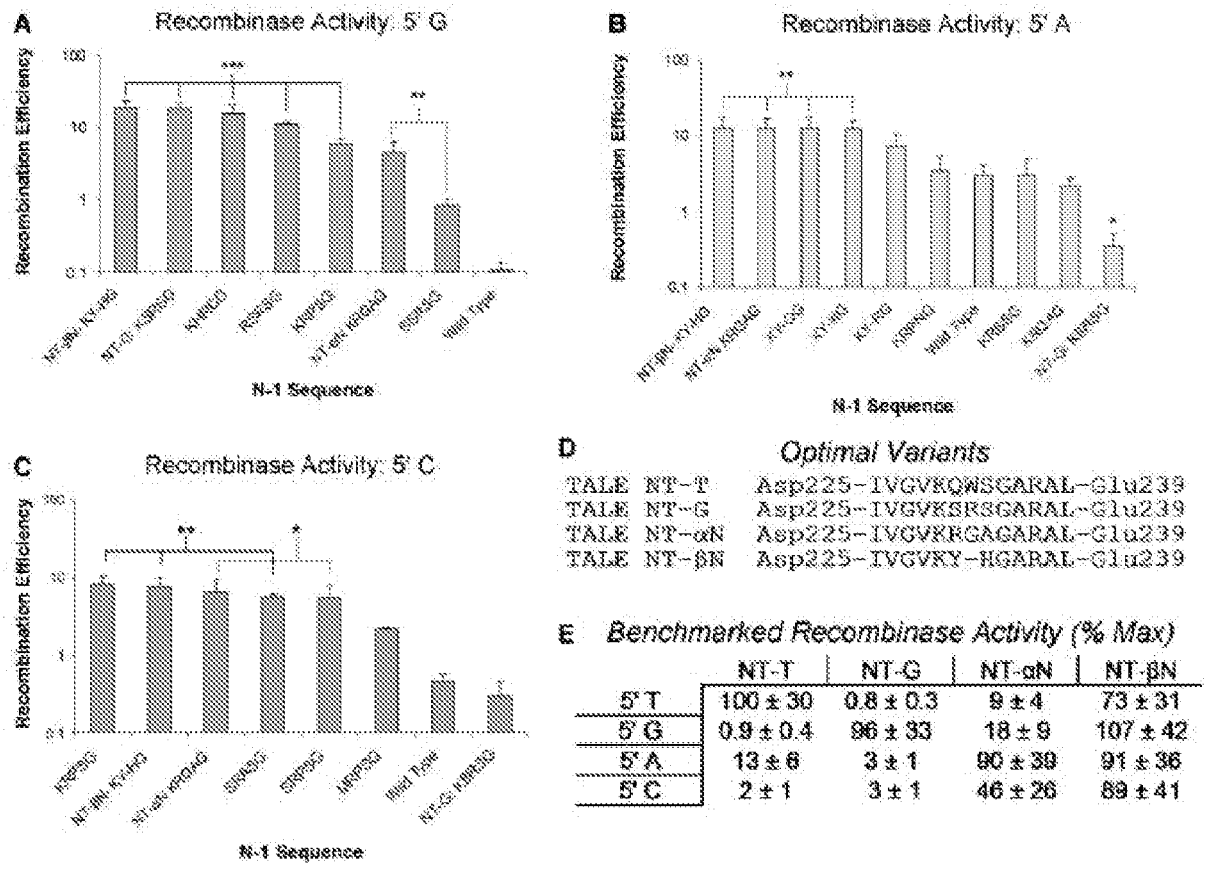


FIG. 18



	NT-T	NT-G	NT-αN	NT-βN
5' T	100 ± 30	0.8 ± 0.3	9 ± 4	73 ± 31
5' G	0.9 ± 0.4	96 ± 33	18 ± 9	107 ± 42
5' A	13 ± 6	3 ± 1	90 ± 39	91 ± 36
5' C	2 ± 1	3 ± 1	46 ± 26	89 ± 41

FIG. 19

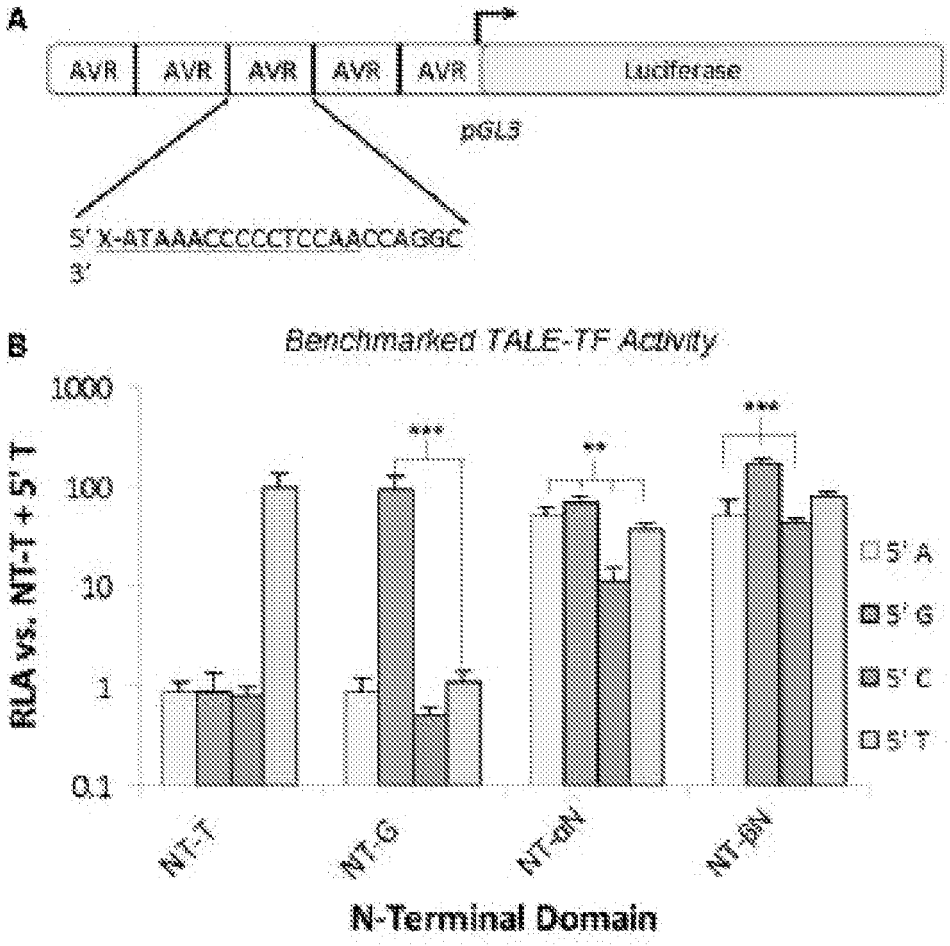
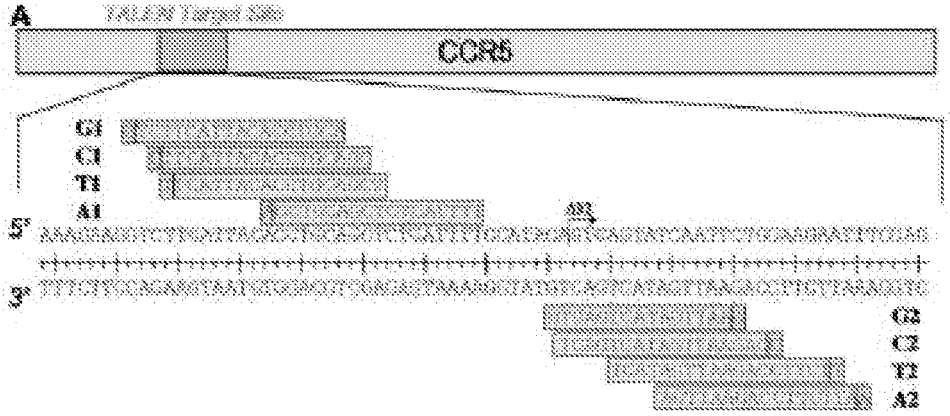

A
AVR
AVR
AVR
AVR
AVR
Luciferase
pGL3
5' X-ATAAACCCCCTCCAACCAGGC
3'
B
Benchmarked TALE-TF Activity
RLA vs. NT-T + 5' T
1000
100
10
1
0.1
***
**
***
5' A
5' G
5' C
5' T
NT-T
NT-G
NT-dN
NT-pN
N-Terminal Domain

FIG. 20

A



B *TALEN Gene Editing Efficiency*

TALEN Pair	NT-T	NT-G	NT-αN	NT-βN	dHax3
T1/T2	18.9 ± 3.7	9.5 ± 2.7	14.5 ± 2.3	11.2 ± 0.3	22.1 ± 2.2
G1/G2	7.1 ± 3.8	17.2 ± 2.8	14.4 ± 1.7	25.1 ± 3.2	8.25 ± 2.9
A1/A2	1 ± 0.2	0.7 ± 0.4	9.1 ± 1.1	0.4 ± 0.2	1.2 ± 0.5
C1/C2	10.4 ± 1.2	9.1 ± 0.4	21.1 ± 4.5	15.1 ± 2.3	11.3 ± 3.1

C *TALEN NTD Activity Enhancement vs. NT-T*

TALEN Pair	NT-T	NT-G	NT-αN	NT-βN	dHax3
T1/T2	100 ± 19	50 ± 14	77 ± 12	59 ± 2	117 ± 12
G1/G2	100 ± 51	242 ± 15	203 ± 23	354 ± 45	116 ± 41
A1/A2	100 ± 20	70 ± 40	910 ± 110	40 ± 20	120 ± 50
C1/C2	100 ± 12	88 ± 4	203 ± 43	145 ± 20	109 ± 30

FIG. 21

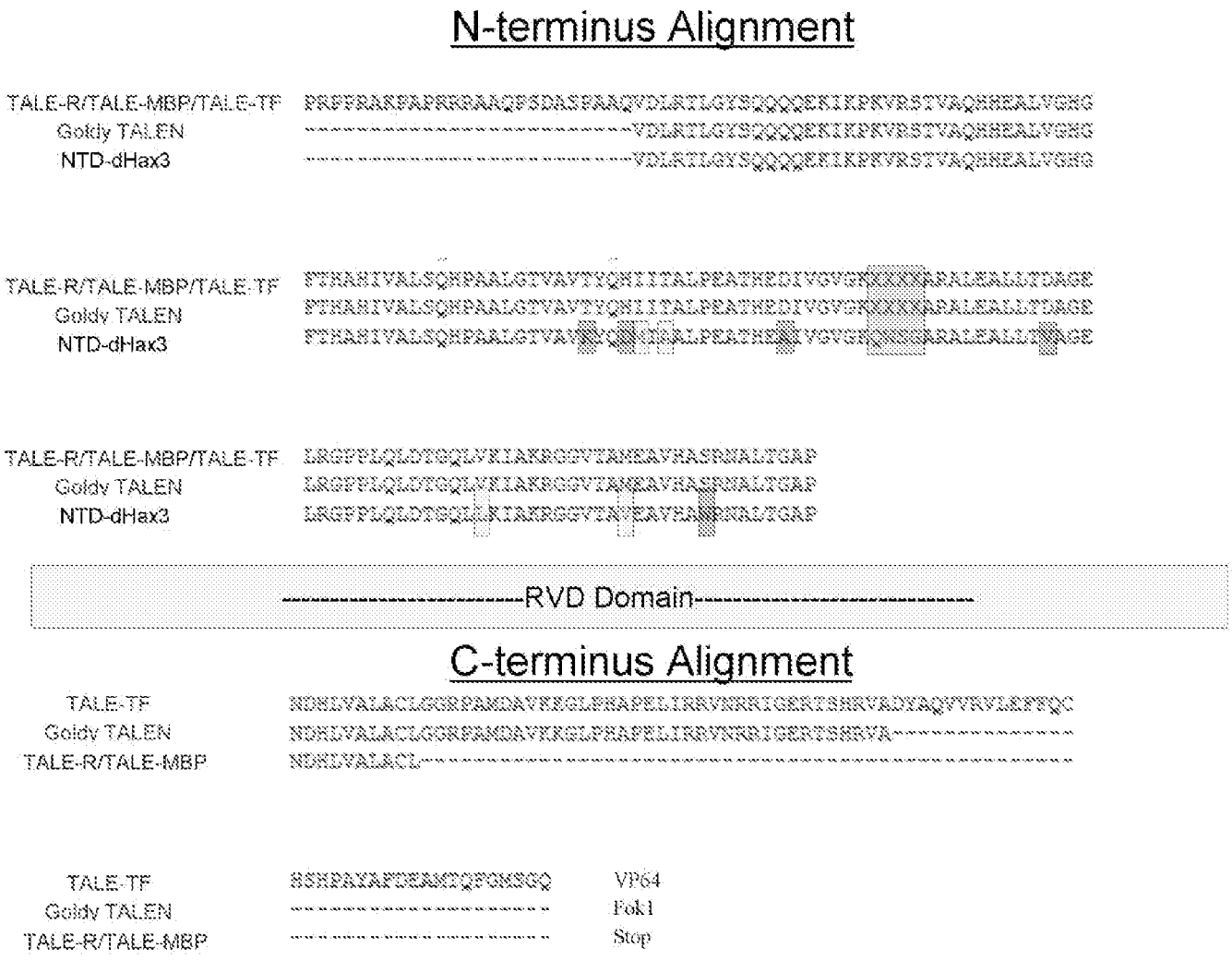

N-terminus Alignment
TALE-R/TALE-MBP/TALE-TF
Goldy TALEN
NTD-dHax3
RVD Domain
C-terminus Alignment
TALE-TF
Goldy TALEN
TALE-R/TALE-MBP
VP64
FokI
Stop

FIG. 22

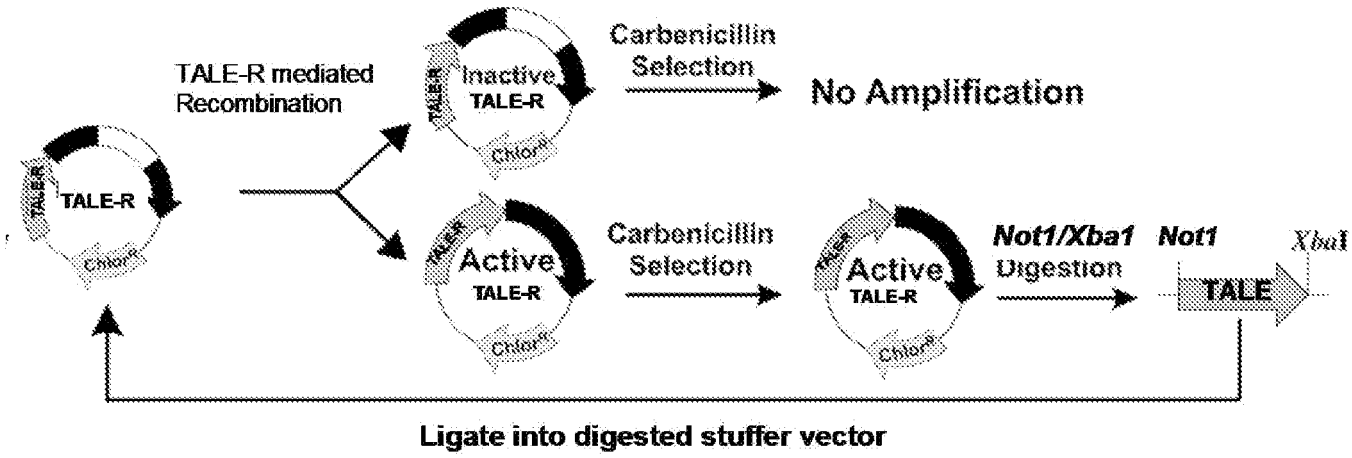

TALE-R mediated
Recombination
TALE-R
Inactive
TALE-R
Active
TALE-R
Carbenicillin
Selection
No Amplification
Carbenicillin
Selection
Active
TALE-R
Not1/Xba1
Digestion
Not1
Xba1
TALE
Ligate into digested stuffer vector

FIG. 23

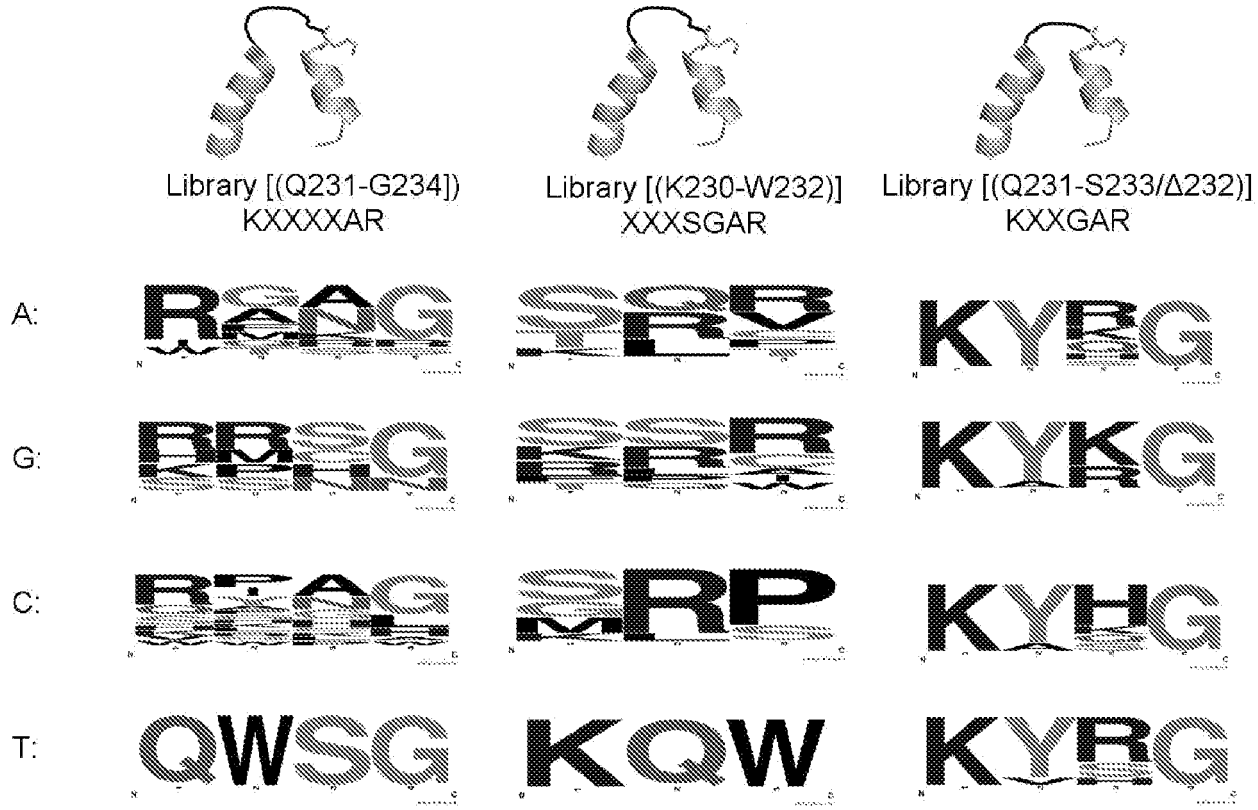


FIG. 24

```
TALE-G          QVDLETLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVTYQ 60

BrgII           ---MAALGYSREQIRRLRQESLSGVAKYHAPLTRHGFTHTDICRISRRWQSLRMVARNVP 57
                   : :****::* .*:* :  * **:* .*. *****:.*  :*::  :*  ** .*

TALE-G          HITTALPEATHEDIVGVGKQWSGARALEALLTDAGELRGPPLQLDTGQLVKIAKRGGVTA 120

BrgII           KLTAALPDLTRTHIVDIAQNSGALEALLPVATALAAAPLRLRASQIAIIAQCGERPA 117
                ::*:***: *: .**.:.*. : *  ******. *  * ..**:* :.*:. **: *  .*

TALE-G          MEAVHASRNALTGAP 135

BrgII           ILALSRLPRKLTGAP 132
                : *:*  *. *****
```

FIG. 25

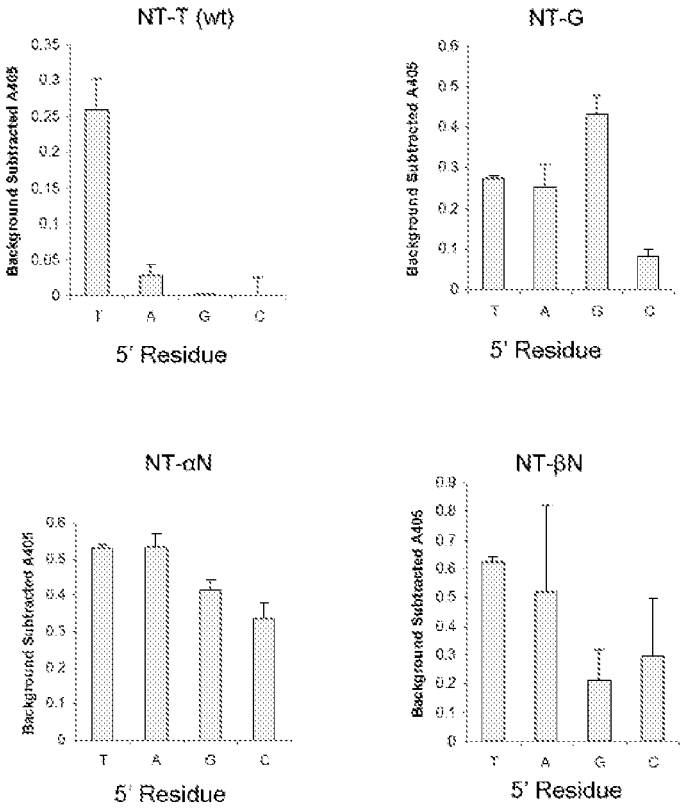
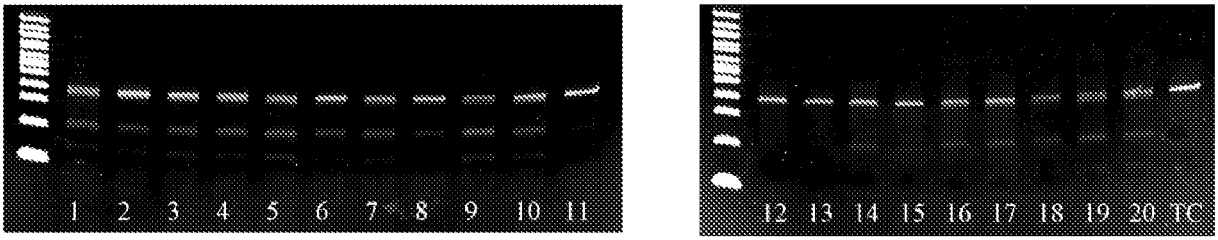


FIG. 26



TALEN Pair	NTD	% Editing	Indels
1: T1/T2	NT-T	16.6%	
2: T1/T2	NT-G	7.5%	
3: T1/T2	NT-βN	11.3%	
4: T1/T2	NT-αN	12.9%	
5: T1/T2	NT-dHax3	20.1%	
6: G1/G2	NT-T	4.4%	
7: G1/G2	NT-G	15.2%	(9/30)
8: G1/G2	NT- αN	4.9%	3x Ave = 14.4
9: G1/G2	NT- βN	28.9%	(10/30)
10: G1/G2	NT-dHax3	6.2%	
11: A1/A2	NT-T	1%	
12: A1/A2	NT-G	0.7%	
13: A1/A2	NT- βN	0.4%	
14: A1/A2	NT- αN	9.1%	(3/30)
15: A1/A2	NT-dHax3	1.2%	
16: C1/C2	NT-T	8.6%	
17: C1/C2	NT-G	8.9%	
18: C1/C2	NT- βN	13.4%	(9/30)
19: C1/C2	NT- αN	17.9%	(8/30)
20: C1/C2	NT-dHax3	9.1%	

FIG. 27

TALEN Pair 7 (9/30)

CCR5 (x20) TGGAAATTCTTCCAGAATTGATACTGACTGTATGGAAAATGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-6 (x3) TGGAAATTCTTCCATAATTGATATTGACTGTATGGA------AGGCTGCGGGTGTAATGAATACCTTCTTTTTGAGATCTGGT
-6 TGGAAATTCTTCCAGAATTGATACTGACTGTATGGAAAA------CTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-6 TGGAAATTCTTCCAGAATTGATACTGACTGTATGGAA------AGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-7 TGGAAATTCTTCCAGAATTGATACTGACTGTATGGAA-------GCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-10 TGGAAATTCTTCCAGAATTGATACTGACTGTAG----------AGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-11 TGGAAATTCTTCCAGAATTGATACTGACTGTAT----------GCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-13 TGGAAATTCTTCCAGAATTGATACTGACTGT----------GCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT

TALEN pair 9 (10/30)
(CCR5 WT X 21)

TGGAAATTCTTCCAGAATTGATACTGACTGTATGGAAAATGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-4 TGGAAATTCTTCCAGAATTGATACTGACTGTATGGAAA----GAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-5 TGGAAATTCTTCCAGAATTGATACTGACTGTATGGAA-----GAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-6 TGGAAATTCTTCCAGAATTGATACTGACTGTATGGA------GAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-9 TGGAAATTCTTCCAGAATTGATACTGACT---------ATGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-10 (x2) TGGAAATTCTTCCAGAATTGATACTGACTGTA----------GAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-11 TGGAAATTCTTCCAGAATTGATACTGAC----------TGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-11 TGGAAATTCTTCCAGGATTGATACTGACT----------GAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-11 TGGAAATTCTTCCAGAATTGATACTGACT----------GAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-11 TGGAAATTCTTCCAGAATTGATACTGACT----------AGGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT

TALEN pair 14 (3/30)
(CCR5 WT X27)

TGGAAATTCTTCCAGAATTGATACTGACTGTATGGAAAATGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-2 TGGAAATTCTTCCAGAATTGATACTGACT--ATGGAAAATGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-3 TGGAAATTCTTCCTCAATTGATACTGA---TATGGAAAATGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-9 TGGAAATTCTTCCAGAATTGATA---------TGGAAAATGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT

TALEN Pair 18 (9/30)
(CCR5 WT X 20)

TGGAAATTCTTCCAGAATTGATACTGACTGTATGGAAAATGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-4 TGGAAATTCTTCCAGAATTGATACTGACTGTATGA----TGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-4 TGGAAATTCTTCCAGAATTGATACTGACTGTATGG----TGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-5 TGGAAATTCTTCCAGAATTGATACTGACTGTA-----AATGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-5 TGGAAATTCTTCCAGAATTGATACTGACTGTATG-----TGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-9 TGGAAATTCTTCCAGAATTGATACTGA---------AAATGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-9 TGGAAATTCTTCCAGAATTGATACTGACTGTATGGA---------CTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-10 TGGAAATTCTTCCAGAATTGATACTGACTG---------TGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
+3/-10 TGGAAATTCTTCCAGAATTGATACTGgta----------TGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-12 TGGAAATTCTTCCAGAATTGATACTGACTGT----------AGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-

TALEN Pair 19 (8/30)
(CCR5 WT (x 23)

TGGAAATTCTTCCAGAATTGATACTGACTGTATGGAAAATGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-7 TGGAAATTCTTCCAGAATTGATACTGACTGTATGGAAA-------CTGC-GGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-7 TGGAAATTCTTCCAGAATTGATACTGACTGTA-------TGAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-9 TGGAAATTCTTCCAGAATTAATACTGACTGT---------GAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-10 TGGAAATTCTTCCAGAATTGATACTGACTG----------GAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-11 TGGAAATTCTTCCAGAATTGATACTGACT----------GAGAGCTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-15 TGGAAATTCTTCCAGAATTGATACTGACTG----------CTGCAGGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-18 TGGAAATTCTTCCAGAATTGATACTGACTGT----------ATGTGTAATGAAGACCTTCTTTTTGAGATCTGGT
-44 TGGAAATTCTTCCAGAATTGATA----------CTTTTTGAGATCTGGT

FIG. 28

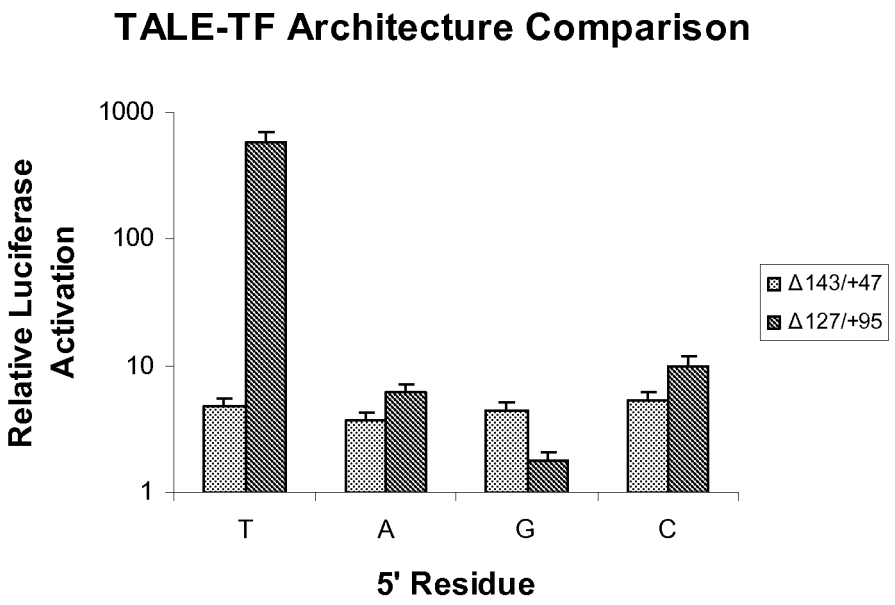

TALE-TF Architecture Comparison
Relative Luciferase Activation
1000
100
10
1
T
A
G
C
5' Residue
Δ143/+47
Δ127/+95

FIG. 29

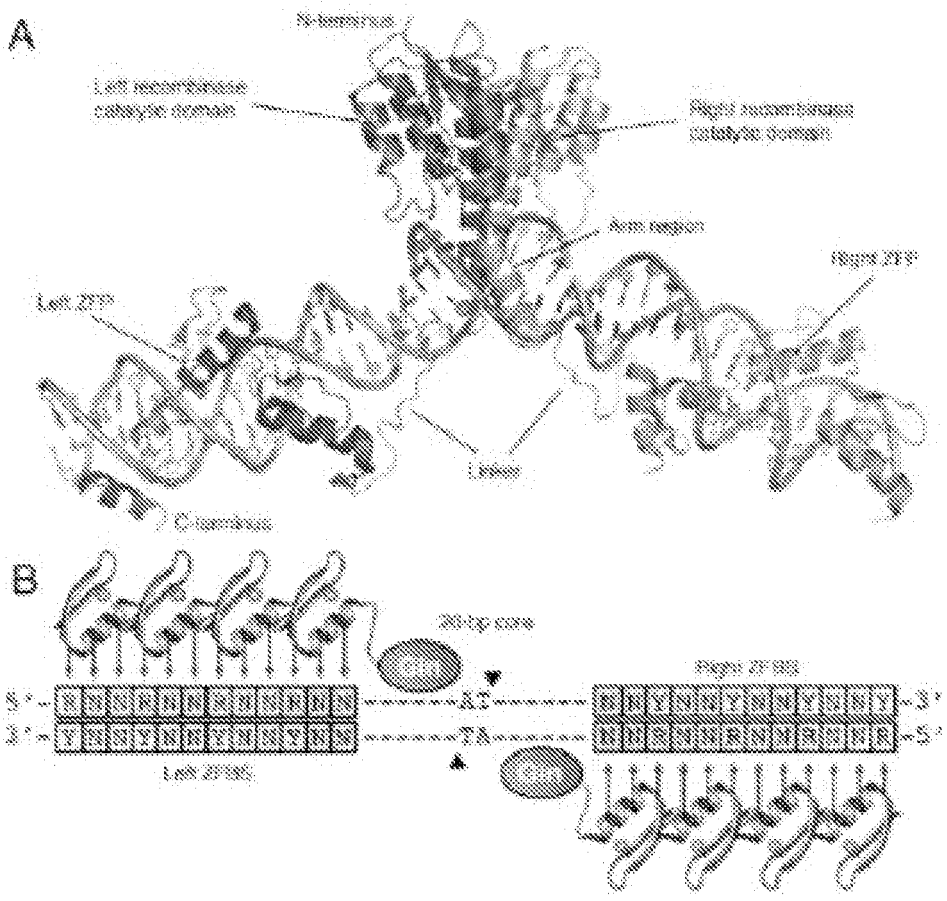

A
B

FIG. 30

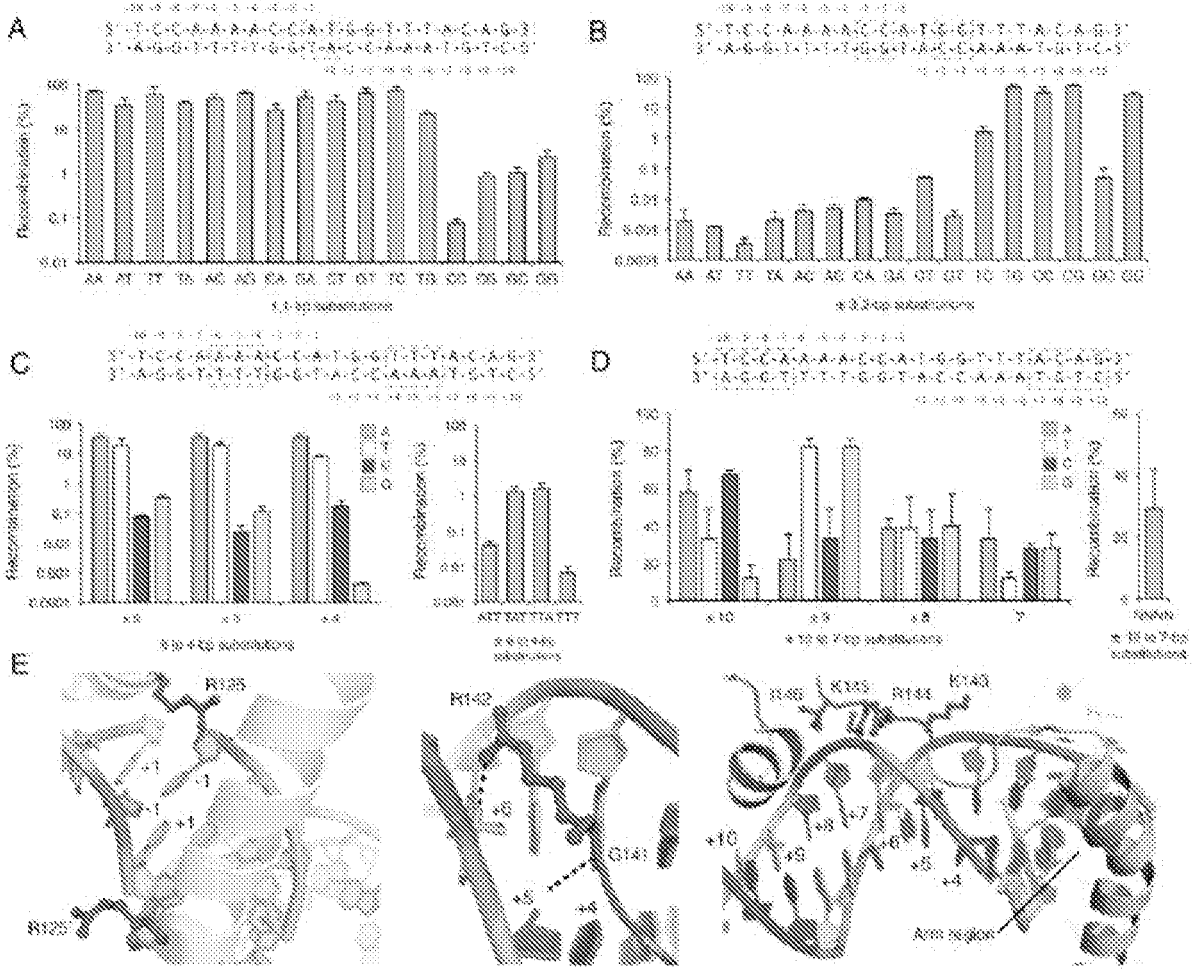


FIG. 31

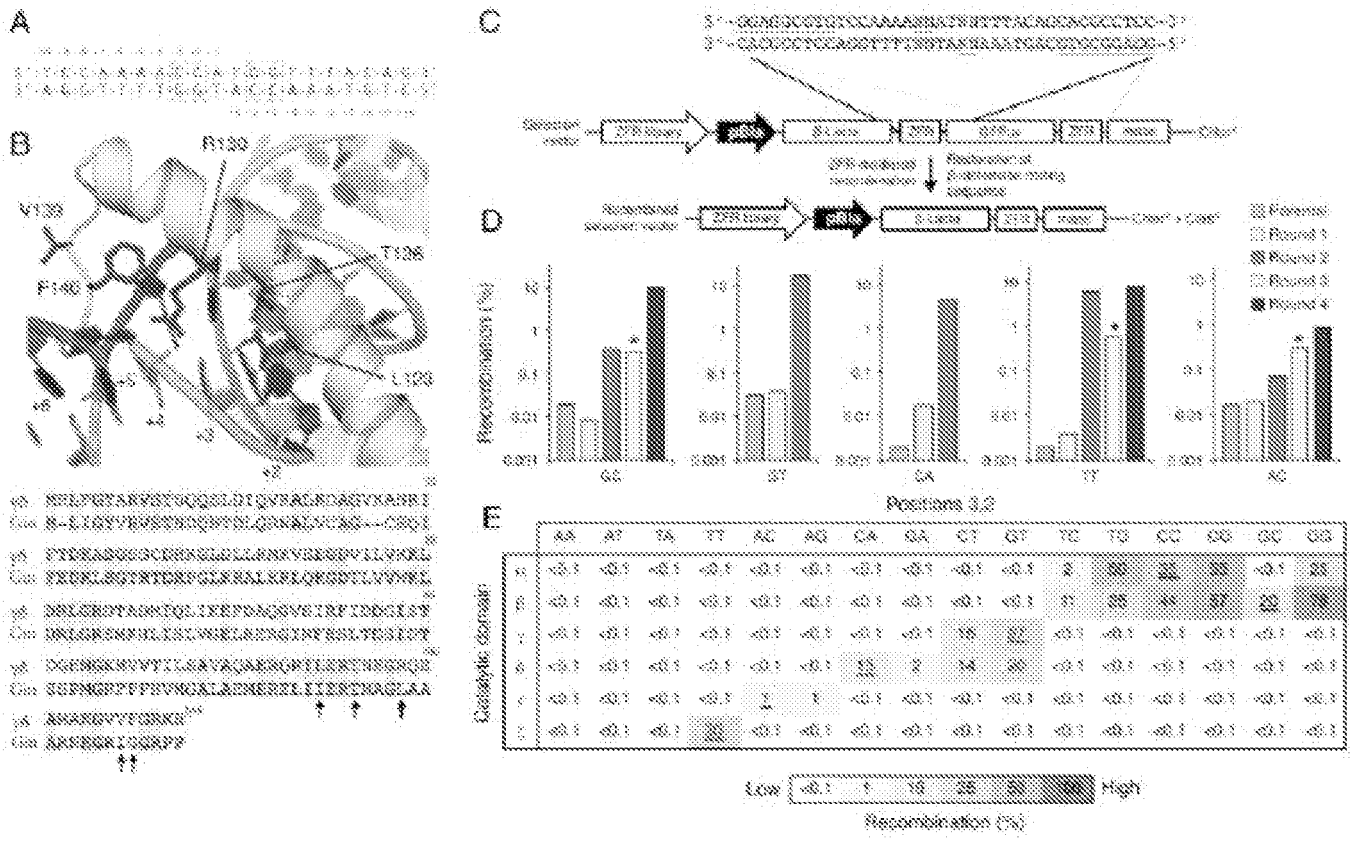


FIG. 32

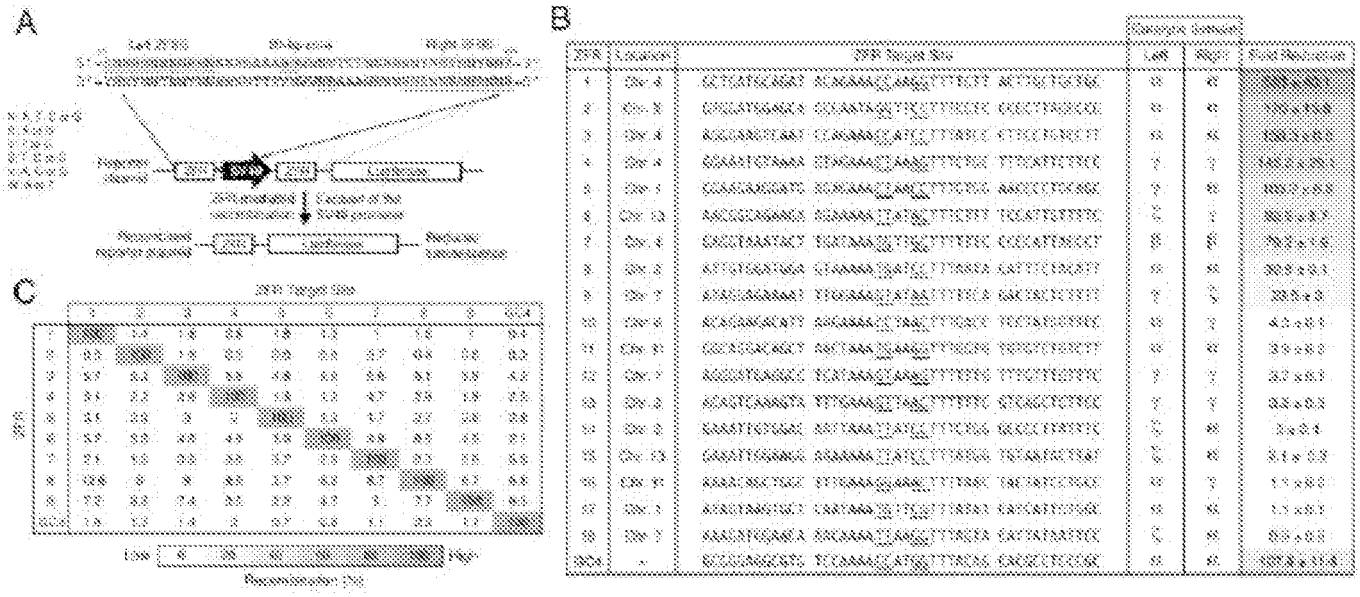


FIG. 33

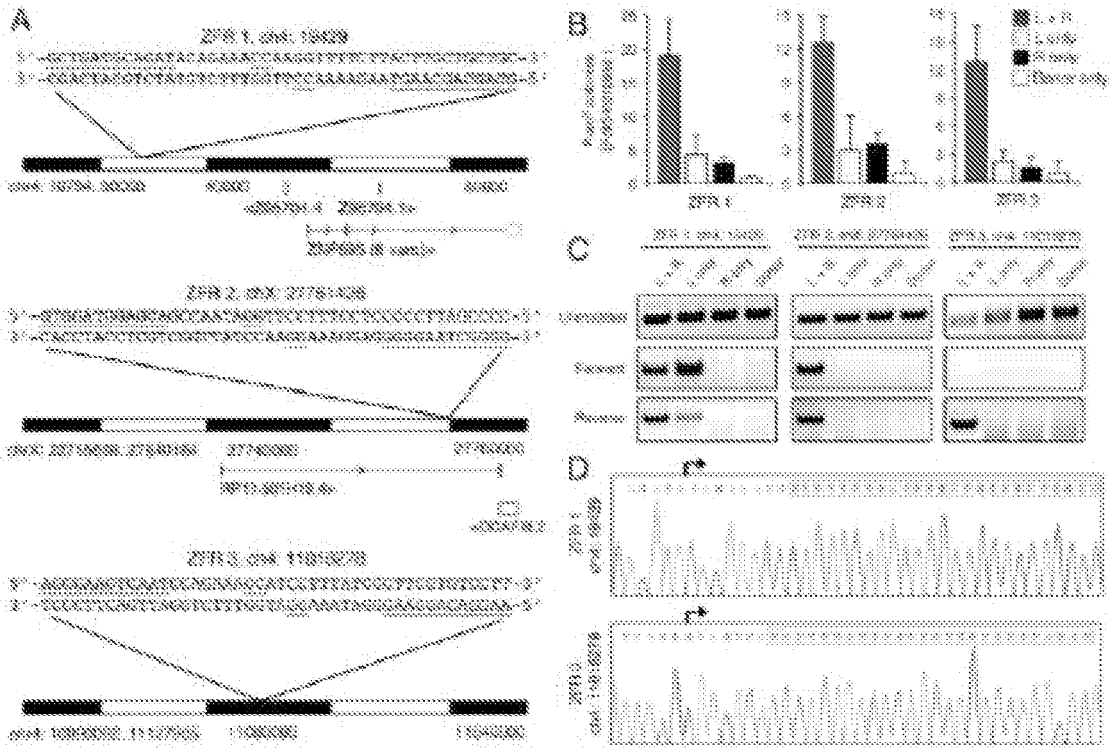


FIG. 34

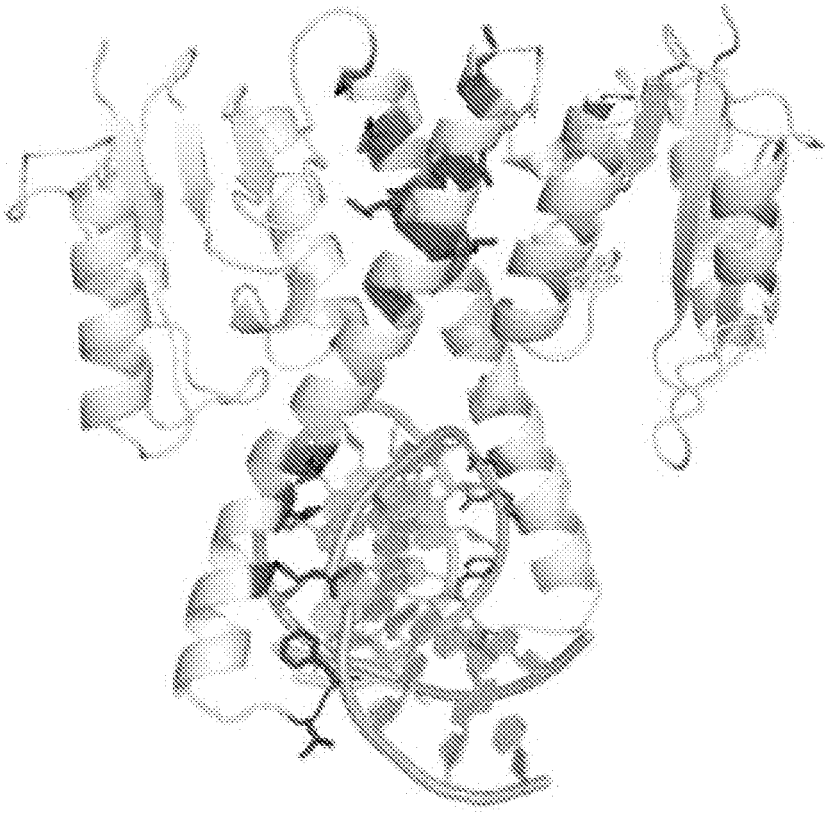


FIG. 35

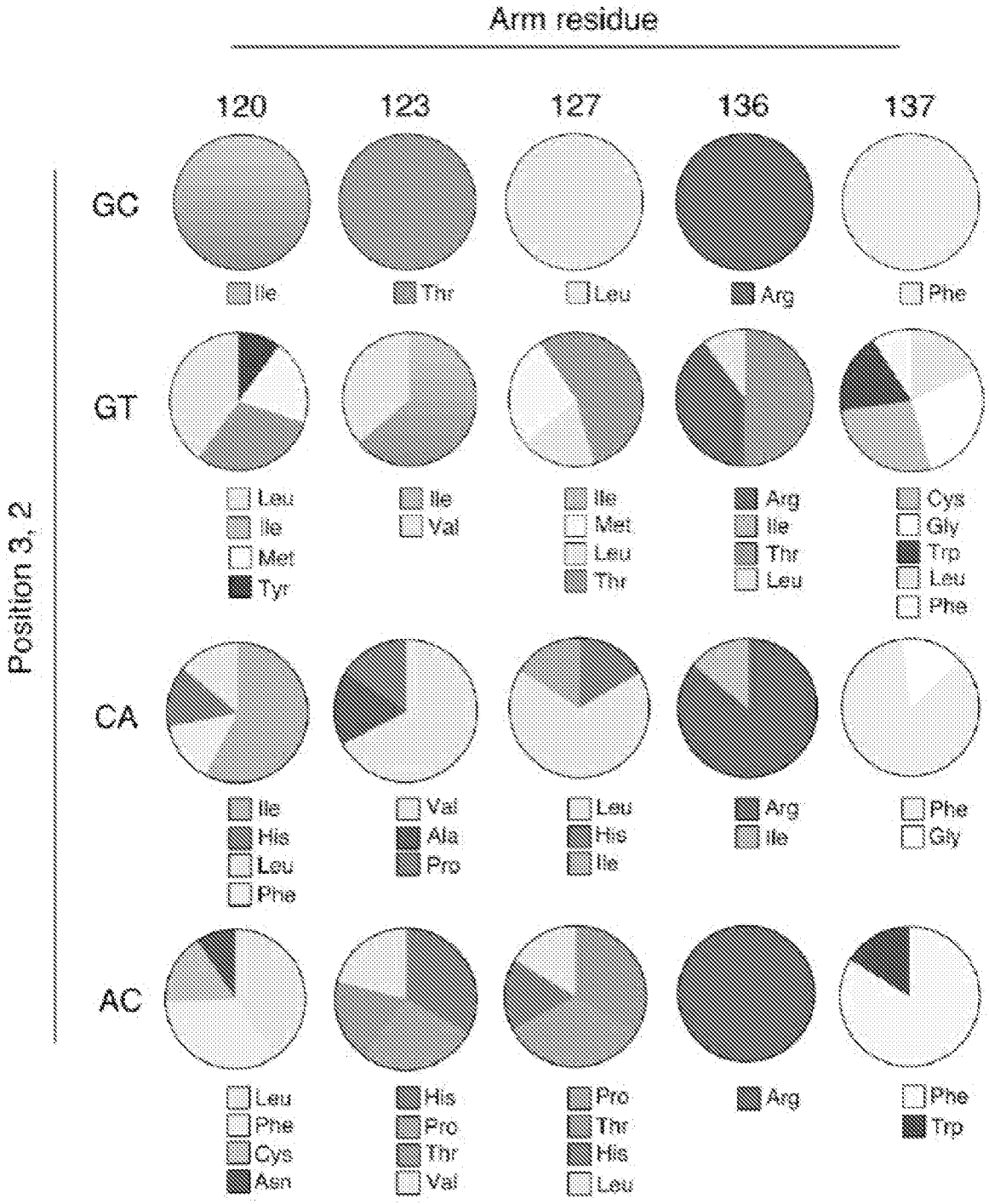

Arm residue
120
123
127
136
137
Position 3, 2
GC
Ile
Thr
Leu
Arg
Phe
GT
Leu
Ile
Met
Tyr
Ile
Val
Ile
Met
Leu
Thr
Arg
Ile
Thr
Leu
Cys
Gly
Trp
Leu
Phe
CA
Ile
His
Leu
Phe
Val
Ala
Pro
Leu
His
Ile
Arg
Ile
Phe
Gly
AC
Leu
Phe
Cys
Asn
His
Pro
Thr
Val
Pro
Thr
His
Leu
Arg
Phe
Trp

FIG. 36

3.2-kb

	Clone	Mutations	CC	GG	CG	GC	GT	TT
	Parental		36	45	35	0.01	0.003	0.06
★	GC-1	L127I, I136R, G137F	44	28	42	20	0.09	0.06

	Clone	Mutations	CC	TT	GA	CA	GT	CT
	Parental		36	0.002	0.008	0.004	0.003	0.089
★	GT-1	I12L, T123V, L127I, I136R, G137W	0.009	0.053	3.06	3.1	27	13.4
	GT-2	L127M, I136R, G137C	0.016	0.0023	0.082	0.237	38	6
	GT-3	L127I	0.0209	0.0029	0.41	0.77	39	13.5
	GT-4	D13A*, I12Y, T123I, L127I, I136T, G137L	0.0018	0.034	-	-	15	-
	GT-5	I12M, T123V	0.0068	0.013	-	-	8	-
	GT-6	I12L, T123I, L127I, G137F	0.0135	0.022	-	-	3	-

	Clone	Mutations	CC	TT	GA	CA	GT	CT
	Parental		36	0.002	0.01	0.004	0.003	0.09
★	CA-1	T123V, I136R, G137F	0.018	0.003	1.2	13	23	14
	CA-2	T123V	0.02	0.08	-	11	27	-
	CA-3	T123N, N132D	0.02	0.6	-	8.2	11.2	-
	CA-4	L92I*, L118S*, I120F, L127H, I136R, G137F	-	-	-	0.87	-	-
	CA-5	I12L, T123V, I136R, G137F	-	-	-	0.43	-	-
	CA-6	E117Q, L118C*, I120H, T123P, L127I	-	-	-	0.5	-	-

	Clone	Mutations	TT	AC	CA	GT	AG	GC
	Parental		0.002	0.003	0.004	0.003	0.008	0.01
★	AC-1	E117L, L118S*, I120L, T123P, L127H, I136R, G137F	0.09	1.7	3.06	6	1	0.008
	AC-2	E117L, I12N, I136R, G137F	0.0002	0.7	-	0.12	-	0.03
	AC-3	R116C*, E117L, L118S*, I120F, T123H, L127T, I136R, G137F	0.0007	0.1	0.0005	0.0014	0.2	-
	AC-4	E117V, I12F, T123P, L127P, I136R, G137F	0.0006	0.01	0.056	0.001	0.018	-

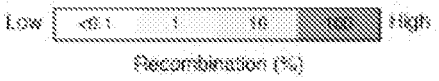


FIG. 37

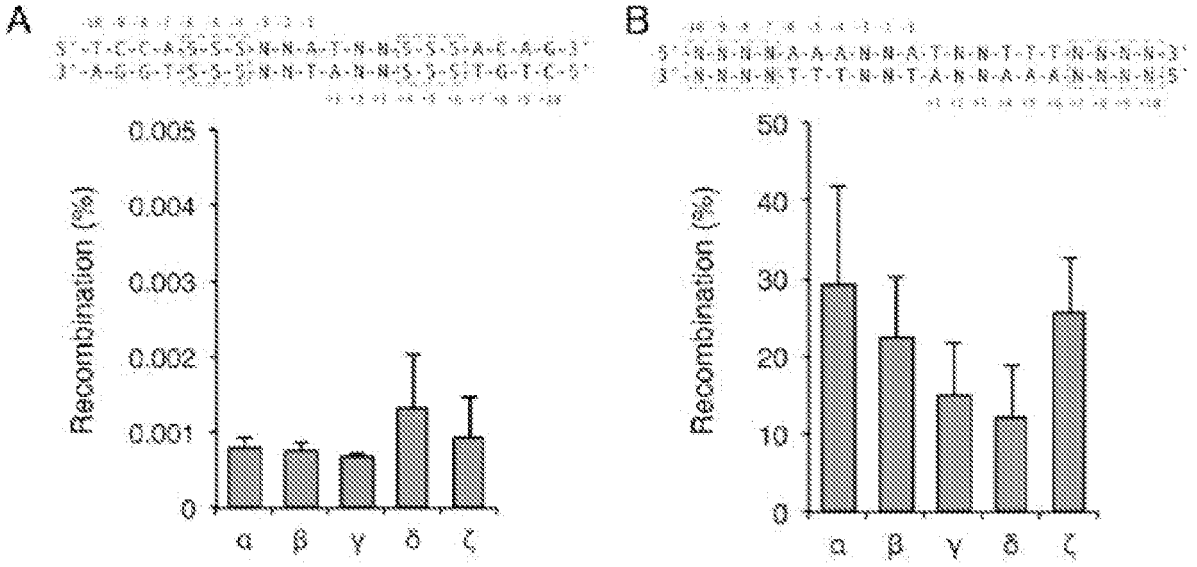

A
B
Recombination (%)
0.005
0.004
0.003
0.002
0.001
0
50
40
30
20
10
α β γ δ ζ

FIG. 38

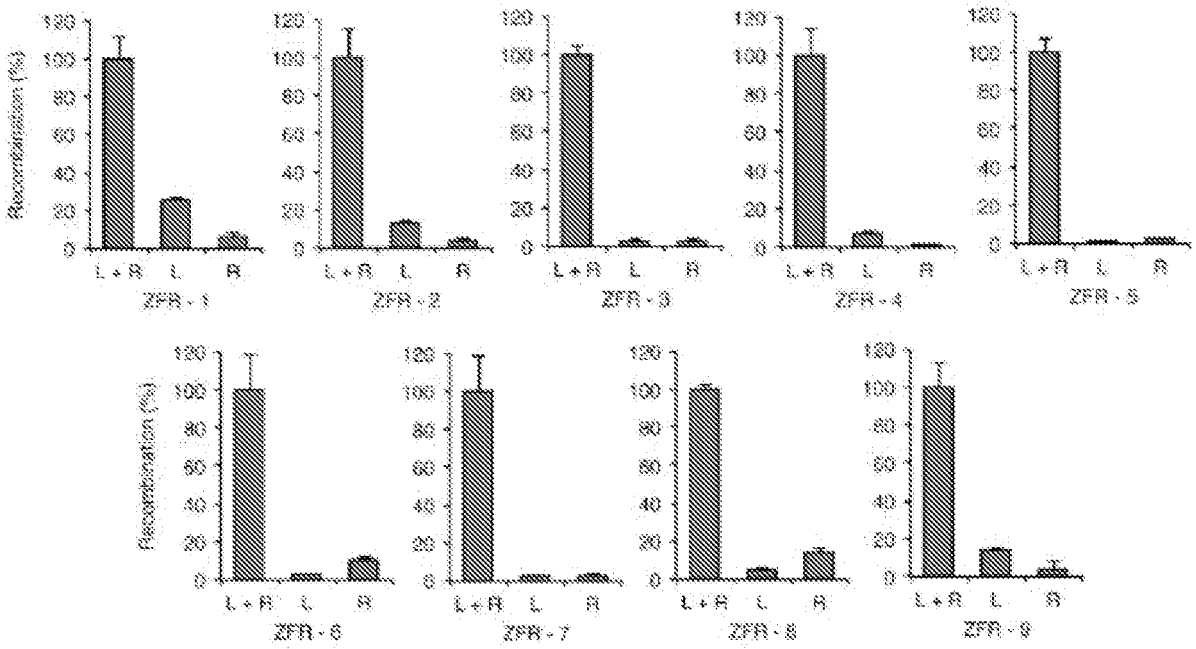

Recombination (%)
L + R
L
R
ZFR - 1
ZFR - 2
ZFR - 3
ZFR - 4
ZFR - 5
ZFR - 6
ZFR - 7
ZFR - 8
ZFR - 9

FIG. 39

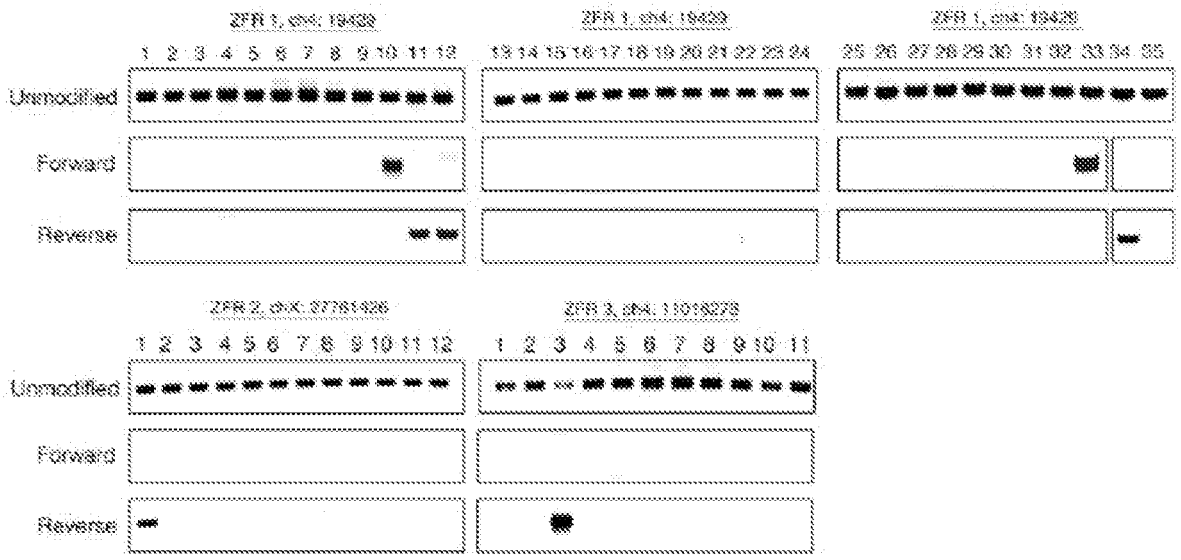

1 2 3 4 5 6 7 8 9 10 11 12
13 14 15 16 17 18 19 20 21 22 23 24
25 26 27 28 29 30 31 32 33 34 35
Unmodified
Forward
Reverse
1 2 3 4 5 6 7 8 9 10 11 12
1 2 3 4 5 6 7 8 9 10 11
Unmodified
Forward
Reverse

FIG. 40