(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
03 March 2022 (03.03.2022)

WIPO | PCT

(10) International Publication Number
WO 2022/046662 A1

(51) International Patent Classification:
C12N 15/63 (2006.01) *C12N 9/22* (2006.01)
C12N 15/10 (2006.01) *C12N 15/113* (2010.01)

(21) International Application Number:
PCT/US2021/047195

(22) International Filing Date:
23 August 2021 (23.08.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

63/069,703	24 August 2020 (24.08.2020)	US
63/186,698	10 May 2021 (10.05.2021)	US
63/232,593	12 August 2021 (12.08.2021)	US

(71) Applicant: **METAGENOMI IP TECHNOLOGIES, LLC** [US/US]; 1545 Park Avenue, Emeryville, California 94608 (US).

(72) Inventors: **THOMAS, Brian**; 1545 Park Avenue, Emeryville, California 94608 (US). **BROWN, Christopher**; 1545 Park Avenue, Emeryville, California 94608 (US). **GOLTSMAN, Daniela S.A.**; 1545 Park Avenue, Emeryville, California 94608 (US). **BUTTERFIELD, Cristina**; 1545 Park Avenue, Emeryville, California 94608 (US). **ALEXANDER, Lisa**; 1545 Park Avenue, Emeryville, California 94608 (US). **LIU, Jason**; 1545 Park Avenue, Emeryville, California 94608 (US).

(74) Agent: **TAVSHANJIAN, Brandon**; Wilson Sonsini Goodrich & Rosati, 650 Page Mill Road, Palo Alto, California 94304 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH,

(54) Title: SYSTEMS AND METHODS FOR TRANSPOSING CARGO NUCLEOTIDE SEQUENCES

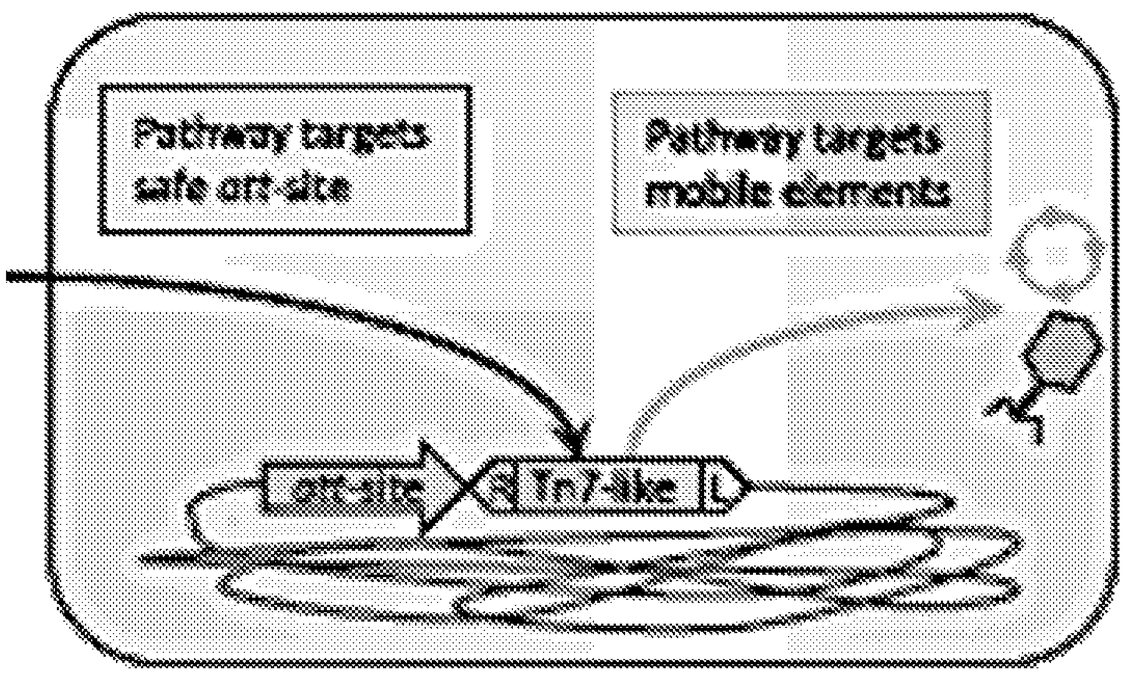


FIG. 3

(57) Abstract: The present disclosure provides systems and methods for transposing a cargo nucleotide sequence to a target nucleic acid site. These systems and methods may comprise a first double-stranded nucleic acid comprising the cargo nucleotide sequence, wherein the cargo nucleotide sequence is configured to interact with a recombinase or transposase complex, a cas effector complex comprising a cas effector and at least one engineered guide polynucleotide configured to hybridize to the target nucleic acid site, and the recombinase or transposase complex wherein said recombinase or transposase complex is configured to recruit the cargo nucleotide to the target nucleic acid site.

[Continued on next page]

WO 2022/046662 A1

GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

SYSTEMS AND METHODS FOR TRANSPOSING CARGO NUCLEOTIDE SEQUENCES

RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Application No. 63/069,703, filed August 24, 2020, entitled "SYSTEMS AND METHODS FOR TRANSPOSING CARGO NUCLEOTIDE SEQUENCES", U.S. Provisional Application No. 63/186,698, filed May 10, 2021, entitled "SYSTEMS AND METHODS FOR TRANSPOSING CARGO NUCLEOTIDE SEQUENCES", and U.S. Provisional Application No. 63/232,593, filed August 12, 2021, entitled "SYSTEMS AND METHODS FOR TRANSPOSING CARGO NUCLEOTIDE SEQUENCES", each of which is incorporated by reference in its entirety herein.

BACKGROUND

[0002] Cas enzymes along with their associated Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) guide ribonucleic acids (RNAs) appear to be a pervasive (~45% of bacteria, ~84% of archaea) component of prokaryotic immune systems, serving to protect such microorganisms against non-self nucleic acids, such as infectious viruses and plasmids by CRISPR-RNA guided nucleic acid cleavage. While the deoxyribonucleic acid (DNA) elements encoding CRISPR RNA elements may be relatively conserved in structure and length, their CRISPR-associated (Cas) proteins are highly diverse, containing a wide variety of nucleic acid-interacting domains. While CRISPR DNA elements have been observed as early as 1987, the programmable endonuclease cleavage ability of CRISPR/Cas complexes has only been recognized relatively recently, leading to the use of recombinant CRISPR/Cas systems in diverse DNA manipulation and gene editing applications.

SEQUENCE LISTING

[0003] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on August 20, 2021, is named 55921-714_601_SL.txt and is 488,452 bytes in size.

SUMMARY

[0004] In some aspects, the present disclosure provides for a system for transposing a cargo nucleotide sequence to a target nucleic acid site comprising: a first double-stranded nucleic acid comprising said cargo nucleotide sequence, wherein said cargo nucleotide sequence is configured to interact with a recombinase or transposase complex; a Cas effector complex

comprising a class II, type II Cas effector and at least one engineered guide polynucleotide configured to hybridize to said target nucleic acid site; and said recombinase or transposase complex, wherein said recombinase or transposase complex is configured to recruit said cargo nucleotide sequence to said target nucleic acid site. In some embodiments, said recombinase or transposase complex binds non-covalently to said Cas effector complex. In some embodiments, said recombinase or transposase complex is covalently linked to said Cas effector complex. In some embodiments, said recombinase or transposase complex is fused to said Cas effector complex in a single polypeptide. In some embodiments, said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence. In some embodiments, the system further comprises a second double-stranded nucleic acid comprising said target nucleic acid site. In some embodiments, the system further comprises a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site. In some embodiments, said PAM sequence is located 3' of said target nucleic acid site. In some embodiments, said recombinase or transposase complex is a Tn7 type transposase complex. In some embodiments, said engineered guide polynucleotide is configured to bind said class II, type II Cas effector. In some embodiments, said class II, type II Cas effector comprises a polypeptide comprising a sequence having at least 80% identity to SEQ ID NO: 1 or a variant thereof. In some embodiments, said recombinase or transposase complex comprises at least one, at least two, at least three, or four polypeptide(s) comprising a sequence having at least 80% identity to any one of SEQ ID NOs: 2-5 or a variant thereof. In some embodiments, said engineered guide polynucleotide comprises a sequence comprising at least 60-80 consecutive nucleotides having at least 80% identity to SEQ ID NO:12 or a variant thereof. In some embodiments, said engineered guide polynucleotide comprises a sequence having at least 80% identity to SEQ ID NO: 11 or a variant thereof In some embodiments, said left-hand recombinase sequence comprises a sequence having at least 80% identity to any one of SEQ ID NOs: 17-18 or a variant thereof. In some embodiments, said right-had recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 19 or a variant thereof. In some embodiments, said class II, type II Cas effector and said recombinase or transposase complex are encoded by polynucleotide sequences comprising fewer than about 10 kilobases

[0005] In some aspects, the present disclosure provides for a method for transposing a cargo nucleotide sequence to a target nucleic acid site comprising a target nucleotide sequence comprising expressing the system of any of the aspects or embodiments described herein within a cell or introducing the system of any of the aspects or embodiments described herein to a cell.

[0006] In some aspects, the present disclosure provides for a system for transposing a cargo nucleotide sequence to a target nucleic acid site comprising: a first double-stranded nucleic acid comprising a cargo nucleotide sequence configured to interact with a Tn7 type transposase complex; a Cas effector complex comprising a class II, type V Cas effector and an engineered guide polynucleotide configured to hybridize to said target nucleotide sequence; and a Tn7 type transposase complex configured to bind said Cas effector complex, wherein said Tn7 type transposase complex comprises a TnsA subunit. In some embodiments, said transposase complex binds non-covalently to said Cas effector complex. In some embodiments, said transposase complex is covalently linked to said Cas effector complex. In some embodiments, said transposase complex is fused to said Cas effector complex in a single polypeptide. In some embodiments, said class II, type V Cas effector is not a Cas12k effector. In some embodiments, said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence. In some embodiments, the system further comprises a second double-stranded nucleic acid comprising said target nucleic acid site. In some embodiments, the system further comprises a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site. In some embodiments, said PAM sequence is located 5' of said target nucleic acid site. In some embodiments, said engineered guide polynucleotide is configured to bind said class II, type V Cas effector. In some embodiments, said TnsA subunit comprises a polypeptide having a sequence having at least 80% identity to SEQ ID NO: 7 or a variant thereof. In some embodiments, said Tn7 type transposase complex comprises at least one, at least two, or three polypeptide(s) comprising a sequence having at least 80% identity to any one of SEQ ID NOs: 8-10, or a variant thereof. In some embodiments, said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 13-16 , or a variant thereof. In some embodiments, said left-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 20 , or a variant thereof. In some embodiments, said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 21, or a variant thereof. In some embodiments, said class II, type V Cas effector is not a Cas12k effector. In some embodiments, said class II, type V Cas effector and said Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 10 kilobases.

[0007] In some aspects, the present disclosure provides for a method for transposing a cargo nucleotide sequence to a target nucleic acid site comprising a target nucleotide sequence comprising expressing the system of any one of any of the aspects or embodiments described

herein within a cell or introducing the system of any one of the aspects or embodiments described herein to a cell.

[0008] In some aspects, the present disclosure provides for a method for transposing a cargo nucleotide sequence to a target nucleic acid site, comprising contacting a first double-stranded nucleic acid comprising a cargo nucleotide sequence with: a Cas effector complex comprising a class II, type II Cas effector and at least one engineered guide polynucleotide configured to hybridize to said target nucleic acid site; a recombinase or transposase complex configured to recruit said cargo nucleotide to said target nucleic acid site; and a second double-stranded nucleic acid comprising said target nucleic acid site. In some embodiments, said recombinase or transposase complex binds non-covalently to said Cas effector complex. In some embodiments, said recombinase or transposase complex is covalently linked to said Cas effector complex. In some embodiments, said recombinase or transposase complex is fused to said Cas effector complex in a single polypeptide. In some embodiments, said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence. In some embodiments, the target nucleic acid further comprises a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site. In some embodiments, said PAM sequence is located 3' of said target nucleic acid site. In some embodiments, said recombinase or transposase complex is a Tn7 type transposase complex. In some embodiments, said engineered guide polynucleotide is configured to bind said class II, type II Cas effector. In some embodiments, said class II, type II Cas effector comprises a polypeptide comprising a sequence having at least 80% identity to SEQ ID NO: 1 or a variant thereof. In some embodiments, said recombinase or transposase complex comprises at least one, at least two, at least three, or four polypeptide(s) comprising a sequence having at least 80% identity to any one of SEQ ID NOs: 2-5 or a variant thereof. In some embodiments, said engineered guide polynucleotide comprises a sequence comprising at least 60-80 consecutive nucleotides having at least 80% identity to SEQ ID NO: 12 or a variant thereof. In some embodiments, said engineered guide polynucleotide comprises a sequence having at least 80% identity to SEQ ID NO: 11 or a variant thereof. In some embodiments, said left-hand recombinase sequence comprises a sequence having at least 80% identity to any one of SEQ ID NOs: 17-18 or a variant thereof. In some embodiments, said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 19 or a variant thereof. In some embodiments, said class II, type II Cas effector and said Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 10 kilobases.

[0009] In some aspects, the present disclosure provides for a method for transposing a cargo nucleotide sequence to a target nucleic acid site, comprising contacting a first double-stranded nucleic acid comprising said cargo nucleotide sequence with: a Cas effector complex comprising a class II, type V Cas effector and at least one engineered guide polynucleotide configured to hybridize to said target nucleotide sequence; a Tn7 type transposase complex configured to bind said Cas effector complex, wherein said Tn7 type transposase complex comprises a TnsA subunit; and a second double-stranded nucleic acid comprising said target nucleic acid site. In some embodiments, said transposase complex binds non-covalently to said Cas effector complex. In some embodiments, said transposase complex is covalently linked to said Cas effector complex. In some embodiments, said transposase complex is fused to said Cas effector complex in a single polypeptide. In some embodiments, said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence. In some embodiments, said target nucleic acid site further comprises a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site. In some embodiments, said PAM sequence is located 3' of said target nucleic acid site. In some embodiments, said engineered guide polynucleotide is configured to bind said class II, type V Cas effector. In some embodiments, said TnsA subunit comprises a polypeptide having a sequence having at least 80% identity to SEQ ID NO: 7or a variant thereof. In some embodiments, said Tn7 type transposase complex comprises at least one, at least two, or three polypeptide(s) comprising a sequence having at least 80% identity to any one of SEQ ID NOs: 8-10, or a variant thereof. In some embodiments, said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 13-16 or a variant thereof. In some embodiments, said left-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 20, or a variant thereof. In some embodiments, said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 21, or a variant thereof. In some embodiments, said class II, type V Cas effector is not a Cas12k effector. In some embodiments, said class II, type V Cas effector and said Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 10 kilobases.

[0010] In some aspects, the present disclosure provides for a system for transposing a cargo nucleotide sequence to a target nucleic acid site comprising: a first double-stranded nucleic acid comprising a cargo nucleotide sequence configured to interact with a Tn7 type transposase complex; a Cas effector complex comprising a class I, type I-F Cas effector and an engineered guide polynucleotide configured to hybridize to said target nucleotide sequence; and a Tn7 type

transposase complex configured to bind said Cas effector complex, wherein said Tn7 type transposase complex comprises a TnsA subunit. In some embodiments, said transposase complex binds non-covalently to said Cas effector complex. In some embodiments, said transposase complex is covalently linked to said Cas effector complex. In some embodiments, said transposase complex is fused to said Cas effector complex in a single polypeptide. In some embodiments, said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence. In some embodiments, the system further comprises a second double-stranded nucleic acid comprising said target nucleic acid site. In some embodiments, the system further comprises a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site. In some embodiments, said PAM sequence is located 3' of said target nucleic acid site. In some embodiments, said PAM sequence is located 5' of said target nucleic acid site. In some embodiments, said engineered guide polynucleotide is configured to bind said class I, type I-F Cas effector. In some embodiments, said class I, type I-F Cas effector comprises a polypeptide comprising a sequence having at least 80% identity to any one of SEQ ID NO: 41-43, or 48-50, or a variant thereof. In some embodiments, said Tn7 type transposase complex comprises at least one, at least two, or three polypeptide(s) comprising a sequence having at least 80% identity to any one of SEQ ID NOs: 44-46, or 51-53, or a variant thereof.

[0011] In some aspects, the present disclosure provides for a method for transposing a cargo nucleotide sequence to a target nucleic acid site comprising a target nucleotide sequence comprising expressing the system of any one of the aspects or embodiments described herein within a cell or introducing the system of any one of the aspects or embodiments described herein to a cell.

[0012] In some aspects, the present disclosure provides for a system for transposing a cargo nucleotide sequence to a target nucleic acid site comprising: a first double-stranded nucleic acid comprising a cargo nucleotide sequence configured to interact with a Tn7 type transposase complex; a Cas effector complex comprising a class II, type V Cas effector and an engineered guide polynucleotide configured to hybridize to said target nucleotide sequence; and a Tn7 type transposase complex configured to bind said Cas effector complex, wherein said Tn7 type transposase complex comprises TnsB, TnsC, and TniQ components, wherein: (a) said class II, type V Cas effector comprises a polypeptide having a sequence having at least 80% sequence identity to any one of SEQ ID NOs:22, 26, 30, 34, 55-89, 104, or 147, or a variant thereof; or (b) said Tn7 type transposase complex comprises a TnsB, TnsC, or TniQ component having a sequence having at least 80% sequence identity to any one of SEQ ID NOs:23-25, 27-29, 31-33,

35-37, 101-103, 105-107, or 148-150, or a variant thereof. In some embodiments, said transposase complex binds non-covalently to said Cas effector complex. In some embodiments, said transposase complex is covalently linked to said Cas effector complex. In some embodiments, said transposase complex is fused to said Cas effector complex in a single polypeptide. In some embodiments, said class II, type V Cas effector comprises a polypeptide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs:22, 26, 30, 34, 55-89, 104, or 147, or a variant thereof. In some embodiments, said Tn7 type transposase complex comprises a TnsB, TnsC, or TniQ component comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs:23-25, 27-29, 31-33, 35-37, 101-103, 105-107, or 148-150, or a variant thereof. In some embodiments, said class II, type V Cas effector is a Cas12k effector. In some embodiments, said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence. In some embodiments, the system further comprises a second double-stranded nucleic acid comprising said target nucleic acid site. In some embodiments, the system further comprises a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site. In some embodiments, said PAM sequence is located 5' of said target nucleic acid site. In some embodiments, said PAM sequence comprises 5'-nGTn-3' or 5'-nGTt-3'. In some embodiments, said engineered guide polynucleotide is configured to bind said class II, type V Cas effector. In some embodiments, said TnsB, TnsC, and TniQ components comprise polypeptides having a sequence having at least 80% identity to any one of SEQ ID NOs:23-25, 27-29, 31-33, 35-37, 101-103, 105-107, or 148-150, respectively. In some embodiments, said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 90, 91, 92, 93, 117, 151, 156-181, or 209-234. In some embodiments, said engineered guide polynucleotide comprises a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 111-114 or 201- 206, 255, 262, 256, 209, 257, 263, 258, 210, or a variant thereof. In some embodiments, said left-hand recombinase sequence comprises a sequence having at least 80% identity to any one of SEQ ID NOs: 125, 127, 123, 129, 131, 133, 153, or 134, or a variant thereof. In some embodiments, said right-hand recombinase sequence comprises a sequence having at least 80% identity to any one of SEQ ID NOs: 126, 155, 128, 124, 130, 132, or 154 , or a variant thereof. In some embodiments, said class II, type V Cas effector and said Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 10 kilobases. In some embodiments: (a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:22 or a variant thereof; (b) said left-hand

recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:125 or a variant thereof; (c) said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 126 or 155, or a variant thereof; (d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about 46-60 nucleotides of SEQ ID NO: 90; or (ii) comprises a sequence having at least 80% sequence identity to non-degenerate nucleotides of any one of SEQ ID NOs: 94, 112, or 202; or (e) said TnsB, TnsC, and TniQ components comprise sequences having at least 80% sequence identity to SEQ ID NOs: 23-25 or variants thereof. In some embodiments: (a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:26 or a variant thereof; (b) said left-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:127 or a variant thereof; (c) said right-hand recombinase sequence comprises a sequence having at least 880% sequence identity to SEQ ID NO:128 or a variant thereof; (d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about 46-60 nucleotides of any one of SEQ ID NOs: 91, 156, or 209; or (ii) comprises a sequence having at least 80% sequence identity to non-degenerate nucleotides of any one of SEQ ID NOs: 95, 113, or 203, or (e) said TnsB, TnsC, and TniQ components comprise sequences having at least 80% sequence identity to SEQ ID NOs: 27-29 or variants thereof. In some embodiments: (a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:60 or a variant thereof; (b) said left-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:131 or a variant thereof; (c) said right-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:132 or a variant thereof; (d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about 46-60 nucleotides of any one of SEQ ID NOs: 117, 161, or 214; or (ii) comprises a sequence having at least 80% sequence identity to non-degenerate nucleotides of SEQ ID NO: 119; or (e) said TnsB, TnsC, and TniQ components comprise sequences having at least 80% sequence identity to SEQ ID NOs: 101-103 or variants thereof. In some embodiments: (a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:147 or a variant thereof; (b) said left-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:153 or a variant thereof; (c) said right-hand recombinase sequence comprises a sequence having at least 880% sequence identity to SEQ ID NO:154 or a variant thereof; (d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about 46-60 nucleotides of any one of SEQ ID NOs: 151, 181, or 234; or (ii) comprises a sequence

having at least 80% sequence identity to non-degenerate nucleotides of SEQ ID NO: 152 or 254; or (e) said TnsB, TnsC, and TniQ components comprise sequences having at least 80% sequence identity to SEQ ID NOs:148-150 or variants thereof. In some embodiments: (a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:34 or a variant thereof; (b) said left-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:129 or a variant thereof; (c) said right-hand recombinase sequence comprises a sequence having at least 880% sequence identity to SEQ ID NO:130 or a variant thereof; (d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about 46-60 nucleotides of any one of SEQ ID NOs: 93, 157, or 210; or (ii) comprises a sequence having at least 80% sequence identity to non-degenerate nucleotides of any one of SEQ ID NOs: 97, 114, or 204, or (e) said TnsB, TnsC, and TniQ components comprise sequences having at least 80% sequence identity to SEQ ID NOs:148-150 or variants thereof. In some embodiments,: (a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:30 or a variant thereof; (b) said left-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:123 or a variant thereof; (c) said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 124, or a variant thereof; (d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about 46-80 nucleotides of SEQ ID NO:92; or (ii) comprises a sequence having at least 80% identity to the non-degenerate nucleotides of SEQ ID NO:111 or 201; (e) said TnsB, TnsC, and TniQ components comprise polypeptides having a sequence having at least 80% identity to SEQ ID NOs: 31, 32, and 33, or variants thereof; or (f) said PAM sequence comprises 5'-nGTn-3' or 5'-nGTt-3'.

[0013] In some aspects, the present disclosure provides for a system for transposing a cargo nucleotide sequence to a target nucleic acid site comprising: a first double-stranded nucleic acid comprising a cargo nucleotide sequence configured to interact with a Tn7 type transposase complex; a Cas effector complex comprising a class II, type V Cas effector and an engineered guide polynucleotide configured to hybridize to said target nucleotide sequence; and a Tn7 type transposase complex configured to bind said Cas effector complex, wherein said Tn7 type transposase complex comprises TnsB and TnsC components but does not comprise a TnsA and/or TniQ component. In some embodiments, said transposase complex binds non-covalently to said Cas effector complex. In some embodiments, said transposase complex is covalently linked to said Cas effector complex. In some embodiments, said transposase complex is fused to said Cas effector complex in a single polypeptide. In some embodiments, said Tn7 type

transposase complex comprises a polypeptide having a sequencing having at least 80% sequence identity to any one of SEQ ID NOs: 39-40 or 109-110. In some embodiments, said TnsB component comprises a polypeptide comprising a sequence having at least 80% sequence identity to SEQ ID NOs: 40 or 109. In some embodiments, said TnsC component comprises a polypeptide comprising a sequence having at least 80% sequence identity to SEQ ID NOs: 39 or 110. In some embodiments, said class II, type V Cas effector is a Cas12k effector. In some embodiments, said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO: 38 or SEQ ID NO:108. In some embodiments, said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence. In some embodiments, the system further comprises a second double-stranded nucleic acid comprising said target nucleic acid site. In some embodiments, said double-stranded nucleic acid comprising said target nucleic acid site or said system is inside a cell. In some embodiments, the system further comprises a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site. In some embodiments, said PAM sequence is located 5' of said target nucleic acid site. In some embodiments, said engineered guide polynucleotide is configured to bind said class II, type V Cas effector. In some embodiments, said TnsB and TnsC components comprise polypeptides having a sequence having at least 80% identity to SEQ ID NOs: 40 and 39 or 109 and 110, respectively. In some embodiments, said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 118, 182, 183, 235, or 236, or a variant thereof. In some embodiments, said engineered guide polynucleotide comprises a sequence having at least 80% identity to non-degenerate nucleotides any one of SEQ ID NOs: 115, 116, 205, 206, 261, 235, 260, or 236, or a variant thereof. In some embodiments, said left-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO:134. In some embodiments, said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NOs: 135, or a variant thereof. In some embodiments, said class II, type V Cas effector and said Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 10 kilobases. In some embodiments (a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:38 or a variant thereof; (b) said left-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:134 or a variant thereof; (c) said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 135, or a variant thereof; (d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about

46-80 nucleotides of SEQ ID NO:182 or 235; or (ii) comprises a sequence having at least 80% identity to the non-degenerate nucleotides of SEQ ID NO:98, 115, 116, 205, or 206; or (e) said TnsB and TnsC components comprise polypeptides having a sequence having at least 80% identity to SEQ ID NOs: 40 and 39, or variants thereof.

[0014] In some aspects, the present disclosure provides for an engineered nuclease system comprising: an endonuclease comprising a RuvC domain and an HNH domain, wherein said endonuclease is derived from an uncultivated microorganism, wherein said endonuclease is a Class II, type II endonuclease comprising a sequence having at least 80% identity to SEQ ID NO: 1 or a variant thereof; and an engineered guide polynucleotide, wherein said engineered guide polynucleotide is configured to form a complex with said endonuclease and said engineered guide polynucleotide comprises a spacer sequence configured to hybridize to a target nucleic acid sequence. In some embodiments, said engineered guide polynucleotide comprises at least 60-80 consecutive nucleotides having at least 80% identity to SEQ ID NO:12 or a variant thereof. In some embodiments, said engineered guide polynucleotide comprises a sequence having at least 80% identity to SEQ ID NO: 11 or a variant thereof.

[0015] In some aspects, the present disclosure provides for an engineered nuclease system comprising: an endonuclease comprising a RuvC domain, wherein said endonuclease is derived from an uncultivated microorganism, and wherein said endonuclease is a Class II, type V endonuclease having at least 80% identity to SEQ ID NO: 5; and an engineered guide polynucleotide, wherein said engineered guide polynucleotide is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence. In some embodiments, said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 13-16 , or a variant thereof.

[0016] In some aspects, the present disclosure provides for an engineered nuclease system comprising: an endonuclease comprising a RuvC domain, wherein said endonuclease is derived from an uncultivated microorganism, and wherein said endonuclease is a Class II, type V-K endonuclease having at least 80% identity to any one of SEQ ID NOs:22, 26, 30, 34, 55-89, 104, or 147, or a variant thereof; and an engineered guide polynucleotide, wherein said engineered guide polynucleotide is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence. In some embodiments, said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 90, 91, 92, 93, 117, 151, 156-181, or 209-234, or a variant thereof. In

some embodiments, said engineered guide polynucleotide comprises a sequence having at least 80% sequence identity to non-degenerate nucleotides of any one of SEQ ID NOs: 111-114 or 201- 206, 255, 262, 256, 209, 257, 263, 258, 210, or a variant thereof.

[0017] In some aspects, the present disclosure provides for an engineered nuclease system comprising: an endonuclease comprising a RuvC domain, wherein said endonuclease is derived from an uncultivated microorganism, and wherein said endonuclease is a Class II, type V-K endonuclease having at least 80% identity to any one of SEQ ID NO: 38 or SEQ ID NO:108, or a variant thereof; and an engineered guide polynucleotide, wherein said engineered guide polynucleotide is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence. In some embodiments, said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 118, 182, 183, 235, or 236, or a variant thereof. In some embodiments, said engineered guide polynucleotide comprises a sequence having at least 80% identity to non-degenerate nucleotides of any one of SEQ ID NOs: 111-114 or 201-206, 255, 262, 256, 209, 257, 263, 258, 210, 115, 116, 205, 206, 261, 235, 260, or 236 , or a variant thereof.

[0018] In some aspects, the present disclosure provides for an engineered nuclease system comprising: a Class I, type I-F Cas endonuclease comprising at least one Cas6, Cas7, or Cas8 polypeptide comprising a sequence having at least 80% identity to any one of SEQ ID NO: 41-43, or 48-50, or a variant thereof; and an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence. In some embodiments, said engineered guide polynucleotide comprises a sequence having at least 80% identity to non-degenerate nucleotides of any one of SEQ ID NOs: 121, 122, 207, or 208.

[0019] Additional aspects and advantages of the present disclosure will become readily apparent to those skilled in this art from the following detailed description, wherein only illustrative embodiments of the present disclosure are shown and described. As will be realized, the present disclosure is capable of other and different embodiments, and its several details are capable of modifications in various obvious respects, all without departing from the disclosure. Accordingly, the drawings and description are to be regarded as illustrative in nature, and not as restrictive.

INCORPORATION BY REFERENCE

[0020] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or

patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0021] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings (also "Figure" and "FIG." herein), of which:

[0022] FIG. 1 depicts typical organizations of CRISPR/Cas loci of different classes and types.

[0023] FIG. 2 depicts the architecture of a natural Class II Type II crRNA/tracrRNA pair, compared to a hybrid sgRNA wherein the crRNA and tracrRNA are joined.

[0024] FIG. 3 depicts the two pathways found in Tn7 and Tn7-like elements.

[0025] FIG. 4 depicts the genomic context of a Type II Tn7 reduced CAST of the family MG36. **FIG. 4A** shows the MG36-5 CAST system consists of a CRISPR array (CRISPR repeats), a Type II nuclease with RuvC and HNH endonuclease domains, and four predicted transposase protein open reading frames. The catalytic transposase TnsB is encoded as two subunits. **FIG. 4B** shows two transposon ends are predicted for the MG36-1 CAST system (TIR-1 and TIR-2). **FIG. 4C** shows alignment of the predicted Type II Tn7 reduced CAST transposon left end (LE) and right end (RE) sequences, with annotated repeats as arrows. The left and right ends were labeled by their orientation.

[0026] FIG. 5 depicts the genomic context of a Type V Tn7 CAST of the family MG39. **FIG. 5A** shows the MG39-1 CAST system consists of a Type V nuclease, four predicted transposon proteins (TnsABC and TniQ), and a CRISPR array. The transposon ends were predicted for the MG39-1 CAST system (TIR-1). **FIG. 5B** shows the alignment of the predicted Type V Tn7 CAST transposon left end (LE) and right end (RE) sequences, with annotated inverted repeats represented as arrows.

[0027] FIG. 6 and **FIG. 7** depict predicted structures (predicted, for example in Example 3) of corresponding sgRNAs of CAST systems described herein.

[0028] FIG. 8 depicts the genomic context of MG108-1, a system described herein. This candidate is a Cas12K CAST which naturally lacks TniQ. Genes in the genomic fragment are represented by arrows.

[0029] FIG. 9 depicts the phylogenetic gene tree of Cas12k effector sequences. The tree was inferred from a multiple sequence alignment of 64 Cas12k sequences recovered here (orange and black branches) and 229 reference Cas12k sequences from public databases (grey branches). Orange branches indicate Cas12k effectors with confirmed association with CAST transposon

components.

[0030] FIG. 10 depicts MG110 Cascade CAST. **A)** Genomic context of the MG110-1 Cascade CAST. Full Tn7 suite (TnsA, TnsB, TnsC/TniB, TniQ) and defective Cascade suite (Cas6, Cas7, fused Cas5-Cas8) are represented by orange arrows. TIR flanking the CAST transposon are represented by connected arrows. **B)** Repeat secondary structure indicates a stem-loop structure of the crRNA. **C)** Sequence alignment of CRISPR repeats from A. wodanis, V. cholerae, and the MG110 family CASTs indicates conserved motifs indicative of the crRNA stem-loop secondary structure.

[0031] FIG. 11A depicts the MG64-3 CRISPR locus. The tracrRNA is encoded upstream from the CRISPR array, while the transposon end is encoded downstream (inner black box). A sequence corresponding to a partial 3' CRISPR repeat and a partial spacer are encoded within the transposon (outer box). The self-matching spacer is encoded outside of the transposon end. **FIG. 11B** depicts tracrRNA sequence alignment for various CASTs provided herein. Alignment of tracrRNA sequences shows regions of conservation. In particular, the sequence "TGCTTTC" at sequence position 92-98 (top box) is suggested to be important for sgRNA tertiary structure and for a non-continuous repeat-anti-repeat pairing with the crRNA. We also suggest that the hairpin "CYCC(n6)GGRG" at positions 265-278 (bottom box) is important for function, possibly positioning the downstream sequence for crRNA pairing. **FIG. 11C** shows presence of other important repeat-anti-repeat (RAR) motifs in e.g. MG64-2, MG64-4, MG64-5, MG64-6, MG64-7, and MG108-1 families.

[0032] FIG. 12A depicts the predicted structure of MG64-2 sgRNA. **FIG. 12B** depicts the predicted structure of MG64-4 sgRNA. **FIG. 12C** depicts the predicted structure of MG64-6 sgRNA. **FIG. 12D** depicts the predicted structure of MG64-7 sgRNA. **FIG. 12E** depicts the predicted structure of MG108-1 sgRNA.

[0033] FIG. 13 depicts PCR, PAM, and Sanger sequencing data which demonstrate that MG64-6 is active in vitro. Using the protocol described for In vitro targeted integrase activity, the effector protein and its TnsB, TnsC, and TniQ proteins were expressed in an in vitro transcription/translation system. After translation, the target DNA, cargo DNA, and sgRNA were added in reaction buffer. Integration was assayed by PCR across the target/donor junctions. **FIG. 13A** depicts a gel image of PCRs of transposition showing apo (no sgRNA) and 64-6 with sgRNA 64-6 sgRNA. The PCR 3 detects the RE junction, PAM distal. PCR 4 is LE junction, PAM distal. PCR 5 is RE junction, PAM proximal. PCR 6 is LE junction, PAM proximal. The PCRs are paired across the different possible orientations (PCR 3 and 6 vs PCR 4 and 5). The LE-PAM proximal and RE-PAM distal orientation is preferred. **FIG. 13B** depicts PAMs from

the in vitro transposition assay, sequencing PCRs 5 and 6. **FIG. 13C** depicts Sanger data which shows the junction of transposition where the excision occurs in the donor DNA. The first panel shows PCRs 3 and 5 (the RE). The second panel shows PCR 4 and 6 (the LE). The Sanger sequencing reaction is of the donor-target product, so the point where the sequencing stops matching the donor DNA is when junction occurs (dark bars underneath sequencing peaks)

[0034] FIG. 14 depicts next-generation sequencing (NGS) results of the in vitro transposition products which reveal the insertion site preferences. The NGS reads were processed in CRISPResso2 compared to a reference sequence with transposition at position 60. Indels from this correspond to transpositions earlier or later than this arbitrary reference sequence.

[0035] FIG. 15 depicts electrophoretic mobility shift assay (EMSA) results of the 64-2 TnsB and its RE DNA sequence. The EMSA results confirm binding and TnsB recognition. The TnsB protein was expressed in an in vitro transcription/translation system, incubated with FAM-labeled DNA containing the RE sequence, and then separated on a native 5% TBE gel. Binding is observed as a shift upwards in the labeled band. Multiple TnsB binding sites leads to multiple shifts in the EMSA. Lane 1: FAM-labeled DNA only. Lane 2: FAM DNA plus the in vitro transcription/translation system (no TnsB protein). Lane 3: FAM DNA plus TnsB. Upshift of the labeled band in Lane 3 indicates binding of the RE sequence by TnsB, indicating it contains an active RE transposition sequence.

BRIEF DESCRIPTION OF THE SEQUENCE LISTING

[0036] The Sequence Listing filed herewith provides exemplary polynucleotide and polypeptide sequences for use in methods, compositions, and systems according to the disclosure. Below are exemplary descriptions of sequences therein.

[0037] MG36

[0038] SEQ ID NO: 1 shows a full-length peptide sequence of a MG36 Cas effector.

[0039] SEQ ID NOs: 2-5 show peptide sequences of MG36 transposition proteins that may comprise a recombinase or transposase complex associated with a MG36 Cas effector. The addition of -B1, -B2, -T1, and -C to the end of the labels denotes similarity to TnsB1, TnsB2, TnsT1, and TniC proteins of Tn7-like systems, respectively.

[0040] SEQ ID NO: 11 shows a nucleotide sequence of an sgRNA engineered to function with an MG36 Cas effector.

[0041] SEQ ID NO: 12 shows a nucleotide sequence of a MG36 tracrRNAs derived from the same loci as a MG36 Cas effector.

[0042] SEQ ID NOs: 17-18 show nucleotide sequences of left-hand transposase recognition sequences associated with a MG36 system.

[0043] SEQ ID NO: 19 shows a nucleotide sequence of a right-hand transposase recognition sequence associated with a MG36 system.

[0044] MG39

[0045] SEQ ID NO: 6 shows the full-length peptide sequence of a MG39-1 Cas effector.

[0046] SEQ ID Nos: 7-10 show the peptide sequences of MG39-1 transposition proteins that may comprise a recombinase or transposase complex associated with the MG39-1 Cas effector.

[0047] SEQ ID NOs: 13-16 show nucleotide sequences of MG39 tracrRNAs derived from the same loci as a MG39 Cas effector.

[0048] SEQ ID NO: 20 shows a nucleotide sequence of a left-hand transposase recognition sequence associated with a MG39 system.

[0049] SEQ ID NO: 21 shows a nucleotide sequence of a right-hand transposase recognition sequence associated with a MG39 system.

[0050] MG64

[0051] SEQ ID NOs: 22, 26, 30, 34, 55-89, 104, and 147 show the full-length peptide sequences of MG64 Cas effectors.

[0052] SEQ ID NOs: 23-25, 27-29, 31-33, 35-37, 101-103, 105-107, and 148-150 show the peptide sequences of MG64 transposition proteins that may comprise a recombinase or transposase complex associated with MG64 Cas effectors. The addition of -A, -B, -C, and -Q to the end of the labels denotes similarity to TnsA, TnsB, TnsC, and TniQ proteins of Tn7-like systems, respectively.

[0053] SEQ ID NOs: 90-93, 117, 151, 156-181, and 209-234 show nucleotide sequences of MG64 tracrRNAs derived from the same loci as a MG64 effector.

[0054] SEQ ID NOs: 94-97, 119, 152, and 184-200 show nucleotide sequences of MG64 target CRISPR repeats.

[0055] SEQ ID NOs: 237-259 show nucleotide sequences of MG64 crRNAs.

[0056] Seq ID NOs: 111-114 and 201-204 show nucleotide sequences of single guide RNAs engineered to function with MG64 Cas effectors.

[0057] SEQ ID NOs: 123, 125, 127, 129, 131, 133, and 153 show nucleotide sequences of left-hand transposase recognition sequences associated with a MG64 system.

[0058] SEQ ID NOs:124, 126, 128, 130, 132, 154, and 155 show nucleotide sequences of right-hand transposase recognition sequences associated with a MG64 system.

[0059] MG108

[0060] SEQ ID NOs: 38, and 108 show the full-length peptide sequences of MG108 Cas effectors.

[0061] SEQ ID NOs: 39-40 and 109-110 show the peptide sequences of MG108 transposition proteins that may comprise a recombinase or transposase complex associated with MG108 Cas effectors. The addition of -A, -B, -C, and -Q to the end of the labels denotes similarity to TnsA, TnsB, TnsC, and TniQ proteins of Tn7-like systems, respectively.

[0062] SEQ ID NO: 98 and 120 show nucleotide sequences of MG108 target CRISPR repeats.

[0063] SEQ ID NO: 260-261 show nucleotide sequences of MG108 crRNAs.

[0064] Seq ID NOs: 115-116 and 205-206 show nucleotide sequences of single guide RNAs engineered to function with MG108 Cas effectors.

[0065] SEQ ID NOs: 118, 182-183, and 235-236 show nucleotide sequences of MG108 tracrRNAs derived from the same loci as a MG108 effector.

[0066] SEQ ID NO: 134 shows a nucleotide sequence of a left-hand transposase recognition sequence associated with a MG108 system.

[0067] SEQ ID NO: 135 shows a nucleotide sequence of a right-hand transposase recognition sequence associated with a MG108 system.

[0068] MG110

[0069] SEQ ID NOs: 41-43 and 48-50 show the full-length peptide sequences of MG110 Cas effectors. The addition of -6, -7, and -8 to the end of the labels denotes similarity to cas6, cas7, and cas8 proteins of class I, type I-F systems, respectively.

[0070] SEQ ID NOs: 44-47 and 51-54 show the peptide sequences of MG110 transposition proteins that may comprise a recombinase or transposase complex associated with MG110 Cas effectors. The addition of -A, -B, -C, and -Q to the end of the labels denotes similarity to TnsA, TnsB, TnsC, and TniQ proteins of Tn7-like systems, respectively.

[0071] SEQ ID NOs: 99-100 show nucleotide sequences of MG110 target CRISPR repeats.

[0072] SEQ ID NOs: 121-122 and 207-208 show nucleotide sequences of MG110 crRNAs.

[0073] SEQ ID NOs: 136 and 138 show nucleotide sequences of left-hand transposase recognition sequences associated with a MG110 system.

[0074] SEQ ID NOs: 137 and 139 show nucleotide sequences of right-hand transposase recognition sequences associated with a MG110 system.

Other Sequences

[0075] SEQ ID NOs: 140-141 show peptide sequences of nuclear localizing signals.

[0076] SEQ ID NOs: 142-143 show peptide sequences of linkers.

[0077] SEQ ID NOs: 144-146 show peptide sequences of epitope tags.

DETAILED DESCRIPTION

[0078] While various embodiments of the invention have been shown and described herein, it

will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions may occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed.

[0079] The practice of some methods disclosed herein employ, unless otherwise indicated, techniques of immunology, biochemistry, chemistry, molecular biology, microbiology, cell biology, genomics, and recombinant DNA. See for example Sambrook and Green, Molecular Cloning: A Laboratory Manual, 4th Edition (2012); the series Current Protocols in Molecular Biology (F. M. Ausubel, et al. eds.); the series Methods In Enzymology (Academic Press, Inc.), PCR 2: A Practical Approach (M.J. MacPherson, B.D. Hames and G.R. Taylor eds. (1995)), Harlow and Lane, eds. (1988) Antibodies, A Laboratory Manual, and Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications, 6th Edition (R.I. Freshney, ed. (2010)) (which is entirely incorporated by reference herein).

[0080] As used herein, the singular forms "a", "an" and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise. Furthermore, to the extent that the terms "including", "includes", "having", "has", "with", or variants thereof are used in either the detailed description and/or the claims, such terms are intended to be inclusive in a manner similar to the term "comprising".

[0081] The term "about" or "approximately" means within an acceptable error range for the particular value as determined by one of ordinary skill in the art, which will depend in part on how the value is measured or determined, i.e., the limitations of the measurement system. For example, "about" can mean within one or more than one standard deviation, per the practice in the art. Alternatively, "about" can mean a range of up to 20%, up to 15%, up to 10%, up to 5%, or up to 1% of a given value.

[0082] As used herein, a "cell" generally refers to a biological cell. A cell may be the basic structural, functional and/or biological unit of a living organism. A cell may originate from any organism having one or more cells. Some non-limiting examples include: a prokaryotic cell, eukaryotic cell, a bacterial cell, an archaeal cell, a cell of a single-cell eukaryotic organism, a protozoa cell, a cell from a plant (e.g., cells from plant crops, fruits, vegetables, grains, soy bean, corn, maize, wheat, seeds, tomatoes, rice, cassava, sugarcane, pumpkin, hay, potatoes, cotton, cannabis, tobacco, flowering plants, conifers, gymnosperms, ferns, clubmosses, hornworts, liverworts, mosses), an algal cell, (e.g.,, Botryococcus braunii, Chlamydomonas reinhardtii, Nannochloropsis gaditana, Chlorella pyrenoidosa, Sargassum patens C. Agardh, and the like), seaweeds (e.g., kelp), a fungal cell (e.g.,, a yeast cell, a cell from a mushroom), an animal cell, a

cell from an invertebrate animal (e.g., fruit fly, cnidarian, echinoderm, nematode, etc.), a cell from a vertebrate animal (e.g., fish, amphibian, reptile, bird, mammal), a cell from a mammal (e.g., a pig, a cow, a goat, a sheep, a rodent, a rat, a mouse, a non-human primate, a human, etc.), and etcetera. Sometimes a cell is not originating from a natural organism (e.g., a cell can be a synthetically made, sometimes termed an artificial cell).

[0083] The term "nucleotide," as used herein, generally refers to a base-sugar-phosphate combination. A nucleotide may comprise a synthetic nucleotide. A nucleotide may comprise a synthetic nucleotide analog. Nucleotides may be monomeric units of a nucleic acid sequence (e.g., deoxyribonucleic acid (DNA) and ribonucleic acid (RNA)). The term nucleotide may include ribonucleoside triphosphates adenosine triphosphate (ATP), uridine triphosphate (UTP), cytosine triphosphate (CTP), guanosine triphosphate (GTP) and deoxyribonucleoside triphosphates such as dATP, dCTP, dITP, dUTP, dGTP, dTTP, or derivatives thereof. Such derivatives may include, for example, [αS]dATP, 7-deaza-dGTP and 7-deaza-dATP, and nucleotide derivatives that confer nuclease resistance on the nucleic acid molecule containing them. The term nucleotide as used herein may refer to dideoxyribonucleoside triphosphates (ddNTPs) and their derivatives. Illustrative examples of dideoxyribonucleoside triphosphates may include, but are not limited to, ddATP, ddCTP, ddGTP, ddITP, and ddTTP. A nucleotide may be unlabeled or detectably labeled, such as using moieties comprising optically detectable moieties (e.g., fluorophores). Labeling may also be carried out with quantum dots. Detectable labels may include, for example, radioactive isotopes, fluorescent labels, chemiluminescent labels, bioluminescent labels, and enzyme labels. Fluorescent labels of nucleotides may include but are not limited fluorescein, 5-carboxyfluorescein (FAM), 2′7′-dimethoxy-4′5-dichloro-6-carboxyfluorescein (JOE), rhodamine, 6-carboxyrhodamine (R6G), N,N,N′,N′-tetramethyl-6-carboxyrhodamine (TAMRA), 6-carboxy-X-rhodamine (ROX), 4-(4′dimethylaminophenylazo) benzoic acid (DABCYL), Cascade Blue, Oregon Green, Texas Red, Cyanine and 5-(2′-aminoethyl)aminonaphthalene-1-sulfonic acid (EDANS). Specific examples of fluorescently labeled nucleotides can include [R6G]dUTP, [TAMRA]dUTP, [R110]dCTP, [R6G]dCTP, [TAMRA]dCTP, [JOE]ddATP, [R6G]ddATP, [FAM]ddCTP, [R110]ddCTP, [TAMRA]ddGTP, [ROX]ddTTP, [dR6G]ddATP, [dR110]ddCTP, [dTAMRA]ddGTP, and [dROX]ddTTP available from Perkin Elmer, Foster City, Calif; FluoroLink DeoxyNucleotides, FluoroLink Cy3-dCTP, FluoroLink Cy5-dCTP, FluoroLink Fluor X-dCTP, FluoroLink Cy3-dUTP, and FluoroLink Cy5-dUTP available from Amersham, Arlington Heights, Ill.; Fluorescein-15-dATP, Fluorescein-12-dUTP, Tetramethyl-rodamine-6-dUTP, IR770-9-dATP, Fluorescein-12-ddUTP, Fluorescein-12-UTP, and Fluorescein-15-2′-dATP available from Boehringer Mannheim,

Indianapolis, Ind.; and Chromosome Labeled Nucleotides, BODIPY-FL-14-UTP, BODIPY-FL-4-UTP, BODIPY-TMR-14-UTP, BODIPY-TMR-14-dUTP, BODIPY-TR-14-UTP, BODIPY-TR-14-dUTP, Cascade Blue-7-UTP, Cascade Blue-7-dUTP, fluorescein-12-UTP, fluorescein-12-dUTP, Oregon Green 488-5-dUTP, Rhodamine Green-5-UTP, Rhodamine Green-5-dUTP, tetramethylrhodamine-6-UTP, tetramethylrhodamine-6-dUTP, Texas Red-5-UTP, Texas Red-5-dUTP, and Texas Red-12-dUTP available from Molecular Probes, Eugene, Oreg. Nucleotides can also be labeled or marked by chemical modification. A chemically-modified single nucleotide can be biotin-dNTP. Some non-limiting examples of biotinylated dNTPs can include, biotin-dATP (e.g., bio-N6-ddATP, biotin-14-dATP), biotin-dCTP (e.g., biotin-11-dCTP, biotin-14-dCTP), and biotin-dUTP (e.g., biotin-11-dUTP, biotin-16-dUTP, biotin-20-dUTP).

[0084] The terms "polynucleotide," "oligonucleotide," and "nucleic acid" are used interchangeably to generally refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof, either in single-, double-, or multi-stranded form. A polynucleotide may be exogenous or endogenous to a cell. A polynucleotide may exist in a cell-free environment. A polynucleotide may be a gene or fragment thereof. A polynucleotide may be DNA. A polynucleotide may be RNA. A polynucleotide may have any three-dimensional structure and may perform any function. A polynucleotide may comprise one or more analogs (e.g., altered backbone, sugar, or nucleobase). If present, modifications to the nucleotide structure may be imparted before or after assembly of the polymer. Some non-limiting examples of analogs include: 5-bromouracil, peptide nucleic acid, xeno nucleic acid, morpholinos, locked nucleic acids, glycol nucleic acids, threose nucleic acids, dideoxynucleotides, cordycepin, 7-deaza-GTP, fluorophores (e.g., rhodamine or fluorescein linked to the sugar), thiol containing nucleotides, biotin linked nucleotides, fluorescent base analogs, CpG islands, methyl-7-guanosine, methylated nucleotides, inosine, thiouridine, pseudouridine, dihydrouridine, queuosine, and wyosine. Non-limiting examples of polynucleotides include coding or non-coding regions of a gene or gene fragment, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), short interfering RNA (siRNA), short-hairpin RNA (shRNA), micro-RNA (miRNA), ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, cell-free polynucleotides including cell-free DNA (cfDNA) and cell-free RNA (cfRNA), nucleic acid probes, and primers. The sequence of nucleotides may be interrupted by non-nucleotide components.

[0085] The terms "transfection" or "transfected" generally refer to introduction of a nucleic acid

into a cell by non-viral or viral-based methods. The nucleic acid molecules may be gene sequences encoding complete proteins or functional portions thereof. See, e.g., Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 18.1-18.88.

[0086] The terms "peptide," "polypeptide," and "protein" are used interchangeably herein to generally refer to a polymer of at least two amino acid residues joined by peptide bond(s). This term does not connote a specific length of polymer, nor is it intended to imply or distinguish whether the peptide is produced using recombinant techniques, chemical or enzymatic synthesis, or is naturally occurring. The terms apply to naturally occurring amino acid polymers as well as amino acid polymers comprising at least one modified amino acid. In some cases, the polymer may be interrupted by non-amino acids. The terms include amino acid chains of any length, including full length proteins, and proteins with or without secondary and/or tertiary structure (e.g., domains). The terms also encompass an amino acid polymer that has been modified, for example, by disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, oxidation, and any other manipulation such as conjugation with a labeling component. The terms "amino acid" and "amino acids," as used herein, generally refer to natural and non-natural amino acids, including, but not limited to, modified amino acids and amino acid analogues. Modified amino acids may include natural amino acids and non-natural amino acids, which have been chemically modified to include a group or a chemical moiety not naturally present on the amino acid. Amino acid analogues may refer to amino acid derivatives. The term "amino acid" includes both D-amino acids and L-amino acids.

[0087] As used herein, the "non-native" can generally refer to a nucleic acid or polypeptide sequence that is not found in a native nucleic acid or protein. Non-native may refer to affinity tags. Non-native may refer to fusions. Non-native may refer to a naturally occurring nucleic acid or polypeptide sequence that comprises mutations, insertions and/or deletions. A non-native sequence may exhibit and/or encode for an activity (e.g., enzymatic activity, methyltransferase activity, acetyltransferase activity, kinase activity, ubiquitinating activity, etc.) that may also be exhibited by the nucleic acid and/or polypeptide sequence to which the non-native sequence is fused. A non-native nucleic acid or polypeptide sequence may be linked to a naturally-occurring nucleic acid or polypeptide sequence (or a variant thereof) by genetic engineering to generate a chimeric nucleic acid and/or polypeptide sequence encoding a chimeric nucleic acid and/or polypeptide.

[0088] The term "promoter", as used herein, generally refers to the regulatory DNA region which controls transcription or expression of a gene and which may be located adjacent to or overlapping a nucleotide or region of nucleotides at which RNA transcription is initiated. A

promoter may contain specific DNA sequences which bind protein factors, often referred to as transcription factors, which facilitate binding of RNA polymerase to the DNA leading to gene transcription. A 'basal promoter', also referred to as a 'core promoter', may generally refer to a promoter that contains all the basic necessary elements to promote transcriptional expression of an operably linked polynucleotide. Eukaryotic basal promoters typically, though not necessarily, contain a TATA-box and/or a CAAT box.

[0089] The term "expression", as used herein, generally refers to the process by which a nucleic acid sequence or a polynucleotide is transcribed from a DNA template (such as into mRNA or other RNA transcript) and/or the process by which a transcribed mRNA is subsequently translated into peptides, polypeptides, or proteins. Transcripts and encoded polypeptides may be collectively referred to as "gene product." If the polynucleotide is derived from genomic DNA, expression may include splicing of the mRNA in a eukaryotic cell.

[0090] As used herein, "operably linked", "operable linkage", "operatively linked", or grammatical equivalents thereof generally refer to juxtaposition of genetic elements, e.g., a promoter, an enhancer, a polyadenylation sequence, etc., wherein the elements are in a relationship permitting them to operate in the expected manner. For instance, a regulatory element, which may comprise promoter and/or enhancer sequences, is operatively linked to a coding region if the regulatory element helps initiate transcription of the coding sequence. There may be intervening residues between the regulatory element and coding region so long as this functional relationship is maintained.

[0091] A "vector" as used herein, generally refers to a macromolecule or association of macromolecules that comprises or associates with a polynucleotide and which may be used to mediate delivery of the polynucleotide to a cell. Examples of vectors include plasmids, viral vectors, liposomes, and other gene delivery vehicles. The vector generally comprises genetic elements, e.g., regulatory elements, operatively linked to a gene to facilitate expression of the gene in a target.

[0092] As used herein, "an expression cassette" and "a nucleic acid cassette" are used interchangeably generally to refer to a combination of nucleic acid sequences or elements that are expressed together or are operably linked for expression. In some cases, an expression cassette refers to the combination of regulatory elements and a gene or genes to which they are operably linked for expression.

[0093] A "functional fragment" of a DNA or protein sequence generally refers to a fragment that retains a biological activity (either functional or structural) that is substantially similar to a biological activity of the full-length DNA or protein sequence. A biological activity of a DNA

sequence may be its ability to influence expression in a manner known to be attributed to the full-length sequence.

[0094] As used herein, an "engineered" object generally indicates that the object has been modified by human intervention. According to non-limiting examples: a nucleic acid may be modified by changing its sequence to a sequence that does not occur in nature; a nucleic acid may be modified by ligating it to a nucleic acid that it does not associate with in nature such that the ligated product possesses a function not present in the original nucleic acid; an engineered nucleic acid may synthesized in vitro with a sequence that does not exist in nature; a protein may be modified by changing its amino acid sequence to a sequence that does not exist in nature; an engineered protein may acquire a new function or property. An "engineered" system comprises at least one engineered component.

[0095] As used herein, "synthetic" and "artificial" are used interchangeably to refer to a protein or a domain thereof that has low sequence identity (e.g., less than 50% sequence identity, less than 25% sequence identity, less than 10% sequence identity, less than 5% sequence identity, less than 1% sequence identity) to a naturally occurring human protein. For example, VPR and VP64 domains are synthetic transactivation domains.

[0096] The term "tracrRNA" or "tracr sequence", as used herein, can generally refer to a nucleic acid with at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or 100% sequence identity and/or sequence similarity to a wild type exemplary tracrRNA sequence (e.g., a tracrRNA from S. pyogenes S. aureus, etc. or SEQ ID NOs: *_*). tracrRNA can refer to a nucleic acid with at most about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% sequence identity and/or sequence similarity to a wild type exemplary tracrRNA sequence (e.g., a tracrRNA from S. pyogenes S. aureus, etc). tracrRNA may refer to a modified form of a tracrRNA that can comprise a nucleotide change such as a deletion, insertion, or substitution, variant, mutation, or chimera. A tracrRNA may refer to a nucleic acid that can be at least about 60% identical to a wild type exemplary tracrRNA (e.g., a tracrRNA from S. pyogenes S. aureus, etc) sequence over a stretch of at least 6 contiguous nucleotides. For example, a tracrRNA sequence can be at least about 60% identical, at least about 65% identical, at least about 70% identical, at least about 75% identical, at least about 80% identical, at least about 85% identical, at least about 90% identical, at least about 95% identical, at least about 98% identical, at least about 99% identical, or 100 % identical to a wild type exemplary tracrRNA (e.g., a tracrRNA from S. pyogenes S. aureus, etc) sequence over a stretch of at least 6 contiguous nucleotides. Type II tracrRNA sequences can be predicted on a genome sequence by identifying regions with complementarity to part of the repeat sequence in an adjacent CRISPR array.

[0097] As used herein, a "guide nucleic acid" can generally refer to a nucleic acid that may hybridize to another nucleic acid. A guide nucleic acid may be RNA. A guide nucleic acid may be DNA. The guide nucleic acid may be programmed to bind to a sequence of nucleic acid site-specifically. The nucleic acid to be targeted, or the target nucleic acid, may comprise nucleotides. The guide nucleic acid may comprise nucleotides. A portion of the target nucleic acid may be complementary to a portion of the guide nucleic acid. The strand of a double-stranded target polynucleotide that is complementary to and hybridizes with the guide nucleic acid may be called the complementary strand. The strand of the double-stranded target polynucleotide that is complementary to the complementary strand, and therefore may not be complementary to the guide nucleic acid may be called noncomplementary strand. A guide nucleic acid may comprise a polynucleotide chain and can be called a "single guide nucleic acid." A guide nucleic acid may comprise two polynucleotide chains and may be called a "double guide nucleic acid." If not otherwise specified, the term "guide nucleic acid" may be inclusive, referring to both single guide nucleic acids and double guide nucleic acids. A guide nucleic acid may comprise a segment that can be referred to as a "nucleic acid-targeting segment" or a "nucleic acid-targeting sequence." A nucleic acid-targeting segment may comprise a sub-segment that may be referred to as a "protein binding segment" or "protein binding sequence" or "Cas protein binding segment".

[0098] The term "sequence identity" or "percent identity" in the context of two or more nucleic acids or polypeptide sequences, generally refers to two (e.g., in a pairwise alignment) or more (e.g., in a multiple sequence alignment) sequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence over a local or global comparison window, as measured using a sequence comparison algorithm. Suitable sequence comparison algorithms for polypeptide sequences include, e.g., BLASTP using parameters of a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix setting gap costs at existence of 11, extension of 1, and using a conditional compositional score matrix adjustment for polypeptide sequences longer than 30 residues; BLASTP using parameters of a wordlength (W) of 2, an expectation (E) of 1000000, and the PAM30 scoring matrix setting gap costs at 9 to open gaps and 1 to extend gaps for sequences of less than 30 residues (these are the default parameters for BLASTP in the BLAST suite available at https://blast.ncbi.nlm.nih.gov); CLUSTALW with parameters of ; the Smith-Waterman homology search algorithm with parameters of a match of 2, a mismatch of -1, and a gap of -1; MUSCLE with default parameters; MAFFT with parameters retree of 2 and maxiterations of 1000; Novafold with default parameters; HMMER hmmalign with default

parameters.

[0099] Included in the current disclosure are variants of any of the enzymes described herein with one or more conservative amino acid substitutions. Such conservative substitutions can be made in the amino acid sequence of a polypeptide without disrupting the three-dimensional structure or function of the polypeptide. Conservative substitutions can be accomplished by substituting amino acids with similar hydrophobicity, polarity, and R chain length for one another. Additionally, or alternatively, by comparing aligned sequences of homologous proteins from different species, conservative substitutions can be identified by locating amino acid residues that have been mutated between species (e.g. non-conserved residues without altering the basic functions of the encoded proteins. Such conservatively substituted variants may include variants with at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity any one of the systems described herein (e.g., MG36 or MG39 systems described herein). In some embodiments, such conservatively substituted variants are functional variants. Such functional variants can encompass sequences with substitutions such that the activity of critical active site residues of the endonuclease is not disrupted. In some embodiments, a functional variant of any of the systems described herein lack substitution of at least one of the conserved or functional residues called out in FIGs. 4 and 5. In some embodiments, a functional variant of any of the systems described herein lacks substitution of all of the conserved or functional residues called out in FIGs. 4 and 5.

[00100] Conservative substitution tables providing functionally similar amino acids are available from a variety of references (see, for example, Creighton, Proteins: Structures and Molecular Properties (W H Freeman & Co.; 2nd Edition (December 1993))). The following eight groups each contain amino acids that are conservative substitutions for one another:

1) Alanine (A), Glycine (G);
2) Aspartic acid (D), Glutamic acid (E);
3) Asparagine (N), Glutamine (Q);
4) Arginine (R), Lysine (K);
5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V);
6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W);
7) Serine (S), Threonine (T); and

8) Cysteine (C), Methionine (M).

[00101] As used herein, the term "RuvC_III domain" generally refers to a third discontinuous segment of a RuvC endonuclease domain (the RuvC nuclease domain being comprised of three discontiguous segments, RuvC_I, RuvC_II, and RuvC_III). A RuvC domain or segments thereof can generally be identified by alignment to known domain sequences, structural alignment to proteins with annotated domains, or by comparison to Hidden Markov Models (HMMs) built based on known domain sequences (e.g., Pfam HMM PF18541 for RuvC_III).

[00102] As used herein, the term "HNH domain" generally refers to an endonuclease domain having characteristic histidine and asparagine residues. An HNH domain can generally be identified by alignment to known domain sequences, structural alignment to proteins with annotated domains, or by comparison to Hidden Markov Models (HMMs) built based on known domain sequences (e.g., Pfam HMM PF01844 for domain HNH).

[00103] As used herein, the term "recombinase" generally refers to a site-specific enzyme that mediates the recombination of DNA between recombinase recognition sequences, which results in the excision, integration, inversion, or exchange (e.g., translocation) of DNA fragments between the recombinase recognition sequences.

[00104] As used herein, the term "recombine," or "recombination," in the context of a nucleic acid modification (e.g., a genomic modification), generally refers to the process by which two or more nucleic acid molecules, or two or more regions of a single nucleic acid molecule, are modified by the action of a recombinase protein. Recombination can result in, *inter alia*, the insertion, inversion, excision, or translocation of a nucleic acid sequence, e.g., in or between one or more nucleic acid molecules.

[00105] As used herein, the term "transposon" generally refers to mobile elements that move in and out of genomes carrying "cargo DNA" with them. In some cases, these transposons may differ on the type of nucleic acid to transpose, the type of repeat at the ends of the transposon, the type of cargo to be carried or by the mode of transposition (i.e. self-repair or host-repair). As used herein, the term "transposase" or "transposases" generally refers to an enzyme that binds to the end of a transposon and catalyzes its movement to another part of the genome. In some cases, the movement may be by a cut and paste mechanism or a replicative transposition

[00106] As used herein, the term "Tn7" or "Tn7-like transposase" generally refers to a family of transposases comprising three main components: a heteromeric transposase (TnsA and/or TnsB) alongside a regulator protein (TnsC). In addition to the TnsABC transposition proteins, Tn7 elements can encode dedicated target site-selection proteins, TnsD and TnsE. In conjunction with TnsABC, the sequence-specific DNA-binding protein TnsD directs transposition into a

conserved site referred to as the "Tn7 attachment site," attTn7. TnsD is a member of a large family of proteins that also includes TniQ. TniQ has been shown to target transposition into resolution sites of plasmids.

[00107] In some cases, the CAST systems described herein may comprise one or more Tn7 or Tn7 like transposases. In certain example embodiments, the Tn7 or Tn7 like transposase comprises a multimeric protein complex. In certain example embodiments, the multimeric protein complex comprises TnsA, TnsB, TnsC, or TniQ. In these combinations, the transposases (TnsA, TnsB, TnsC, TniQ) may form complexes or fusion proteins with each other.

[00108] As used herein, the term "Cas12k"(alternatively "class II, type V-K") generally refers to a subtype of Type V CRISPR systems that have been found to be defective in nuclease activity (e.g. they may comprise at least one defective RuvC domain that lacking at least one catalytic residue important for DNA cleavage). Such subtype of effectors have been generally associated with CAST systems.

[00109] As used herein, the term "type I-F" (alternatively class I, type I-F CRISPR) generally refers to a subtype of class I, type I CRISPR systems. Such systems generally comprise multi-component CRISPR effectors comprising Cas8, Cas7, and Cas6 proteins. In some cases, such systems are found associated with CAST systems. In some cases, type I-F CRISPR systems comprise crRNAs comprising an 8-nt 5′ handles for Cas8 and/or Cas5 binding, 32-nt spacers bound by six copies of Cas7 for target recognition, or a 20-nt 3′ hairpins for Cas6 binding and pre-crRNA processing. In some cases, type-F systems utilize a 5′-CC PAM on the non-target strand for target binding.

[00110] *Overview*

[00111] The discovery of new Cas enzymes with unique functionality and structure may offer the potential to further disrupt deoxyribonucleic acid (DNA) editing technologies, improving speed, specificity, functionality, and ease of use. Relative to the predicted prevalence of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) systems in microbes and the sheer diversity of microbial species, relatively few functionally characterized CRISPR/Cas enzymes exist in the literature. This is partly because a huge number of microbial species may not be readily cultivated in laboratory conditions. Metagenomic sequencing from natural environmental niches that represent large numbers of microbial species may offer the potential to drastically increase the number of new CRISPR/Cas systems known and speed the discovery of new oligonucleotide editing functionalities. A recent example of the fruitfulness of such an approach is demonstrated by the 2016 discovery of CasX/CasY CRISPR systems from metagenomic analysis of natural microbial communities.

[00112] CRISPR/Cas systems are RNA-directed nuclease complexes that have been described to function as an adaptive immune system in microbes. In their natural context, CRISPR/Cas systems occur in CRISPR (clustered regularly interspaced short palindromic repeats) operons or loci, which generally comprise two parts: (i) an array of short repetitive sequences (30-40bp) separated by equally short spacer sequences, which encode the RNA-based targeting element; and (ii) ORFs encoding the Cas encoding the nuclease polypeptide directed by the RNA-based targeting element alongside accessory proteins/enzymes. Efficient nuclease targeting of a particular target nucleic acid sequence generally requires both (i) complementary hybridization between the first 6-8 nucleic acids of the target (the target seed) and the crRNA guide; and (ii) the presence of a protospacer-adjacent motif (PAM) sequence within a defined vicinity of the target seed (the PAM usually being a sequence not commonly represented within the host genome). Depending on the exact function and organization of the system, CRISPR-Cas systems are commonly organized into 2 classes, 5 types and 16 subtypes based on shared functional characteristics and evolutionary similarity (see **FIG. 1**).

[00113] Class I CRISPR-Cas systems have large, multisubunit effector complexes, and comprise Types I, III, and IV.

[00114] Type I CRISPR-Cas systems are considered of moderate complexity in terms of components. In Type I CRISPR-Cas systems, the array of RNA-targeting elements is transcribed as a long precursor crRNA (pre-crRNA) that is processed at repeat elements to liberate short, mature crRNAs that direct the nuclease complex to nucleic acid targets when they are followed by a suitable short consensus sequence called a protospacer-adjacent motif (PAM). This processing occurs via an endoribonuclease subunit (Cas6) of a large endonuclease complex called Cascade, which also comprises a nuclease (Cas3) protein component of the crRNA-directed nuclease complex. Cas I nucleases function primarily as DNA nucleases.

[00115] Type III CRISPR systems may be characterized by the presence of a central nuclease, known as Cas10, alongside a repeat-associated mysterious protein (RAMP) that comprises Csm or Cmr protein subunits. Like in Type I systems, the mature crRNA is processed from a pre-crRNA using a Cas6-like enzyme. Unlike type I and II systems, type III systems appear to target and cleave DNA-RNA duplexes (such as DNA strands being used as templates for an RNA polymerase).

[00116] Type IV CRISPR-Cas systems possess an effector complex that consists of a highly reduced large subunit nuclease (csf1), two genes for RAMP proteins of the Cas5 (csf3) and Cas7 (csf2) groups, and, in some cases, a gene for a predicted small subunit; such systems are commonly found on endogenous plasmids.

[00117] Class II CRISPR-Cas systems generally have single-polypeptide multidomain nuclease effectors, and comprise Types II, V and VI.

[00118] Type II CRISPR-Cas systems are considered the simplest in terms of components. In Type II CRISPR-Cas systems, the processing of the CRISPR array into mature crRNAs does not require the presence of a special endonuclease subunit, but rather a small trans-encoded crRNA (tracrRNA) with a region complementary to the array repeat sequence; the tracrRNA interacts with both its corresponding effector nuclease (e.g. Cas9) and the repeat sequence to form a precursor dsRNA structure, which is cleaved by endogenous RNAse III to generate a mature effector enzyme loaded with both tracrRNA and crRNA. Cas II nucleases are known as DNA nucleases. Type 2 effectors generally exhibit a structure consisting of a RuvC-like endonuclease domain that adopts the RNase H fold with an unrelated HNH nuclease domain inserted within the folds of the RuvC-like nuclease domain. The RuvC-like domain is responsible for the cleavage of the target (e.g., crRNA complementary) DNA strand, while the HNH domain is responsible for cleavage of the displaced DNA strand.

[00119] Type V CRISPR-Cas systems are characterized by a nuclease effector (e.g. Cas12) structure similar to that of Type II effectors, comprising a RuvC-like domain. Similar to Type II, most (but not all) Type V CRISPR systems use a tracrRNA to process pre-crRNAs into mature crRNAs; however, unlike Type II systems which requires RNAse III to cleave the pre-crRNA into multiple crRNAs, type V systems are capable of using the effector nuclease itself to cleave pre-crRNAs. Like Type-II CRISPR-Cas systems, Type V CRISPR-Cas systems are again known as DNA nucleases. Unlike Type II CRISPR-Cas systems, some Type V enzymes (e.g., Cas12a) appear to have a robust single-stranded nonspecific deoxyribonuclease activity that is activated by the first crRNA directed cleavage of a double-stranded target sequence.

[00120] Type VI CRIPSR-Cas systems have RNA-guided RNA endonucleases. Instead of RuvC-like domains, the single polypeptide effector of Type VI systems (e.g. Cas13) comprises two HEPN ribonuclease domains. Differing from both Type II and V systems, Type VI systems also appear to not need a tracrRNA for processing of pre-crRNA into crRNA. Similar to type V systems, however, some Type VI systems (e.g., C2C2) appear to possess robust single-stranded nonspecific nuclease (ribonuclease) activity activated by the first crRNA directed cleavage of a target RNA.

[00121] Because of their simpler architecture, Class II CRISPR-Cas have been most widely adopted for engineering and development as designer nuclease/genome editing applications.

[00122] One of the early adaptations of such a system for in vitro use can be found in Jinek et al. (Science. 2012 Aug 17;337(6096):816-21, which is entirely incorporated herein by reference).

The Jinek study first described a system that involved (i) recombinantly-expressed, purified full-length Cas9 (e.g., a Class II, Type II Cas enzyme) isolated from *S. pyogenes* SF370, (ii) purified mature ~42 nt crRNA bearing a ~20 nt 5' sequence complementary to the target DNA sequence desired to be cleaved followed by a 3' tracr-binding sequence (the whole crRNA being in vitro transcribed from a synthetic DNA template carrying a T7 promoter sequence); (iii) purified tracrRNA in vitro transcribed from a synthetic DNA template carrying a T7 promoter sequence, and (iv) Mg^{2+}. Jinek later described an improved, engineered system wherein the crRNA of (ii) is joined to the 5' end of (iii) by a linker (e.g., GAAA) to form a single fused synthetic guide RNA (sgRNA) capable of directing Cas9 to a target by itself (compare top and bottom panel of **FIG. 2**).

[00123] Mali et al. (Science. 2013 Feb 15; 339(6121): 823–826.), which is entirely incorporated herein by reference, later adapted this system for use in mammalian cells by providing DNA vectors encoding (i) an ORF encoding codon-optimized Cas9 (e.g., a Class II, Type II Cas enzyme) under a suitable mammalian promoter with a C-terminal nuclear localization sequence (e.g., SV40 NLS) and a suitable polyadenylation signal (e.g., TK pA signal); and (ii) an ORF encoding an sgRNA (having a 5' sequence beginning with G followed by 20 nt of a complementary targeting nucleic acid sequence joined to a 3' tracr-binding sequence, a linker, and the tracrRNA sequence) under a suitable Polymerase III promoter (e.g., the U6 promoter).

[00124] Transposons are mobile elements that can move between positions in a genome. Such transposons have evolved to limit the negative effects they exert on the host. A variety of regulatory mechanisms are used to maintain transposition at a low frequency and sometimes coordinate transposition with various cell processes. Some prokaryotic transposons also can mobilize functions that benefit the host or otherwise help maintain the element. Certain transposons may have also evolved mechanisms of tight control over target site selection, the most notable example being the Tn7 family.

[00125] Transposon Tn7 and similar elements may be reservoirs for antibiotic resistance and pathogenesis functions in clinical settings, as well as encoding other adaptive functions in natural environments. The Tn7 system, for example, has evolved mechanisms to almost completely avoid integrating into important host genes, but also maximize dispersal of the element by recognizing mobile plasmids and bacteriophage capable of moving Tn7 between host bacteria.

[00126] Tn7 and Tn7-like elements may control where and when they insert, possessing one pathway that directs insertion into a single conserved position in bacterial genomes and a second pathway that appears to be adapted to maximizing targeting into mobile plasmids capable of

transporting the element between bacteria (see **FIG. 3**). The association between Tn7-like transposons and CRISPR-Cas systems suggests that the transposons might have hijacked CRISPR effectors to generate R-loops in target sites and facilitate the spread of transposons via plasmids and phages.

[00127] *MG36 Systems*

[00128] In one aspect, the present disclosure provides for a system for transposing a cargo nucleotide sequence to a target nucleic acid site. This system may comprise a first double-stranded nucleic acid. The first double-stranded nucleic acid may comprise the cargo nucleotide sequence wherein the cargo nucleotide sequence is configured to interact with a recombinase complex. The system may comprise a Cas effector complex. This Cas effector complex may comprise a class II, type II Cas effector and at least one engineered guide polynucleotide configured to hybridize to the target nucleic acid site. The class II, type II Cas effector can comprise a RuvC domain and an HNH domain. The system may comprise the recombinase or transposase complex, wherein the recombinase or transposase complex is configured to recruit the cargo nucleotide sequence to the target nucleic acid site.

[00129] In some cases, the cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence. In some cases, the cargo nucleotide sequence is flanked by a right-hand transposase recognition sequence. In some cases, the cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence. In some cases, the system further comprises a second double-stranded nucleic acid comprising the target nucleic acid site. In some cases, the system further comprises a PAM sequence compatible with the Cas effector complex adjacent to the target nucleic acid site. In some cases, the PAM sequence is located 3' of the target nucleic acid site. In some cases, the recombinase or transposase complex is a Tn7 type transposase complex. In some cases, the engineered guide polynucleotide is configured to bind the class II, type II Cas effector. In some cases, the class II, type II Cas effector comprises a polypeptide which has at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 1 or a variant thereof. In some cases, the class II, type II Case effector comprises a polypeptide substantially identical to SEQ ID NO: 1.

[00130] In some cases, the recombinase or transposase complex comprises at least one polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 2-5 or a variant thereof. In some cases, the recombinase or transposase complex comprises at least one polypeptide comprising a sequence substantially identical to any one of SEQ ID NOs: 2-5. In some cases, the recombinase or transposase complex comprises at least two polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 2-5 or a variant thereof. In some cases, the recombinase or transposase complex comprises at least two polypeptides comprising a sequence substantially identical to any one of SEQ ID NOs: 2-5 or a variant thereof. In some cases, the recombinase or transposase complex comprises at least three polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID Nos: 2-5 or a variant thereof. In some cases, the recombinase or transposase complex comprises at least three polypeptides comprising a sequence substantially identical to any one of SEQ ID NOs: 2-5 or a variant thereof. In some cases, the recombinase or transposase complex comprises four polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 2-5 or a

variant thereof. In some cases, the recombinase or transposase complex comprises four polypeptides comprising a sequence substantially identical to any one of SEQ ID NOs: 2-5 or a variant thereof. In some cases, the recombinase or transposase complex comprises a TnsB1 polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 2 or a variant thereof. In some cases, the recombinase or transposase complex comprises a TnsB1 polypeptide comprising a sequence substantially identical to SEQ ID NOs: 2 or a variant thereof. In some cases, the recombinase or transposase complex comprises a TnsB2 polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NO: 3 or a variant thereof. In some cases, the recombinase or transposase complex comprises a TnsB2 polypeptide comprising a sequence substantially identical to SEQ ID NO: 3 or a variant thereof. In some cases, the recombinase or transposase complex comprises a TnsT1 polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NO: 4 or a variant thereof. In some cases, the recombinase or transposase complex comprises a TnsT1 polypeptide comprising a sequence substantially identical to SEQ ID NO: 4 or a variant thereof. In some cases, the recombinase or transposase complex comprises a TnsC polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at

least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NO: 5 or a variant thereof. In some cases, the recombinase or transposase complex comprises a TnsC polypeptide comprising a sequence substantially identical to SEQ ID NO: 5 or a variant thereof.

[00131] In some cases, the engineered guide polynucleotide comprises a sequence comprising at least 60-80 consecutive nucleotides having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO:11 or a variant thereof.

[00132] In some cases, the left-hand recombinase sequence comprises a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 17-18 or a variant thereof. In some cases, the right-had recombinase sequence comprises a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 19 or a variant thereof.

[00133] In some cases, the class II, type II Cas effector and the recombinase or transposase complex are encoded by polynucleotide sequences comprising fewer than about 20 kilobases, fewer than about 15 kilobases, fewer than about 10 kilobases, or fewer than about 5 kilobases.

[00134] In one aspect, the present disclosure provides a method for transposing a cargo nucleotide sequence to a target nucleic acid site comprising a target nucleotide sequence comprising expressing a system described herein within a cell or introducing a system described herein to a cell.

MG39 Systems

[00135] In one aspect, the present disclosure provides for a system for transposing a cargo nucleotide sequence to a target nucleic acid site. The system may comprise a first double-stranded nucleic acid comprising a cargo nucleotide sequence. This cargo nucleotide sequence may be configured to interact with a Tn7 type transposase complex. The system may comprise a Cas effector complex. The Cas effector complex may comprise a class II, type V Cas effector and an engineered guide polynucleotide configured to hybridize to the target nucleotide sequence. The class II, type V Cas effector can comprise a RuvC domain. The system may comprise a Tn7 type transposase complex configured to bind the Cas effector complex, wherein the Tn7 type transposase complex comprises a TnsA subunit.

[00136] In some cases, the cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence. In some cases, the cargo nucleotide sequence is flanked by a right-hand transposase recognition sequence. In some cases, the cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence. In some cases, the system further comprises a second double-stranded nucleic acid comprising the target nucleic acid site. In some cases, the system further comprises a PAM sequence compatible with the Cas effector complex adjacent to the target nucleic acid site. In some cases, the PAM sequence is located 3' of the target nucleic acid site.

[00137] In some cases, the engineered guide polynucleotide is configured to bind the class II, type V Cas effector. In some cases, the class II, type V Cas effector comprises a polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 5, or a variant thereof. In some cases, the class II, type V Cas effector comprises a polypeptide comprising a sequence substantially identical to SEQ ID NO: 5 or a variant thereof. In some cases, the TnsA subunit comprises a polypeptide having a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 7or a variant thereof. In

some cases, the TnsA subunit comprises a polypeptide having a sequence substantially identical to SEQ ID NO: 7 or a variant thereof.

[00138] In some cases, the Tn7 type transposase complex comprises at least one polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 8-10, or a variant thereof. In some cases, the recombinase or transposase complex comprises at least one polypeptide comprising a sequence substantially identical to any one of SEQ ID NOs: 8-10, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least two polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 8-10, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least two polypeptides comprising a sequence substantially identical to any one of SEQ ID NOs: 8-10, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least three polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 8-10 or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least three polypeptides comprising a sequence substantially identical to any one of SEQ ID NOs: 8-10 or a variant thereof.

[00139] In some cases, the Tn7 type transposase complex comprises a TnsA polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%,

at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 7 or a variant thereof. In some cases, the Tn7 type transposase complex comprises a TnsA polypeptide comprising a sequence substantially identical to any one of SEQ ID NOs: 7 or a variant thereof. In some cases, the Tn7 type transposase complex comprises a TnsB polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 8 or a variant thereof. In some cases, the Tn7 type transposase complex comprises a TnsB polypeptide comprising a sequence substantially identical to any one of SEQ ID NOs: 8 or a variant thereof. In some cases, the Tn7 type transposase complex comprises a TnsC polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 9 or a variant thereof. In some cases, the Tn7 type transposase complex comprises a TnsC polypeptide comprising a sequence substantially identical to any one of SEQ ID NOs: 9 or a variant thereof. In some cases, the Tn7 type transposase complex comprises a TniQ polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 10 or a variant thereof. In some cases, the Tn7 type transposase complex comprises a TniQ polypeptide comprising a sequence substantially identical to any one of SEQ ID NOs: 10 or a variant thereof.

[00140] In some cases, the engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least about 20%, at least about 25%, at least

about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 13-16, or a variant thereof. In some cases, the engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides substantially identical to any one of SEQ ID NOs: 13-16 or a variant thereof.

[00141] In some cases, the left-hand recombinase sequence comprises a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 20, or a variant thereof. In some cases, the left-hand recombinase sequence comprises a sequence substantially identical to SEQ ID NO: 20, or a variant thereof.

[00142] In some cases, the right-hand recombinase sequence comprises a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 21, or a variant thereof. In some cases, the right-hand recombinase sequence comprises a sequence substantially identical to SEQ ID NO: 21, or a variant thereof.

[00143] In some cases, the class II, type V Cas effector and the Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 20 kilobases, fewer than about 15 kilobases, fewer than about 10 kilobases, or fewer than about 5 kilobases.

[00144] In one aspect, the present disclosure provides for a method for transposing a cargo nucleotide sequence to a target nucleic acid site comprising a target nucleotide sequence comprising expressing a system described herein within a cell or introducing a system described herein to a cell.

[00145] In one aspect, the present disclosure provides for a method for transposing a cargo nucleotide sequence to a target nucleic acid site, comprising contacting a first double-stranded nucleic acid comprising a cargo nucleotide sequence with a Cas effector complex. The Cas

effector complex may comprise a class II, type II Cas effector and at least one engineered guide polynucleotide configured to hybridize to the target nucleic acid site. The method may comprise contacting the first double-stranded nucleic acid comprising the cargo nucleotide sequence with a recombinase or transposase complex configured to recruit the cargo nucleotide to the target nucleic acid site. The method may comprise contacting the first double-stranded nucleic acid comprising the cargo nucleotide sequence with a second double-stranded nucleic acid comprising the target nucleic acid site.

[00146] In some cases, the cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence. In some cases, the cargo nucleotide sequence is flanked by a right-hand transposase recognition sequence. In some cases, the cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence. In some cases, the Cas effector complex further comprises a PAM sequence compatible with the Cas effector complex adjacent to the target nucleic acid site. In some cases, the PAM sequence is located 3' of the target nucleic acid site. In some cases, the PAM sequence is located 5' of the target nucleic acid site.

[00147] *MG64 Systems*

[00148] In one aspect, the present disclosure provides for a system for transposing a cargo nucleotide sequence to a target nucleic acid site. The system may comprise a first double-stranded nucleic acid comprising a cargo nucleotide sequence. This cargo nucleotide sequence may be configured to interact with a Tn7 type transposase complex. The system may comprise a Cas effector complex. The Cas effector complex may comprise a class II, type V Cas effector and an engineered guide polynucleotide configured to hybridize to the target nucleotide sequence. The system may comprise a Tn7 type transposase complex configured to bind the Cas effector complex. The class II, type V Cas effector can comprise a RuvC domain.

[00149] In some cases, the cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence. In some cases, the cargo nucleotide sequence is flanked by a right-hand transposase recognition sequence. In some cases, the cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence. In some cases, the system further comprises a second double-stranded nucleic acid comprising the target nucleic acid site. In some cases, the system further comprises a PAM sequence compatible with the Cas effector complex adjacent to the target nucleic acid site. In some cases, the PAM sequence is located 3' of the target nucleic acid site. In some cases, the PAM sequence is located 5' of the target nucleic acid site. In some cases, the PAM sequence comprises 5'-nGTn-3' or 5'-nGTt-3'.

[00150] In some cases, the engineered guide polynucleotide is configured to bind the class II, type V Cas effector. In some cases, the class II, type V Cas effector comprises a polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs:22, 26, 30, 34, 55-89, 104, or 147, or a variant thereof. In some cases, the class II, type V Cas effector comprises a polypeptide comprising a sequence substantially identical to any one of SEQ ID NOs:22, 26, 30, 34, 55-89, 104, or 147, or a variant thereof.

[00151] In some cases, the Tn7 type transposase complex comprises at least one polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs:23-25, 27-29, 31-33, 35-37, 101-103, 105-107, or 148-150, or a variant thereof. In some cases, the recombinase or transposase complex comprises at least one polypeptide comprising a sequence substantially identical to any one of SEQ ID NOs:23-25, 27-29, 31-33, 35-37, 101-103, 105-107, or 148-150, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least two polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs:23-25, 27-29, 31-33, 35-37, 101-103, 105-107, or 148-150, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least two polypeptides comprising a sequence substantially identical to any one of SEQ ID NOs:23-25, 27-29, 31-33, 35-37, 101-103, 105-107, or 148-150, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least three polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least

about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs:23-25, 27-29, 31-33, 35-37, 101-103, 105-107, or 148-150, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least three polypeptides comprising a sequence substantially identical to any one of SEQ ID NOs:23-25, 27-29, 31-33, 35-37, 101-103, 105-107, or 148-150, or a variant thereof.

[00152] In some cases, the Tn7 type transposase complex comprises TnsB, TnsC, and TniQ polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs:23-25, 27-29, 31-33, 35-37, 101-103, 105-107, or 148-150, or a variant thereof, respectively. In some cases, the Tn7 type transposase complex comprises a TnsB polypeptide comprising a sequence substantially identical to any one of SEQ ID NOs: 8 or a variant thereof. In some cases, the Tn7 type transposase complex comprises TnsB, TnsC, and TniQ polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs:23-25, 27-29, 31-33, 35-37, 101-103, 105-107, or 148-150, or a variant thereof, respectively.

[00153] In some cases, the engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 90, 91, 92, 93, 117, 151, 156-181, or 209-234, or a variant thereof. In some cases, the engineered guide

polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides substantially identical any one of SEQ ID NOs: 90, 91, 92, 93, 117, 151, 156-181, or 209-234, or a variant thereof.

[00154] In some cases, the engineered guide polynucleotide comprises a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to non-degenerate nucleotides of any one of SEQ ID NOs: 111-114 or 201- 204, or a variant thereof. In some cases, the engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides substantially identical to the non-degenerate nucleotides of any one of SEQ ID NOs: 111-114 or 201- 204, or a variant thereof.

[00155] In some cases, the left-hand recombinase sequence comprises a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 125, 127, 123, 129, 131, 133, 153, or 134, or a variant thereof. In some cases, the left-hand recombinase sequence comprises a sequence substantially identical to any one of SEQ ID NOs: 125, 127, 123, 129, 131, 133, 153, or 134, or a variant thereof.

[00156] In some cases, the right-hand recombinase sequence comprises a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 126, 155, 128, 124, 130, 132, or 154, or a variant thereof. In some cases, the right-hand recombinase sequence comprises a sequence substantially identical to any one of SEQ ID NOs: 126, 155, 128, 124, 130, 132, or 154, or a variant thereof.

[00157] In some cases, the class II, type V Cas effector and the Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 20 kilobases, fewer than about 15 kilobases, fewer than about 10 kilobases, or fewer than about 5 kilobases.

[00158] *MG108 Systems*

[00159] In one aspect, the present disclosure provides for a system for transposing a cargo nucleotide sequence to a target nucleic acid site. The system may comprise a first double-stranded nucleic acid comprising a cargo nucleotide sequence. This cargo nucleotide sequence may be configured to interact with a Tn7 type transposase complex. The system may comprise a Cas effector complex. The Cas effector complex may comprise a class II, type V Cas effector and an engineered guide polynucleotide configured to hybridize to the target nucleotide sequence. The class II, type V Cas effector can comprise a RuvC domain. The system may comprise a Tn7 type transposase complex configured to bind the Cas effector complex. In some cases, the Tn7 type transposase complex comprises TnsB and TnsC components but does not comprise a TnsA and/or TniQ component.

[00160] In some cases, the cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence. In some cases, the cargo nucleotide sequence is flanked by a right-hand transposase recognition sequence. In some cases, the cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence. In some cases, the system further comprises a second double-stranded nucleic acid comprising the target nucleic acid site. In some cases, the system further comprises a PAM sequence compatible with the Cas effector complex adjacent to the target nucleic acid site. In some cases, the PAM sequence is located 3' of the target nucleic acid site. In some cases, the PAM sequence is located 5' of the target nucleic acid site.

[00161] In some cases, the engineered guide polynucleotide is configured to bind the class II, type V Cas effector. In some cases, the class II, type V Cas effector comprises a polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 38 or SEQ ID NO:108, or a variant thereof. In some cases, the class II, type V Cas effector comprises a polypeptide comprising a sequence substantially identical to SEQ ID NO: 38 or SEQ ID NO:108, or a variant thereof.

[00162] In some cases, the Tn7 type transposase complex comprises at least one polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 39-40 or 109-110, or a variant thereof. In some cases, the recombinase or transposase complex comprises at least one polypeptide comprising a sequence substantially identical to any one of SEQ ID NOs: 39-40 or 109-110, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least two polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 39-40 or 109-110, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least two polypeptides comprising a sequence substantially identical to any one of SEQ ID NOs: 39-40 or 109-110, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least three polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 39-40 or 109-110, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least three polypeptides comprising a sequence substantially identical to any one of SEQ ID NOs: 39-40 or 109-110, or a variant thereof.

[00163] In some cases, the Tn7 type transposase complex comprises a TnsB component comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least

about 98%, or at least about 99% identity to any one of SEQ ID NOs: 40 or 109, or a variant thereof. In some cases, the recombinase or transposase complex comprises a TnsB component comprising a sequence substantially identical to any one of SEQ ID NOs: 40 or 109, or a variant thereof.

[00164] In some cases, the Tn7 type transposase complex comprises a TnsC component comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NOs: 39 or 110, or a variant thereof. In some cases, the recombinase or transposase complex comprises a TnsC component comprising a sequence substantially identical to any one of SEQ ID NOs: 39 or 110, or a variant thereof.

[00165] In some cases, the Tn7 type transposase complex comprises TnsB and TnsC components comprising sequences having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NOs: 40 and 39 or 109 and 110, or a variant thereof, respectively. In some cases, the recombinase or transposase complex comprises TnsB and TnsC components comprising sequences substantially identical to any one of SEQ ID NOs: 40 and 39 or 109 and 110, or a variant thereof, respectively.

[00166] In some cases, the engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 118, 182, 183, 235, or 236, or a variant thereof. In some cases, the engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides substantially identical any one of SEQ ID NOs: 118, 182, 183, 235, or 236, or a variant thereof.

[00167] In some cases, the engineered guide polynucleotide comprises a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to non-degenerate nucleotides of any one of SEQ ID NOs: 115, 116, 205, or 206, or a variant thereof. In some cases, the engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides substantially identical to the non-degenerate nucleotides of any one of SEQ ID NOs: 115, 116, 205, or 206, or a variant thereof.

[00168] In some cases, the left-hand recombinase sequence comprises a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NOs: 134, or a variant thereof. In some cases, the left-hand recombinase sequence comprises a sequence substantially identical to SEQ ID NO:134, or a variant thereof.

[00169] In some cases, the right-hand recombinase sequence comprises a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 135, or a variant thereof. In some cases, the right-hand recombinase sequence comprises a sequence substantially identical to SEQ ID NO: 135, or a variant thereof.

[00170] In some cases, the class II, type V Cas effector and the Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 20 kilobases, fewer than about 15 kilobases, fewer than about 10 kilobases, or fewer than about 5 kilobases.

[00171] *MG110 Systems*

[00172] In one aspect, the present disclosure provides for a system for transposing a cargo nucleotide sequence to a target nucleic acid site. The system may comprise a first double-stranded nucleic acid comprising a cargo nucleotide sequence. This cargo nucleotide sequence

may be configured to interact with a Tn7 type transposase complex. The system may comprise a Cas effector complex. The Cas effector complex may comprise a class I, type I Cas effector and an engineered guide polynucleotide configured to hybridize to the target nucleotide sequence. The system may comprise a Tn7 type transposase complex configured to bind the Cas effector complex.

[00173] In some cases, the cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence. In some cases, the cargo nucleotide sequence is flanked by a right-hand transposase recognition sequence. In some cases, the cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence. In some cases, the system further comprises a second double-stranded nucleic acid comprising the target nucleic acid site. In some cases, the system further comprises a PAM sequence compatible with the Cas effector complex adjacent to the target nucleic acid site. In some cases, the PAM sequence is located 3' of the target nucleic acid site.

[00174] In some cases, the engineered guide polynucleotide is configured to bind the class I, type I Cas effector. In some cases, the class I, type I Cas effector comprises a polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NO: 41-43, or 48-50, or a variant thereof. In some cases, the class I, type I Cas effector comprises a polypeptide comprising a sequence substantially identical to any one of SEQ ID NO: 41-43, or 48-50, or a variant thereof.

[00175] In some cases, the engineered guide polynucleotide is configured to bind the class I, type I Cas effector. In some cases, the class I, type I Cas effector comprises Cas6, Cas7, and Cas8 effectors comprising sequences having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NO: 41-43, or 48-50, or a variant thereof. In some cases, the class I, type I Cas effector comprises Cas6,

Cas7, and Cas8 effectors comprising sequences substantially identical to any one of SEQ ID NO: 41-43, or 48-50, or a variant thereof.

[00176] In some cases, the Tn7 type transposase complex comprises at least one polypeptide comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 44-47, or 51-54, or a variant thereof. In some cases, the recombinase or transposase complex comprises at least one polypeptide comprising a sequence substantially identical to any one of SEQ ID NOs: 44-47, or 51-54, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least two polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 44-47, or 51-54, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least two polypeptides comprising a sequence substantially identical to any one of SEQ ID NOs: 44-47, or 51-54, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least three polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 44-47, or 51-54, or a variant thereof. In some cases, the Tn7 type transposase complex comprises at least three polypeptides comprising a sequence substantially identical to any one of SEQ ID NOs: 44-47, or 51-54, or a variant thereof. In some cases, the Tn7 type transposase complex comprises four polypeptides comprising a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%,

at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 44-47, or 51-54, or a variant thereof, or a variant thereof. In some cases, Tn7 type transposase complex comprises four polypeptides comprising a sequence substantially identical to any one of SEQ ID NOs: 44-47, or 51-54, or a variant thereof.

[00177] In some cases, the Tn7 type transposase complex comprises TnsA, TnsB, TnsC, and TniQ components comprising sequences having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to any one of SEQ ID NOs: 44-47, or 51-54, or a variant thereof, respectively.

[00178] In some cases, the engineered guide polynucleotide comprises a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to non-degenerate nucleotides of any one of SEQ ID NOs: 121, 122, 207, or 208, or a variant thereof. In some cases, the engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides substantially identical to the non-degenerate nucleotides of any one of SEQ ID NOs: 121, 122, 207, or 208, or a variant thereof.

[00179] In some cases, the left-hand recombinase sequence comprises a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 136 or 138, or a variant thereof. In some cases, the left-hand recombinase sequence comprises a sequence substantially identical to SEQ ID NO: 136 or 138, or a variant thereof.

[00180] In some cases, the right-hand recombinase sequence comprises a sequence having at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%,

at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% identity to SEQ ID NO: 137 or 139, or a variant thereof, or a variant thereof. In some cases, the right-hand recombinase sequence comprises a sequence substantially identical to SEQ ID NO: 137 or 139, or a variant thereof, or a variant thereof.

[00181] In some cases, the class I, type I Cas effector and the Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 20 kilobases, fewer than about 15 kilobases, fewer than about 10 kilobases, or fewer than about 5 kilobases.

[00182] In one aspect, the present disclosure provides for a method for transposing a cargo nucleotide sequence to a target nucleic acid site comprising a target nucleotide sequence comprising expressing a system described herein within a cell or introducing a system described herein to a cell.

[00183] In accordance with IUPAC conventions, the following abbreviations are used throughout the examples:

A = adenine

C = cytosine

G = guanine

T = thymine

R = adenine or guanine

Y = cytosine or thymine

S = guanine or cytosine

W = adenine or thymine

K = guanine or thymine

M = adenine or cytosine

B = C, G, or T

D = A, G, or T

H = A, C, or T

V = A, C, or G

EXAMPLES

Example 1 – (General Protocol) PAM sequence identification/confirmation for systems described herein

[00184] Putative endonucleases were expressed in an *E. coli* lysate-based expression system (myTXTL, Arbor Biosciences). PAM sequences were determined by sequencing plasmids containing randomly-generated potential PAM sequences that could be cleaved by the putative nucleases. In this system, an *E. coli* codon optimized nucleotide sequence encoding the putative nuclease was transcribed and translated *in vitro* from a PCR fragment under control of a T7 promoter. A second PCR fragment with a minimal CRISPR array composed of a T7 promoter followed by a repeat-spacer-repeat sequence was transcribed in the same reaction. Successful expression of the endonuclease and repeat-spacer-repeat sequence in the TXTL system followed by CRISPR array processing provided active *in vitro* CRISPR nuclease complexes.

[00185] A library of target plasmids containing a spacer sequence matching that in the minimal array preceded by 8N mixed bases (potential PAM sequences) was incubated with the output of the TXTL reaction. After 1-3 hr, the reaction was stopped and the DNA was recovered via a DNA clean-up kit, e.g., Zymo DCC, AMPure XP beads, QiaQuick etc. Adapter sequences were blunt-end ligated to DNA with active PAM sequences that were cleaved by the endonuclease, whereas DNA that was not cleaved is inaccessible for ligation. DNA segments comprising active PAM sequences were then amplified by PCR with primers specific to the library and the adapter sequence. The PCR amplification products were resolved on a gel to identify amplicons that correspond to cleavage events. The amplified segments of the cleavage reaction were also used as templates for preparation of an NGS library or as a substrate for Sanger sequencing. Sequencing this resulting library, which is a subset of the starting 8N library, revealed sequences with PAM activity compatible with the CRISPR complex. For PAM testing with a processed RNA construct, the same procedure was repeated except that an *in vitro* transcribed RNA was added along with the plasmid library and the minimal CRISPR array template is omitted.

[00186] For endonucleases which are binding-competent but nuclease deficient, the PAM was determined via a modification of the above procedure. After expression in TXTL, the sgRNA or crRNA and PAM library were added. Upon binding of the effector in a sgRNA-dependent manner to the spacer sequence, the spacer sequence was sequestered within the effector protein. The appropriate restriction enzyme that targets within the spacer sequence was added and all unprotected plasmids within the library were cleaved. The uncleaved (endonuclease-bound)

members of the library which contain the PAM were identified by PCR and subsequent NGS library preparation of the band.

Example 2 – In vitro targeted integrase activity

[00187] Integrase activity was preferentially assayed with a previously identified PAM but may be conducted with a PAM library substrate instead, with reduced efficiency. One arrangement of components for *in vitro* testing involved three plasmids other than that containing the donor sequence: (1) an expression plasmid with effector (or effectors) under a T7 promoter; (2) an expression plasmid with integrase genes under a T7 promoter; a sgRNA or crRNA and tracrRNA; (3) a target plasmid which contained the spacer site and appropriate PAM; and (4) a donor plasmid which contained the required left end (LE) and right end (RE) DNA sequences for transposition around a cargo gene (e.g. a selection marker such as a Tet resistance gene). Using an *in vitro* transcription/translation (TXTL) system (e.g. *E. coli* lysate- or reticulocyte lysate-based system), the effector and integrase genes were expressed. After expression, the RNA, target DNA, and donor DNA were added and incubated to allow for transposition to occur. Transposition was detected via PCR across the junction of the integrase site, with one primer on the target DNA and one primer on the donor DNA. The resulting PCR product was sequenced via NGS to determine the exact insertion topology relative to the sgRNA/crRNA targeted site. The primers were located downstream such that a variety of insertion sites can be accommodated and detected. Primers were designed such that integration is detected in either orientation of cargo or on either side of the spacer, as the integration direction was also not known initially.

[00188] Integration efficiency was measured via quantitative PCR (qPCR) measurements of the experimental output of target DNA with integrated cargo, normalized to the amount of unmodified target DNA also measured via qPCR.

[00189] This assay may be conducted with purified protein components rather than from lysate-based expression. In this case the proteins were expressed in an *E. coli* protease deficient B strain under a T7 inducible promoter, the cells were lysed using sonication, and the His-tagged protein of interest was purified using HisTrap FF (GE Lifescience) Ni-NTA affinity chromatography on the AKTA Avant FPLC (GE Lifescience). Purity was determined using densitometry in ImageLab software (Bio-Rad) of the protein bands resolved on SDS-PAGE and InstantBlue Ultrafast (Sigma-Aldrich) coomassie stained acrylamide gels (Bio-Rad). The protein was desalted in storage buffer composed of 50 mM Tris-HCl, 300 mM NaCl, 1 mM TCEP, 5% glycerol; pH 7.5 (or other buffers as determined for maximum stability) and stored at -80 °C.

After purification the effector(s) and integrase(s) were added to the sgRNA, target DNA, and donor DNA as described above in a reaction buffer, for example 26 mM HEPES pH 7.5, 4.2 mM TRIS pH 8, 50 µg/mL BSA, 2 mM ATP, 2.1 mM DTT, 0.05 mM EDTA, 0.2 mM $MgCl_2$, 28 mM NaCl, 21 mM KCl, 1.35% glycerol,(final pH 7.5) supplemented with 15 mM $Mg(OAc)_2$.

Example 3 – Predicted RNA folding

[00190] Predicted RNA folding of the active single RNA sequence was computed at 37˚ using the method of Andronescu 2007. All hairpin-loop secondary structures were singly deleted from the structure and iteratively compiled into a smaller single guide. In a second approach, the tracrRNA of MG64-1 was aligned to known type V-k tracrRNA, and areas of unique insertions were mutated out of the single guide and minimized by 57 bases. **FIG. 12A** depicts the predicted structure of MG64-2 sgRNA (SEQ ID NO:202). **FIG. 12B** depicts the predicted structure of MG64-4 sgRNA (SEQ ID NO:203). **FIG. 12C** depicts the predicted structure of MG64-6 sgRNA (SEQ ID NO:201). **FIG. 12D** depicts the predicted structure of MG64-7 sgRNA (SEQ ID NO:204). **FIG. 12E** depicts the predicted structure of MG108-1 sgRNA (SEQ ID NO:206). The shading of the bases corresponds to the probability of base pairing of that base.

Example 4 – Transposon end verification via gel shift

[00191] The transposon ends were tested for TnsB binding via an electrophoretic mobility shift assay (EMSA). In this case the potential LE or RE was synthesized as a DNA fragment (100-500 bp) and end-labeled with FAM via PCR with FAM-labeled primers. The TnsB protein was synthesized in an in vitro transcription/translation system (e.g. PURExpress). After synthesis, 1 µL of TnsB protein was added to 50 nM of the labeled RE or LE in a 10 µL reaction in binding buffer (20 mM HEPES pH 7.5, 2.5 mM Tris pH 7.5, 10 mM NaCl, 0.0625 mM EDTA, 5 mM TCEP, 0.005% BSA, 1 ug/mL poly(dI-dC), and 5% glycerol). The binding was incubated at 30˚ for 40 minutes, then 2uL of 6X loading buffer (60 mM KCl, 10 mM Tris pH 7,6, 50% glycerol) was added. The binding reaction was separated on a 5% TBE gel and visualized. Shifts of the LE or RE in the presence of TnsB were attributed to successful binding and were indicative of transposase activity.

[00192] FIG. 15 shows an example of this experiment, where the RE DNA sequence for MG64-2 (e.g. SEQ ID NO:155) was end-labeled with FAM by the above procedure and incubated with the corresponding MG64-2 TnsB-like component (e.g. SEQ ID NO: 23). Upshift of the labeled band in Lane 3 indicates binding of the RE sequence by TnsB, indicating it contains an active RE transposition sequence.

Example 5 – Integrase activity in *E. coli* (prophetic)

[00193] As *E. coli* lacks the capacity to efficiently repair genomic double-stranded DNA breaks, transformation of *E. coli* by agents able to cause double-stranded breaks in the E. coli genome causes cell death. Exploiting this phenomenon, endonuclease or effector-assisted integrase activity is tested in *E. coli* by recombinantly expressing either the endonuclease or effector-assisted integrase and a guide RNA (determined e.g. as in Example 3) in a target strain with spacer/target and PAM sequences integrated into its genomic DNA.

[00194] Engineered strains are then transformed with a plasmid containing the nuclease or effector with single guide RNA, a plasmid expressing the integrase and accessory genes, and a plasmid containing a temperature sensitive origin of replication with a selectable marker flanked by left end (LE) and right end (RE) transposon motifs for integration. Transformants induced for expression of these genes are then screened for transfer of the marker to the genomic target by selection at restrictive temperature for plasmid replication and the marker integration in the genome is confirmed by PCR.

[00195] Off-target integrations are screened using an unbiased approach. In brief, purified gDNA is fragmented with Tn5 integrase or shearing, and DNA of interest is then PCR amplified using primers specific to a ligated adaptor and the selectable marker. The amplicons are then prepared for NGS sequencing. Analysis of the resulting sequences is trimmed of the transposon sequences and flanking sequences are mapped to the genome to determine insertion position, and off target insertion rates are determined.

Example 6 –Colony PCR screen of Transposase Activity (Prophetic)

[00196] For testing of nuclease or effector assisted integrase activity in bacterial cells, strain MGB0032 is constructed from BL21(DE3) E. coli cells which are engineered to contain the target and corresponding PAM sequence specific to MG64_1. MGB0032 E. coli cells are then transformed with pJL56 (plasmid that expresses the MG64_1 effector and helper suite, ampicillin resistant) and pTCM 64_1 sg, a chloramphenicol-resistant plasmid that expresses the single guide RNA sequence for the engineered target of interest driven by a T7 promoter.

[00197] An MGB0032 culture containing both plasmids is then grown to a saturation, diluted at least 1:10 into growth culture with appropriate antibiotics, and incubated at 37°C until OD of approximately 1. Cells from this growth stage are made electrocompetent and transformed with streamlined 64_1 pDonor, a plasmid bearing a tetracycline resistance marker flanked by left end (LE) and right end (RE) transposon motifs for integration. Electroporated cells are then

recovered for 2 hours on LB medium in the presence or absence of IPTG at a final concentration of 100 μM before being plated on LB-agar-ampicillin-chloramphenicol-tetracycline and incubated 4 days at 37°C. Sterile toothpicks are used to sample each resultant CFU, which is mixed into water. To this solution is added Q5 High Fidelity PCR mastermix (New England Biolabs) and primers LA155 (5'-GCTCTTCCGATCTNNNNNGATGAGCGCATTGTTAGATTTCAT-3') and oJL50 (5'-AAACCGACATCGCAGGCTTC-3'). These primers flank the predicted insertion junction. The predicted product size is 609 bp. DNA amplified PCR product is visualized on a 2% agarose gel. Sanger sequencing of PCR products confirms the transposition event.

Example 7 –In cell expression/in vitro assay (Prophetic)

[00198] To test the functionality of the NLS constructs in a physiologically relevant environment, constructs cloned with active NLS-tagged CAST components are integrated into K562 cells using lentiviral transduction. Briefly, constructs cloned into lentiviral transfer plasmids are transfected into 293T cells with envelope and packaging plasmids, and virus containing supernatant is harvested from the media after 72hr incubation. Media containing virus is then incubated with K562 cell lines with 8 μg/mL of polybrene for 72 hrs, and transfected cells are then selected for integration in bulk using Puromycin at 1 μg/mL for 4 days. Cell lines undergoing selection are harvested at the end of 4 days, and differentially lysed for nuclear and cytoplasmic fractions. Subsequent fractions are then tested for transposition capability with a complementary set of in vitro expressed components.

[00199] 10 million cells are centrifuged and washed once with 1xPBS pH7.4. Supernatant wash is aspirated completely to the cell pellet, and flash frozen at -80C for 16 hrs. After thawing on ice, cell pellet size is measured by mass, and appropriate extraction volumes of cell fractionation and nuclear extraction reagent (NE-PER) is used to natively extract proteins in cell fractions. Briefly, cytoplasmic extraction reagent is used at 1:10 mass of cells to volume of extraction reagent. Cell suspension is mixed by vortexing and lysed with non-ionic detergent. Cells are then centrifuged at 16,000xg at 4°C for 5 minutes. Cytoplasmic extraction supernatant is then decanted and saved for in vitro testing. Nuclear extraction reagent is then added 1:2 original cell mass to nuclear extraction reagent and incubated on ice for 1 hr on ice with intermittent vortexing. Nuclear suspension is then centrifuged at 16,000 x g for 10 minutes at 4°C and supernatant nuclear extract is decanted and tested for in vitro transposition activity. Using 4 μL of each cell and nuclear extract for each condition, we perform the in vitro transposition reaction

with a complementary set of in vitro expressed proteins, donor DNA, pTarget, and buffer. Evidence of transposition activity is assayed by PCR amplification of donor-target junctions.

Example 8 – Activity in mammalian cells (prophetic)

[00200] To show targeting and cleavage activity in mammalian cells, nuclear localization sequences are fused to the C terminus of each of the nuclease or effector proteins and integrase proteins and the fusion proteins are purified. A single guide RNA targeting a genomic locus of interest is synthesized and incubated with the nuclease/effector protein to form a ribonucleoprotein complex. Cells are transfected with a plasmid containing a selectable neomycin resistance marker (NeoR) or a fluorescent marker flanked by the left end (LE) and right end (RE) motifs, recovered for 4-6 hours, and subsequently electroporated with nuclease RNP and integrase proteins. Integration of a plasmid into the genome is quantified by counting G418-resistant colonies or fluorescence activated cell cytometry. Genomic DNA is extracted 72 hours after electroporation and used for the preparation of an NGS-library. Off target frequency is assayed by fragmenting the genome and preparing amplicons of the transposon marker and flanking DNA for NGS library preparation. At least 40 different target sites are chosen for testing each targeting system's activity.

Example 9 – Activity of targeted nuclease

[00201] In situ expression and protein sequence analyses suggest that some RNA guided effectors are active nucleases. They contain predicted endonuclease-associated domains (matching RuvC and HNH_endonuclease domains) and predicted HNH and RuvC catalytic residues (see e.g. **FIG. 4A**, which shows predicted catalytic residues of the MG36-5 effector).

[00202] Candidate activity is tested with engineered single guide RNA sequences using the myTXTL system and *in vitro* transcribed RNA. Active proteins are identified as those that successfully cleave the library to yield a band around 170 bp in agarose gel electrophoresis

Example 10 – Identification of transposons

[00203] Transposons are predicted to be active when they contain one or more protein sequences with integrase and/or integrase function between the left and right ends of the transposon. A typical Tn7 transposon generally comprises a catalytic integrase TnsB, but may also contain TnsA, TnsC, TnsD, TnsE, TniQ, and/or other integrases or integrases. The transposon ends comprise predicted integrase binding sites, which contain direct and/or inverted repeats of 15 bp to 150 bp in length flanking the integrase proteins and other 'cargo' genes.

Protein sequence analysis indicated that the integrases contain integrase domains, integrase domains and/or integrase catalytic residues, suggesting that they are active (e.g. **FIG. 4A**, which shows a locus diagram for an example MG36-5 effector-based CAST system containing TnsB elements; and **FIG. 5A**, which shows a locus diagram for an exemplary MG39-1 effector-based CAST system containing TnsA, TnsB, TnsC, and TniQ elements).

Example 11 – Identification of CRISPR-associated transposons

[00204] Putative CRISPR-associated transposons (CAST) contain a DNA and/or RNA targeting CRISPR effector and proteins with predicted integrase function in the vicinity of a CRISPR array. In some systems, the effector is predicted to have nuclease activity based on the presence of endonuclease-associated catalytic domains and/or catalytic residues (e.g. **FIG. 4A**, which shows predicted catalytic residues of the MG36-5 effector in the context of a CAST system locus containing TnsB elements). The integrases were predicted to be associated with the active nucleases when the CRISPR loci (CRISPR nuclease and array) and the integrase proteins are located between the predicted transposon left and right ends (e.g. **FIG. 4B** and **4C**). In this case, the effector was predicted to direct DNA integration to specific genomic locations based on a guide RNA.

[00205] In some systems, the effector was predicted to have homology with known CRISPR effector proteins, but to be inactive based on the absence of endonuclease domains and/or catalytic residues (**FIG. 5A**). The integrases are predicted to be associated with the effector when the CRISPR loci (inactive CRISPR nuclease and array) and the integrase proteins are located within the predicted transposon left and right ends (**FIG. 5A** and **5B**).

Example 12 – CAST Discovery

[00206] CRISPR-associated transposons (CAST) are systems that comprise a transposon that has evolved to interact with a CRISPR system to promote targeted integration of DNA cargo.

[00207] CASTs are genomic sequences encoding one or more protein sequences involved in DNA transposition within the signature left and right ends of the transposon. A typical Tn7 transposon, generally comprises a catalytic transposase TnsB, but may also contain a catalytic transposase TnsA, a loader protein TnsC or TniB, and target recognition proteins TnsD, TnsE, TniQ, and/or other transposon-associated components. The transposon ends comprise predicted transposase binding sites, which contain direct and/or inverted repeats of 15 bp to 150 bp in length flanking the transposon machinery and other ‘cargo’ genes.

[00208] In addition, CASTs also encode a DNA and/or RNA targeting CRISPR nuclease or effector in the vicinity of a CRISPR array. In some systems, the effector is predicted to be an active nuclease based on the presence of endonuclease-associated catalytic domains and/or catalytic residues. In some systems, the effector was predicted to have sequence similarity with known CRISPR effector proteins, but to be inactive based on the absence of endonuclease domains and/or catalytic residues. The transposons are predicted to be associated with the effector when the CRISPR locus and the transposon-associated proteins are located within the predicted transposon left and right ends. In this case, the effector is predicted to direct DNA integration to specific genomic locations based on a guide RNA.

Example 13a – Cas12k CAST

[00209] Cas12k CAST systems encode a nuclease-defective CRISPR Cas12k effector, a CRISPR array, a tracrRNA, and Tn7-like transposition proteins (see e.g. ,**FIG. 8**, which shows a locus organization diagram for MG108-1 CAST system containing Cas12k). Cas12k effectors are phylogenetically diverse and features that confirm their association with CASTs have been confirmed for several (see e.g., **FIG. 9**, which shows how MG64-1, MG64-2, MG64-3, MG 64-5, MG64-6, MG64-7, MG64-13, MG64-54, MG64-56, MG108-1, and MG108-2 effectors are part of this group). One such characteristic feature was transposon ends identified in the context of the MG64-3 CRISPR locus; the transposon left end was identified downstream from the MG64-3 CRISPR locus, as shown by terminal inverted repeats and self-matching spacer sequences (**FIG. 11A**). Another such characteristic that was identified includes Cas12k CAST CRISPR repeats (crRNA) which contain a conserved motif 5'-GNNGGNNTGAAAG-3' (see e.g. MG64-2, MG64-4, MG64-5, MG64-6, MG64-7, and MG108-1 and **FIG. 11B**). Short repeat-antirepeats (RAR) within the crRNA motif aligned with different regions of the tracrRNA, and RAR motifs appeared to define the start and end of the tracrRNA. **FIG. 13C** shows presence of these RAR motifs in e.g. MG64-2, MG64-4, MG64-5, MG64-6, MG64-7, and MG108-1 families.

Example 13b – Class I Type I-F CAST

[00210] Some CASTs encode nuclease-defective CRISPR Type I-F Cascade effector proteins, a CRISPR array, and Tn7-like transposition proteins (see e.g. **FIG. 10A**, which shows a locus organization diagram of a MG110-1 effector-based Type I-F CAST system). Type I-F Cascade CAST were predicted to function with a single guide RNA encoded by the crRNA, which contains a conserved motif 5'-CTGCCGNNTAGGNAGC-3' likely involved in formation of a

stem-loop structure (see e.g., **FIG. 10B-C**, which shows an alignment of this feature in MG110-1 and MG110-2 family crRNAs SEQ ID NOs: 207 and 208). Based in part on its having these same features, the MG110-2 effector-containing and family was also identified as a Type I-F CAST system.

Example 14 – Transposon end prediction

[00211] Transposon ends were estimated from intergenic regions flanking the effector and the transposon machinery. For example, for Cas12k CAST, the intergenic region located directly upstream from TnsB and directly downstream from the CRISPR locus, were predicted as containing the Tn7 transposon left and right ends (LE and RE) (see e.g. **FIG. 11A**, which shows LE and RE analysis in the context of an MG64-3 family CAST locus diagram).

[00212] Direct and inverted repeats (DR/IR) of ~12 bp were predicted on the contig, with up to 2 mismatches. In addition, the Dotplot algorithm was used to find short (~ 10-20 bp) DR/IR flanking CAST transposons. Matching DR/IR located in intergenic regions flanking CAST effector and transposon genes were predicted to encode transposon binding sites. LE and RE extracted from intergenic regions, which encode putative transposon binding sites, were aligned to define the transposon ends boundaries. Putative transposon LE and RE ends are identified as regions: a) located within 400 bp upstream and downstream from the first and last predicted transposon encoded genes; b) sharing multiple short inverted repeats; and c) sharing > 65% nucleotide id. This process was repeated to identify putative LE/RE sequences for MG36-5 (SEQ ID NOs: 17-18), MG39-1 (SEQ ID NOs:20-21), MG64-2 (SEQ ID NOs:125-126), MG64-4 (SEQ ID NOs:127-128), MG64-6 (SEQ ID NOs:123-124), MG64-7 (SEQ ID NOs:129-130), MG64-13 (SEQ ID NOs:131-132), MG64-54 (SEQ ID NO:133), MG108-1 (SEQ ID NOs:134-135), MG110-1 (SEQ ID NOs:136-137), and MG110-2 (SEQ ID NOs:138-139).

Example 15 – Single Guide Design for Class II, Type V CAST Systems

[00213] Analysis of the intergenic regions surrounding the Cas effector and CRISPR array for MG 64 sub-families identified a potential anti-repeat sequence and a conserved "CYCC(N6)GGRG" stem-loop structure neighboring the antirepeat corresponding to the sequence of the tracrRNA (**FIG. 11B**). TracrRNA and crRNA repeat were folded and trimmed, adding a tetraloop sequence of GAAA to maintain the stem loop region of the crRNA-tracrRNA complementary sequence, in order to generate the sgRNA. These sequences are outlined in Table 1 below.

Table 1: Corresponding crRNA-tracrRNA sequences for MG families described herein.

Description	SEQ ID NO:	Sequence
MG64-2 crRNA	255	See sequence listing
MG64-2 tracrRNA	262	AAUUAAUAGCGCCGCCGUUCAU GCUUCUAGGAGCCUCUGAAAGG UGACAAAUGCGGGUUAGUUUGG CUGUUGUCAGACAGUCUUGCUU UCUGACCCUGGUAGCUGCCCAC CCCGAAGCUGCUGUUCCUUGUG AACAGGAAUUAGGUGCGCCCCC AGUAAUAAGGGUAUGGGUUUAC CACAGUGGUGGCUACUGAAUCA CCUCCGAGCAAGGAGGAACCCA CU
MG64-4 crRNA	256	See sequence listing
MG64-4 tracrRNA	209	See sequence listing
MG64-6 crRNA	257	See sequence listing
MG64-6 tracrRNA	263	AUAACAGCGCCGCAGGUCAUGC CGUCAAAAGCCUCUGAACUGUG UUAAAUGGGGGUUAGUUUGACU GUUGAAAGACAGUUGUGCUUUC UGACCCUGGUAGCUGCCCACCC UGAUGCUGCUAUCUUUCGGGAU AGGAAUAAGGUGCGCUCCCAGU AAUAGGGGUGUAGAUGUACUAC AGUGGUGGCUACUAAAUCACCU CCGACCAAGGAGGAAUCCAUCC UUAAUUUUUUAUUUUUU
MG64-7 crRNA	258	See sequence listing
MG64-7 tracrRNA	210	See sequence listing
MG108-1 crRNA	261	See sequence listing
MG108-1 tracrRNA	235	See sequence listing
MG108-2 crRNA	260	See sequence listing
MG108-2 tracrRNA	236	See sequence listing

Example 16 – In vitro integration activity using targeted nuclease

[00214] In situ expression and protein sequence analyses indicated that some RNA guided effectors are active nucleases. They contain predicted endonuclease-associated domains (matching RuvC and HNH_endonuclease domains), and/or predicted HNH and RuvC catalytic residues. Candidate activity was tested with engineered single guide RNA sequences using the myTXTL system and in vitro transcribed RNA. Active proteins are identified as those that successfully cleave the library to yield a band around 170 bp in agarose gel electrophoresis.

Example 17 – Programmable DNA Integration

[00215] CAST activity was tested by combining five types of components in a single reaction: (1) a Cas effector protein expressed by myTXTL or PURExpress; (2) a target DNA fragment or plasmid containing the target sequence and PAM corresponding to the Cas enzyme; (3) a donor DNA fragment containing a marker or fragment of DNA flanked by the predicted LE and RE of the transposase system in a DNA fragment or plasmid; (4) any combination of additional transposase proteins predicted to be part of the array expressed using myTXTL or PURExpress; and (5) an engineered in vitro transcribed single guide RNA sequence. Active systems that successfully transposed the donor fragment were assayed by PCR amplification of the donor-target junction.

[00216] FIG. 13 shows example data demonstrating that the MG64-6 system comprising the MG64-6 effector, TnsB, TnsC, and TniQ proteins (SEQ ID NOs:30-33) using the predicted LE/RE donor sequences (SEQ ID NOs: 123-124) and in silico designed sgRNA (SEQ ID NO:201) is active. After performing the transposition reaction by combining all the MG64-6 components, PCR amplification of the junction showed that proper donor-target formation occurred and the transposition reaction was sg dependent. (**FIG. 13A**). Presence of amplified bands in PCR reactions #3 and #4 (spanning the LE/RE junctions when the LE/RE is inserted distal to the PAM, respectively) indicated that both orientations of the donor relative to the target are made: one where the LE is closer to the PAM, and one where the RE is closer to the PAM. While both transposition orientations were made, there was a preference for donor integration in the target where the LE is closer to the PAM, represented by strong band present for reactions #4 and #5 (which span the LE junction when it is inserted distal to the PAM and the RE junction when it is inserted proximal to the PAM, respectively).

[00217] Sanger sequencing of the preferred orientation product was performed. Of the integrations that occur with the LE closer to the PAM, there was a clear degradation of the

sequencing chromatogram signal from either the forward or reverse direction over the target/donor junction (**FIG. 13C**). This indicated that, of the products that are oriented with the LE closer to the PAM, integration occurred over a range of nucleotides, with the primary product of LE-closer-to-PAM products as a 61 bp integration from the PAM (**FIG. 14**). Sequencing that originated from the donor over the donor-target junction defined the composition of the essential outer bounds of the LE and RE sequences. Further investigation of the LE and RE domains will determine the inner limits of the LE and RE sequences and thus the minimal LE/RE that are essential for transposition. Sequencing of the RE on LE-closer-to-PAM products showed a 3 bp duplication downstream of the donor RE. This is in part due to the Tn7 transposase integration event that cleaves and ligates the donor fragment at a staggered cut site. A 3 bp duplication is smaller than the expected 5 bp of duplication from other Tn7 transposases.

[00218] Sanger sequencing of the PCR amplified product over the 8N library of the target plasmid also elucidated that the PAM preference of the MG64-6 effector as a nGTn/nGTt on the 5' end of the spacer. NGS analysis of the PAM library target corroborated the nGTn motif preference at the 5' end (**FIG. 13B**).

Example 18 – Integration window determination

[00219] PCR junctions of the PAM that were amplified in Example 17 above were indexed for NGS libraries and sequenced on a MiSeq with a V2 300 read kit. Reads were mapped and quantified using CRISPResso using an amplicon sequence of a putative transposition sequence with a 60bp distance of integration from the PAM (guideseq = 20bp 3' end of LE or RE, center of window = 0, window size = 20) Indel histogram was normalized to total indel reads detected, and frequencies were plotted relative to the 60bp reference sequence (**FIG. 14**).

[00220] Both PCR reactions 5 (LE proximal to PAM, **FIG. 13a**) and PCR 4 (RE distal to PAM, **FIG. 13b**) were plotted on the sequence and distance from the PAM for MG64-6 (**FIG. 14**). Analysis of the integration window indicated that 95% of the integrations that occurred at the spacer PAM site are within a 10 bp window between 58 and 68 nucleotides away from the PAM. Differences in the integration distance between the distal and the proximal frequencies reflected the integration site duplication - a 3-5 base pair duplication as a result of staggered nuclease activity of the transposase upon integration.

Example 19 – Transposon end verification via gel shift

[00221] In order to verify the activity of TnsB on the predicted transposon end sequence, the RE of MG64-6 was amplified using FAM labeled oligos. MG64-6 TnsB protein was expressed

using a cell free transcription/translation system and incubated with the RE FAM labeled product. After incubation for 30 minutes, binding was observed on a native 5% TBE gel (**FIG. 15**). Multiple bands of fluorescent product within the co-incubated lane (**FIG. 15**, lane 3) indicated a minimum of 3 TnsB binding sites.

Example 20 – Colony PCR screen of Transposase Activity (Prophetic)

[00222] Transposition activity is assayed via a colony PCR screen. After transformation with the pDonor plasmids, E. coli are plated onto LB- agar containing ampicillin, chloramphenicol, and tetracycline. Select CFUs are added to a solution containing PCR reagents and primers that flank the insertion junction.

Example 22 – LE-RE Minimization (Prophetic)

[00223] Sequencing of the target-transposition junction helps to identify the terminal inverted repeats by identifying the outmost sequence from the donor plasmid that are incorporated into the target reaction. By performing repeat analysis of 14 bp with variability of 10%, short repeats contained within the terminal ends are identified; identifying the minimal sequences to be included in truncations of these that preserve the repeats while deleting superfluous sequence. Prediction and cloning is done in multiple iterations, with each interaction tested with in vitro transposition. Transposition is predicted to be active down to a LE region of 68 bp combined with a RE region of 96bp.

Example 23 – Overhang influence of transposition (Prophetic)

[00224] In order to test whether superfluous sequence outside of the TnsB binding motifs are necessary for transposition, oligos designed for the TGTACA or TGTCGA motifs of both LE and RE are designed and synthesized with 0, 1, 2, 3, 5 and 10 bp extra base pairs. These synthesized oligos are used to generate donor PCR fragments with overhangs and tested for their ability to transpose into the target site.

Example 24 – CAST NLS design (Prophetic)

[00225] Eukaryotic genome editing for therapeutic purposes is dependent on the import of editing enzymes into the nucleus. Small polypeptide stretches of larger proteins signal to cellular components for protein import across the nuclear membrane. Placement of these tags may require optimization, as these NLS tags need to provide import function while also maintaining function of the protein to which it is fused. In order to test functional orientations of the NLS to

each of the components of the CAST complex, constructs fusing Nucleoplasmin NLS to the N-terminus and SV40 NLS to the C-terminus of each of the components of the MG CAST are synthesized. Proteins of these constructs are expressed in cell free in vitro transcription/ translation reactions and tested for in vitro transposition activity with a complement set of untagged components. NLS-tagged constructs are assessed for maintenance of activity by PCR of the donor-target junction using PCR 4 (Assessing RE distal transpositions) and the cognate transposition event, PCR 5(Assessing LE proximal transposition).

Example 25 – Cas12k and TniQ protein fusion construct design and testing (Prophetic)

[00226] To simplify/minimize the expression of the protein components and facilitate delivery of these components into cells, fusion constructs between the Cas12k effector and the TniQ protein with various linkers, linker lengths, and domain boundaries are designed, synthesized, and tested. Both orientations of the TniQ fused to the Cas12k are designed and synthesized, a C-terminal fusion, Cas-TniQ, and an N terminal fusion, TniQ-Cas.

[00227] Two other linkers are also employed to fuse the effector and TniQ genes. P2A, a self-stopping translation sequence is active in a Cas-NLS-P2A-NLS-TniQ construct, and an MCV Internal Ribosome Entry Sequence (IRES) mRNA-based linker allows for independent translation of the two components in cells.

Example 26 – Intracellular expression coupled with in vitro transposition testing (Prophetic)

[00228] To test the functionality of the NLS constructs in a physiologically relevant environment, constructs cloned with active NLS-tagged CAST components are integrated into K562 cells using lentiviral transduction. Briefly, constructs cloned into lentiviral transfer plasmids are transfected into 293T cells with envelope and packaging plasmids, and virus containing supernatant are harvested from the media after 72hr incubation. Media containing virus is then incubated with K562 cell lines with 8 µg/mL of polybrene for 72 hrs, and transfected cells are then selected for integration in bulk using Puromycin at 1 µg/mL for 4 days. Cell lines undergoing selection are harvested at the end of 4 days, and differentially lysed for nuclear and cytoplasmic fractions. Subsequent fractions are then tested for transposition capability with a complementary set of in vitro expressed components.

[00229] Both NLS-TnsB and TnsB-NLS are tested by cell fractionation and in vitro transposition, and transposition is detected across both cytoplasmic and nuclear fractions

[00230] Cas12k fusions in the cell are similarly fractionated and tested for transposition. Cas-NLS Cas-NLS-P2A-NLS-TniQ are transduced into cells, fractionated, and tested in vitro for subcellular activity. Cas-NLS-P2A-NLS-TniQ is able to transpose in the cytoplasm with the addition of single guide to the reaction. By supplementing holo Cas protein (+sgRNA) or additional TniQ with sgRNA, we are able to complement the Cas-NLS-P2A-NLS-TniQ construct in the nuclear fraction.

[00231] Systems of the present disclosure may be used for various applications, such as, for example, nucleic acid editing (e.g., gene editing) or binding to a nucleic acid molecule (e.g., sequence-specific binding). Such systems may be used, for example, for remediating (e.g., removing or replacing) a genetically inherited mutation that may cause a disease in a subject; inactivating a gene in order to ascertain its function in a cell; as a diagnostic tool to detect disease-causing genetic elements (e.g. via cleavage of reverse-transcribed viral RNA or an amplified DNA sequence encoding a disease-causing mutation); as deactivated enzymes in combination with a probe to target and detect a specific nucleotide sequence (e.g. sequence encoding antibiotic resistance int bacteria); to render viruses inactive or incapable of infecting host cells by targeting viral genomes; to add genes or amend metabolic pathways to engineer organisms to produce valuable small molecules, macromolecules, or secondary metabolites; to establish a gene drive element for evolutionary selection, and/or to detect cell perturbations by foreign small molecules and nucleotides as a biosensor.

[00232] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. It is not intended that the invention be limited by the specific examples provided within the specification. While the invention has been described with reference to the aforementioned specification, the descriptions and illustrations of the embodiments herein are not meant to be construed in a limiting sense. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. Furthermore, it shall be understood that all aspects of the invention are not limited to the specific depictions, configurations or relative proportions set forth herein which depend upon a variety of conditions and variables. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is therefore contemplated that the invention shall also cover any such alternatives, modifications, variations, or equivalents. It is intended that the following claims define the scope of the invention and that

methods and structures within the scope of these claims and their equivalents be covered thereby.

CLAIMS

WHAT IS CLAIMED IS:

1. A system for transposing a cargo nucleotide sequence to a target nucleic acid site comprising:

a first double-stranded nucleic acid comprising said cargo nucleotide sequence, wherein said cargo nucleotide sequence is configured to interact with a recombinase or transposase complex;

a Cas effector complex comprising a class II, type II Cas effector and at least one engineered guide polynucleotide configured to hybridize to said target nucleic acid site; and

said recombinase or transposase complex, wherein said recombinase or transposase complex is configured to recruit said cargo nucleotide sequence to said target nucleic acid site.

2. The system of claim 1, wherein said recombinase or transposase complex binds non-covalently to said Cas effector complex.
3. The system of claim 1, wherein said recombinase or transposase complex is covalently linked to said Cas effector complex.
4. The system of claim 3, wherein said recombinase or transposase complex is fused to said Cas effector complex in a single polypeptide.
5. The system of any one of claims 1-4, wherein said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence.
6. The system of any one of claims 1-5, further comprising a second double-stranded nucleic acid comprising said target nucleic acid site.
7. The system of claim 6, further comprising a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site.
8. The system of claim 7, wherein said PAM sequence is located 3' of said target nucleic acid site.
9. The system of any one of claims 1-8, wherein said recombinase or transposase complex is a Tn7 type transposase complex.
10. The system of any one of claims 1-9, wherein said engineered guide polynucleotide is configured to bind said class II, type II Cas effector.

11. The system of any one of claims 1-10, wherein said class II, type II Cas effector comprises a polypeptide comprising a sequence having at least 80% identity to SEQ ID NO: 1 or a variant thereof.
12. The system of any one of claims 1-11, wherein said recombinase or transposase complex comprises at least one, at least two, at least three, or four polypeptide(s) comprising a sequence having at least 80% identity to any one of SEQ ID NOs: 2-5 or a variant thereof.
13. The system of any one of claims 1-12, wherein said engineered guide polynucleotide comprises a sequence comprising at least 60-80 consecutive nucleotides having at least 80% identity to SEQ ID NO:12 or a variant thereof.
14. The system of any one of claims 1-12, wherein said engineered guide polynucleotide comprises a sequence having at least 80% identity to SEQ ID NO: 11 or a variant thereof
15. The system of any one of claims 2-14, wherein said left-hand recombinase sequence comprises a sequence having at least 80% identity to any one of SEQ ID NOs: 17-18 or a variant thereof.
16. The system of any one of claims 2-15, wherein said right-had recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 19 or a variant thereof.
17. The system of any one of claims 1-16, wherein said class II, type II Cas effector and said recombinase or transposase complex are encoded by polynucleotide sequences comprising fewer than about 10 kilobases.
18. A method for transposing a cargo nucleotide sequence to a target nucleic acid site comprising a target nucleotide sequence comprising expressing the system of any one of claims 1-17 within a cell or introducing the system of any one of claims 1-17 to a cell.
19. A system for transposing a cargo nucleotide sequence to a target nucleic acid site comprising:

a first double-stranded nucleic acid comprising a cargo nucleotide sequence configured to interact with a Tn7 type transposase complex;

a Cas effector complex comprising a class II, type V Cas effector and an engineered guide polynucleotide configured to hybridize to said target nucleotide sequence; and

a Tn7 type transposase complex configured to bind said Cas effector complex, wherein said Tn7 type transposase complex comprises a TnsA subunit.

20. The system of claim 19, wherein said transposase complex binds non-covalently to said Cas effector complex.
21. The system of claim 19, wherein said transposase complex is covalently linked to said Cas effector complex.
22. The system of claim 21, wherein said transposase complex is fused to said Cas effector complex in a single polypeptide.
23. The system of any one of claims 19-22, wherein said class II, type V Cas effector is not a Cas12k effector.
24. The system of any one of claims 19-23, wherein said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence.
25. The system of any one of claims 19-24, further comprising a second double-stranded nucleic acid comprising said target nucleic acid site.
26. The system of claim 19-25, further comprising a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site.
27. The system of claim 26, wherein said PAM sequence is located 5' of said target nucleic acid site.
28. The system of any one of claims 19-27, wherein said engineered guide polynucleotide is configured to bind said class II, type V Cas effector.
29. The system of any one of claims 19-28, wherein said TnsA subunit comprises a polypeptide having a sequence having at least 80% identity to SEQ ID NO: 7 or a variant thereof.
30. The system of any one of claims 19-29, wherein said Tn7 type transposase complex comprises at least one, at least two, or three polypeptide(s) comprising a sequence having at least 80% identity to any one of SEQ ID NOs: 8-10, or a variant thereof.
31. The system of any one of claims 19-30, wherein said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 13-16 , or a variant thereof.
32. The system of any one of claims 24-31, wherein said left-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 20 , or a variant thereof.
33. The system of any one of claims 24-32, wherein said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 21, or a variant thereof.

34. The system of any one of claims 19-33, wherein said class II, type V Cas effector is not a Cas12k effector.
35. The system of any one of claims 19-34, wherein said class II, type V Cas effector and said Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 10 kilobases.
36. A method for transposing a cargo nucleotide sequence to a target nucleic acid site comprising a target nucleotide sequence comprising expressing the system of any one of claims 19-34 within a cell or introducing the system of any one of claims 19-34 to a cell.
37. A method for transposing a cargo nucleotide sequence to a target nucleic acid site, comprising contacting a first double-stranded nucleic acid comprising a cargo nucleotide sequence with:

    a Cas effector complex comprising a class II, type II Cas effector and at least one engineered guide polynucleotide configured to hybridize to said target nucleic acid site;

    a recombinase or transposase complex configured to recruit said cargo nucleotide to said target nucleic acid site; and

    a second double-stranded nucleic acid comprising said target nucleic acid site.
38. The system of claim 37, wherein said recombinase or transposase complex binds non-covalently to said Cas effector complex.
39. The system of claim 37, wherein said recombinase or transposase complex is covalently linked to said Cas effector complex.
40. The system of claim 39, wherein said recombinase or transposase complex is fused to said Cas effector complex in a single polypeptide.
41. The method of any one of claims 37-40, wherein said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence.
42. The method of any one of claims 37-41, further comprising a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site.
43. The method of claim 42, wherein said PAM sequence is located 3’ of said target nucleic acid site.
44. The system of any one of claims 37-43, wherein said recombinase or transposase complex is a Tn7 type transposase complex.
45. The method of any one of claims 37-44, wherein said engineered guide polynucleotide is configured to bind said class II, type II Cas effector.

46. The method of any one of claims 37-45, wherein said class II, type II Cas effector comprises a polypeptide comprising a sequence having at least 80% identity to SEQ ID NO: 1 or a variant thereof.
47. The method of any one of claims 37-46, wherein said recombinase or transposase complex comprises at least one, at least two, at least three, or four polypeptide(s) comprising a sequence having at least 80% identity to any one of SEQ ID NOs: 2-5 or a variant thereof.
48. The method of any one of claims 37-47, wherein said engineered guide polynucleotide comprises a sequence comprising at least 60-80 consecutive nucleotides having at least 80% identity to SEQ ID NO: 12 or a variant thereof.
49. The method of any one of claims 37-48, wherein said engineered guide polynucleotide comprises a sequence having at least 80% identity to SEQ ID NO: 11 or a variant thereof.
50. The method of any one of claims 41-49, where said left-hand recombinase sequence comprises a sequence having at least 80% identity to any one of SEQ ID NOs: 17-18 or a variant thereof.
51. The method of any one of claims 41-50, wherein said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 19 or a variant thereof.
52. The method of any one of claims 37-51, wherein said class II, type II Cas effector and said Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 10 kilobases.
53. A method for transposing a cargo nucleotide sequence to a target nucleic acid site, comprising contacting a first double-stranded nucleic acid comprising said cargo nucleotide sequence with:

a Cas effector complex comprising a class II, type V Cas effector and at least one engineered guide polynucleotide configured to hybridize to said target nucleotide sequence;

a Tn7 type transposase complex configured to bind said Cas effector complex, wherein said Tn7 type transposase complex comprises a TnsA subunit; and

a second double-stranded nucleic acid comprising said target nucleic acid site.

54. The system of claim 53, wherein said transposase complex binds non-covalently to said Cas effector complex.

55. The system of claim 53, wherein said transposase complex is covalently linked to said Cas effector complex.
56. The system of claim 55, wherein said transposase complex is fused to said Cas effector complex in a single polypeptide.
57. The method of any one of claims 53-56, wherein said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence.
58. The method of any one of claims 53-57, further comprising a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site.
59. The method of claim 58, wherein said PAM sequence is located 3' of said target nucleic acid site.
60. The method of any one of claims 53-59, wherein said engineered guide polynucleotide is configured to bind said class II, type V Cas effector.
61. The method of any one of claims 53-60, wherein said TnsA subunit comprises a polypeptide having a sequence having at least 80% identity to SEQ ID NO: 7or a variant thereof.
62. The method of any one of claims 53-61, wherein said Tn7 type transposase complex comprises at least one, at least two, or three polypeptide(s) comprising a sequence having at least 80% identity to any one of SEQ ID NOs: 8-10, or a variant thereof.
63. The method of any one of claims 53-62, wherein said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 13-16 or a variant thereof.
64. The method of any one of claims 57-63, wherein said left-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 20, or a variant thereof.
65. The method of any one of claims 57-64, wherein said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 21, or a variant thereof.
66. The method of any one of claims 53-65, wherein said class II, type V Cas effector is not a Cas12k effector
67. The method of any one of claims 53-66, wherein said class II, type V Cas effector and said Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 10 kilobases.

68. A system for transposing a cargo nucleotide sequence to a target nucleic acid site comprising:

a first double-stranded nucleic acid comprising a cargo nucleotide sequence configured to interact with a Tn7 type transposase complex;

a Cas effector complex comprising a class I, type I-F Cas effector and an engineered guide polynucleotide configured to hybridize to said target nucleotide sequence; and

a Tn7 type transposase complex configured to bind said Cas effector complex, wherein said Tn7 type transposase complex comprises a TnsA subunit.

69. The system of claim 68, wherein said transposase complex binds non-covalently to said Cas effector complex.

70. The system of claim 68, wherein said transposase complex is covalently linked to said Cas effector complex.

71. The system of claim 70, wherein said transposase complex is fused to said Cas effector complex in a single polypeptide.

72. The system of any one of claims 68-71, wherein said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence.

73. The system of any one of claims 68-72, further comprising a second double-stranded nucleic acid comprising said target nucleic acid site.

74. The system of claim 68-73, further comprising a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site.

75. The system of claim 74, wherein said PAM sequence is located 3' of said target nucleic acid site.

76. The system of claim 74, wherein said PAM sequence is located 5' of said target nucleic acid site.

77. The system of any one of claims 68-76, wherein said engineered guide polynucleotide is configured to bind said class I, type I-F Cas effector.

78. The system of any one of claims 68-77, wherein said class I, type I-F Cas effector comprises a polypeptide comprising a sequence having at least 80% identity to any one of SEQ ID NO: 41-43, or 48-50, or a variant thereof.

79. The system of any one of claims 68-78, wherein said Tn7 type transposase complex comprises at least one, at least two, or three polypeptide(s) comprising a sequence having at least 80% identity to any one of SEQ ID NOs: 44-46, or 51-53, or a variant thereof.

80. A method for transposing a cargo nucleotide sequence to a target nucleic acid site comprising a target nucleotide sequence comprising expressing the system of any one of claims 68-79 within a cell or introducing the system of any one of claims 68-79 to a cell.

81. A system for transposing a cargo nucleotide sequence to a target nucleic acid site comprising:

a first double-stranded nucleic acid comprising a cargo nucleotide sequence configured to interact with a Tn7 type transposase complex;

a Cas effector complex comprising a class II, type V Cas effector and an engineered guide polynucleotide configured to hybridize to said target nucleotide sequence; and

a Tn7 type transposase complex configured to bind said Cas effector complex,

wherein said Tn7 type transposase complex comprises TnsB, TnsC, and TniQ components, wherein:

(a) said class II, type V Cas effector comprises a polypeptide having a sequence having at least 80% sequence identity to any one of SEQ ID NOs:22, 26, 30, 34, 55-89, 104, or 147, or a variant thereof; or

(b) said Tn7 type transposase complex comprises a TnsB, TnsC, or TniQ component having a sequence having at least 80% sequence identity to any one of SEQ ID NOs:23-25, 27-29, 31-33, 35-37, 101-103, 105-107, or 148-150, or a variant thereof.

82. The system of claim 81, wherein said transposase complex binds non-covalently to said Cas effector complex.

83. The system of claim 81, wherein said transposase complex is covalently linked to said Cas effector complex.

84. The system of claim 83, wherein said transposase complex is fused to said Cas effector complex in a single polypeptide.

85. The system of any one of claims 81-84, wherein said class II, type V Cas effector comprises a polypeptide comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs:22, 26, 30, 34, 55-89, 104, or 147, or a variant thereof.

86. The system of any one of claims 81-85, wherein said Tn7 type transposase complex comprises a TnsB, TnsC, or TniQ component comprising a sequence having at least 80% sequence identity to any one of SEQ ID NOs:23-25, 27-29, 31-33, 35-37, 101-103, 105-107, or 148-150, or a variant thereof.

87. The system of any one of claims 81-86, wherein said class II, type V Cas effector is a Cas12k effector.
88. The system of any one of claims 81-87, wherein said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence.
89. The system of any one of claims 81-88, further comprising a second double-stranded nucleic acid comprising said target nucleic acid site.
90. The system of any one of claims 81-89, further comprising a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site.
91. The system of claim 90, wherein said PAM sequence is located 5' of said target nucleic acid site.
92. The system of claim 91, wherein said PAM sequence comprises 5'-nGTn-3' or 5'-nGTt-3'.
93. The system of any one of claims 81-92, wherein said engineered guide polynucleotide is configured to bind said class II, type V Cas effector.
94. The system of any one of claims 81-93, wherein said TnsB, TnsC, and TniQ components comprise polypeptides having a sequence having at least 80% identity to any one of SEQ ID NOs:23-25, 27-29, 31-33, 35-37, 101-103, 105-107, or 148-150, respectively.
95. The system of any one of claims 81-94, wherein said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 90, 91, 92, 93, 117, 151, 156-181, or 209-234.
96. The system of any one of claims 81-95, wherein said engineered guide polynucleotide comprises a sequence having at least 80% sequence identity to any one of SEQ ID NOs: 111-114 or 201- 206, 255, 262, 256, 209, 257, 263, 258, 210, or a variant thereof.
97. The system of any one of claims 88-96, wherein said left-hand recombinase sequence comprises a sequence having at least 80% identity to any one of SEQ ID NOs: 125, 127, 123, 129, 131, 133, 153, or 134, or a variant thereof.
98. The system of any one of claims 88-97, wherein said right-hand recombinase sequence comprises a sequence having at least 80% identity to any one of SEQ ID NOs: 126, 155, 128, 124, 130, 132, or 154 , or a variant thereof.
99. The system of any one of claims 81-97, wherein said class II, type V Cas effector and said Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 10 kilobases.

100. The system of claim 81 and 88, wherein:

(a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:22 or a variant thereof;

(b) said left-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:125 or a variant thereof;

(c) said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 126 or 155, or a variant thereof;

(d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about 46-60 nucleotides of SEQ ID NO: 90; or (ii) comprises a sequence having at least 80% sequence identity to non-degenerate nucleotides of any one of SEQ ID NOs: 94, 112, or 202; or

(e) said TnsB, TnsC, and TniQ components comprise sequences having at least 80% sequence identity to SEQ ID NOs: 23-25 or variants thereof.

101. The system of claim 81 and 88, wherein:

(a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:26 or a variant thereof;

(b) said left-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:127 or a variant thereof;

(c) said right-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:128 or a variant thereof;

(d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about 46-60 nucleotides of any one of SEQ ID NOs: 91, 156, or 209; or (ii) comprises a sequence having at least 80% sequence identity to non-degenerate nucleotides of any one of SEQ ID NOs: 95, 113, or 203, or

(e) said TnsB, TnsC, and TniQ components comprise sequences having at least 80% sequence identity to SEQ ID NOs: 27-29 or variants thereof.

102. The system of claim 81 and 88, wherein:

(a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:60 or a variant thereof;

(b) said left-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:131 or a variant thereof;

(c) said right-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:132 or a variant thereof;

(d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about 46-60 nucleotides of any one of SEQ ID NOs: 117, 161, or 214; or (ii) comprises a sequence having at least 80% sequence identity to non-degenerate nucleotides of SEQ ID NO: 119; or

(e) said TnsB, TnsC, and TniQ components comprise sequences having at least 80% sequence identity to SEQ ID NOs: 101-103 or variants thereof.

103. The system of claim 81 and 88, wherein:

(a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:147 or a variant thereof;

(b) said left-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:153 or a variant thereof;

(c) said right-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:154 or a variant thereof;

(d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about 46-60 nucleotides of any one of SEQ ID NOs: 151, 181, or 234; or (ii) comprises a sequence having at least 80% sequence identity to non-degenerate nucleotides of SEQ ID NO: 152 or 254; or

(e) said TnsB, TnsC, and TniQ components comprise sequences having at least 80% sequence identity to SEQ ID NOs:148-150 or variants thereof.

104. The system of claim 81 and 88, wherein:

(a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:34 or a variant thereof;

(b) said left-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:129 or a variant thereof;

(c) said right-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:130 or a variant thereof;

(d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about 46-60 nucleotides of any one of SEQ ID NOs: 93, 157, or 210; or (ii) comprises a sequence having at least 80% sequence identity to non-degenerate nucleotides of any one of SEQ ID NOs: 97, 114, or 204.

(e) said TnsB, TnsC, and TniQ components comprise sequences having at least 80% sequence identity to SEQ ID NOs:148-150 or variants thereof

105. The system of claim 81, 88, and 91, wherein:

(a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:30 or a variant thereof;

(b) said left-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:123 or a variant thereof;

(c) said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 124, or a variant thereof;

(d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about 46-80 nucleotides of SEQ ID NO:92; or (ii) comprises a sequence having at least 80% identity to the non-degenerate nucleotides of SEQ ID NO:111 or 201;

(e) said TnsB, TnsC, and TniQ components comprise polypeptides having a sequence having at least 80% identity to SEQ ID NOs: 31, 32, and 33, or variants thereof; or

(f) said PAM sequence comprises 5'-nGTn-3' or 5'-nGTt-3'.

106. A system for transposing a cargo nucleotide sequence to a target nucleic acid site comprising:

a first double-stranded nucleic acid comprising a cargo nucleotide sequence configured to interact with a Tn7 type transposase complex;

a Cas effector complex comprising a class II, type V Cas effector and an engineered guide polynucleotide configured to hybridize to said target nucleotide sequence; and

a Tn7 type transposase complex configured to bind said Cas effector complex, wherein said Tn7 type transposase complex comprises TnsB and TnsC components but does not comprise a TnsA and/or TniQ component.

107. The system of claim 106, wherein said transposase complex binds non-covalently to said Cas effector complex.

108. The system of claim 106, wherein said transposase complex is covalently linked to said Cas effector complex.

109. The system of claim 108, wherein said transposase complex is fused to said Cas effector complex in a single polypeptide.

110. The system of any one of claims 106-109, wherein said Tn7 type transposase complex comprises a polypeptide having a sequencing having at least 80% sequence identity to any one of SEQ ID NOs: 39-40 or 109-110.

111. The system of any one of claims 106-110, wherein said TnsB component comprises a polypeptide comprising a sequence having at least 80% sequence identity to SEQ ID NOs: 40 or 109.

112. The system of any one of claims 106-111, wherein said TnsC component comprises a polypeptide comprising a sequence having at least 80% sequence identity to SEQ ID NOs: 39 or 110.

113. The system of any one of claims 106-112, wherein said class II, type V Cas effector is a Cas12k effector.

114. The system of any one of claims 106-112, wherein said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO: 38 or SEQ ID NO:108.

115. The system of any one of claims 106-114, wherein said cargo nucleotide sequence is flanked by a left-hand transposase recognition sequence and a right-hand transposase recognition sequence.

116. The system of any one of claims 106-115, further comprising a second double-stranded nucleic acid comprising said target nucleic acid site.

117. The system of any one of claims 106-116, wherein said double-stranded nucleic acid comprising said target nucleic acid site or said system is inside a cell.

118. The system of any one of claims 106-117, further comprising a PAM sequence compatible with said Cas effector complex adjacent to said target nucleic acid site.

119. The system of claim 118, wherein said PAM sequence is located 5' of said target nucleic acid site.

120. The system of any one of claims 106-119, wherein said engineered guide polynucleotide is configured to bind said class II, type V Cas effector.

121. The system of any one of claims 106-120, wherein said TnsB and TnsC components comprise polypeptides having a sequence having at least 80% identity to SEQ ID NOs: 40 and 39 or 109 and 110, respectively.

122. The system of any one of claims 106-121, wherein said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 118, 182, 183, 235, or 236, or a variant thereof.

123. The system of any one of claims 106-121, wherein said engineered guide polynucleotide comprises a sequence having at least 80% identity to non-degenerate

nucleotides any one of SEQ ID NOs: 115, 116, 205, 206, 261, 235, 260, or 236, or a variant thereof.

124. The system of any one of claims 115-123, wherein said left-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO:134.

125. The system of any one of claims 115-124, wherein said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NOs: 135, or a variant thereof.

126. The system of any one of claims 106-125, wherein said class II, type V Cas effector and said Tn7 type transposase complex are encoded by polynucleotide sequences comprising fewer than about 10 kilobases.

127. The system of claim 106 and 115, wherein:

(a) said class II, type V Cas effector comprises a sequence having at least 80% sequence identity to SEQ ID NO:38 or a variant thereof;

(b) said left-hand recombinase sequence comprises a sequence having at least 80% sequence identity to SEQ ID NO:134 or a variant thereof;

(c) said right-hand recombinase sequence comprises a sequence having at least 80% identity to SEQ ID NO: 135, or a variant thereof;

(d) said engineered guide polynucleotide: (i) comprises a sequence having at least 80% sequence identity to at least about 46-80 nucleotides of SEQ ID NO:182 or 235; or (ii) comprises a sequence having at least 80% identity to the non-degenerate nucleotides of SEQ ID NO:98, 115, 116, 205, or 206; or

(e) said TnsB and TnsC components comprise polypeptides having a sequence having at least 80% identity to SEQ ID NOs: 40 and 39, or variants thereof.

128. A method for transposing a cargo nucleotide sequence to a target nucleic acid site comprising a target nucleotide sequence comprising expressing the system of any one of claims 81-127 within a cell or introducing the system of any one of claims 81-127 to a cell

129. An engineered nuclease system comprising:

an endonuclease comprising a RuvC domain and an HNH domain, wherein said endonuclease is derived from an uncultivated microorganism, wherein said endonuclease is a Class II, type II endonuclease comprising a sequence having at least 80% identity to SEQ ID NO: 1 or a variant thereof; and

an engineered guide polynucleotide, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence.

130. The engineered nuclease system of claim 129, wherein said engineered guide polynucleotide comprises at least 60-80 consecutive nucleotides having at least 80% identity to SEQ ID NO:12 or a variant thereof.

131. The engineered nuclease system of claim 129 and 130, wherein said engineered guide polynucleotide comprises a sequence having at least 80% identity to SEQ ID NO: 11 or a variant thereof.

132. An engineered nuclease system comprising:

an endonuclease comprising a RuvC domain, wherein said endonuclease is derived from an uncultivated microorganism, and wherein said endonuclease is a Class II, type V endonuclease having at least 80% identity to SEQ ID NO: 5; and

an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence.

133. The engineered nuclease system of claim 132, wherein said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to SEQ ID NOs: 13-16 , or a variant thereof.

134. An engineered nuclease system comprising:

an endonuclease comprising a RuvC domain, wherein said endonuclease is derived from an uncultivated microorganism, and wherein said endonuclease is a Class II, type V-K endonuclease having at least 80% identity to any one of SEQ ID NOs:22, 26, 30, 34, 55-89, 104, or 147, or a variant thereof; and

an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence.

135. The engineered nuclease system of claim 134, wherein said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 90, 91, 92, 93, 117, 151, 156-181, or 209-234, or a variant thereof.

136. The engineered nuclease system of claim 134 or 135, wherein said engineered guide polynucleotide comprises a sequence having at least 80% sequence identity to non-

degenerate nucleotides of any one of SEQ ID NOs: 111-114 or 201- 206, 255, 262, 256, 209, 257, 263, 258, 210, or a variant thereof.

137. An engineered nuclease system comprising:

an endonuclease comprising a RuvC domain, wherein said endonuclease is derived from an uncultivated microorganism, and wherein said endonuclease is a Class II, type V-K endonuclease having at least 80% identity to any one of SEQ ID NO: 38 or SEQ ID NO:108, or a variant thereof; and

an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence.

138. The engineered nuclease system of claim 137, wherein said engineered guide polynucleotide comprises a sequence comprising at least about 46-80 consecutive nucleotides having at least 80% identity to any one of SEQ ID NOs: 118, 182, 183, 235, or 236, or a variant thereof.

139. The engineered nuclease system of claim 137, wherein said engineered guide polynucleotide comprises a sequence having at least 80% identity to non-degenerate nucleotides of any one of SEQ ID NOs: 111-114 or 201-206, 255, 262, 256, 209, 257, 263, 258, 210, 115, 116, 205, 206, 261, 235, 260, or 236 , or a variant thereof.

140. An engineered nuclease system comprising:

a Class I, type I-F Cas endonuclease comprising at least one Cas6, Cas7, or Cas8 polypeptide comprising a sequence having at least 80% identity to any one of SEQ ID NO: 41-43, or 48-50, or a variant thereof; and

an engineered guide RNA, wherein said engineered guide RNA is configured to form a complex with said endonuclease and said engineered guide RNA comprises a spacer sequence configured to hybridize to a target nucleic acid sequence.

141. The engineered nuclease system of claim 140, wherein said engineered guide polynucleotide comprises a sequence having at least 80% identity to non-degenerate nucleotides of any one of SEQ ID NOs: 121, 122, 207, or 208.

.

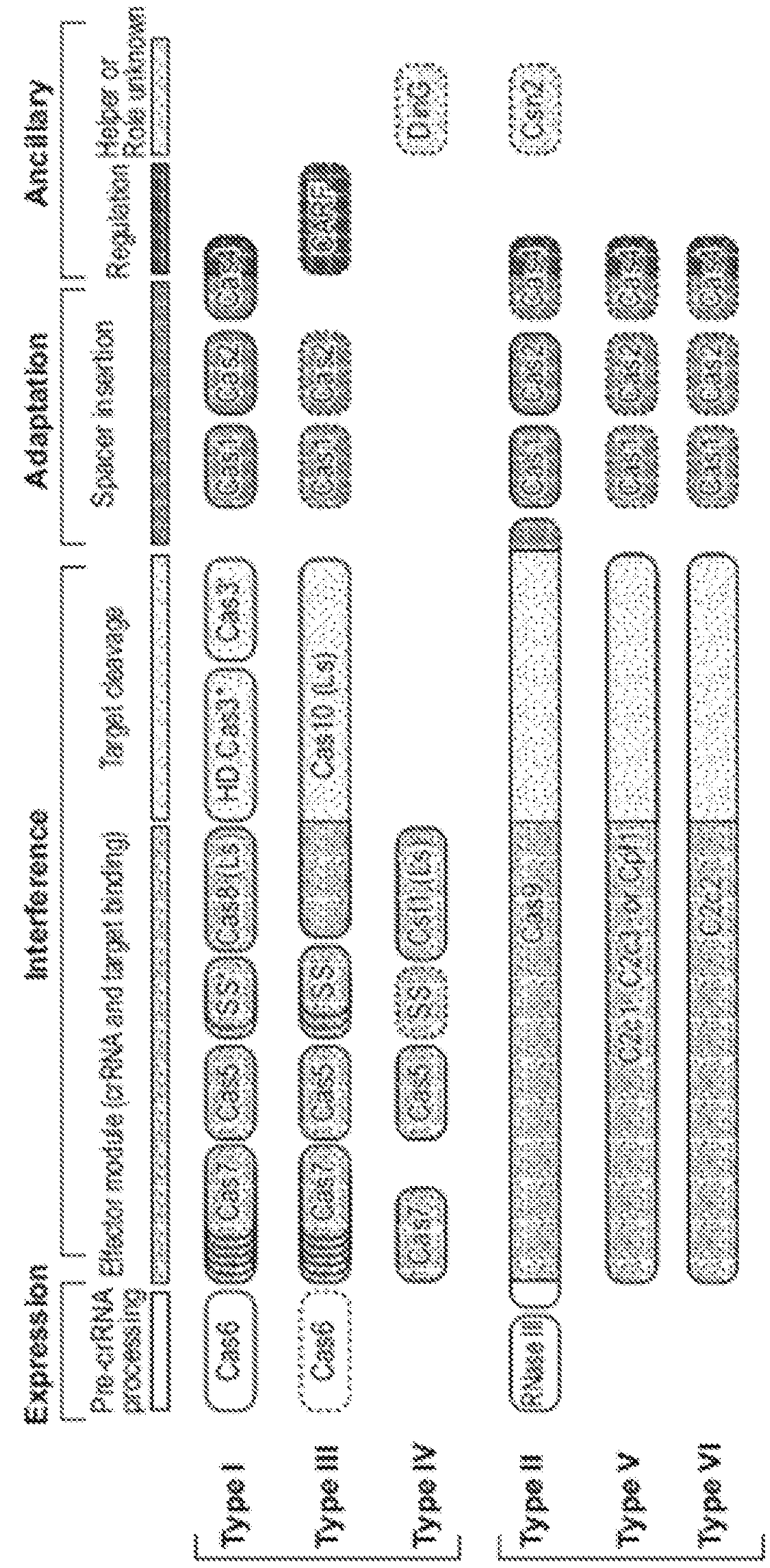


FIG. 1

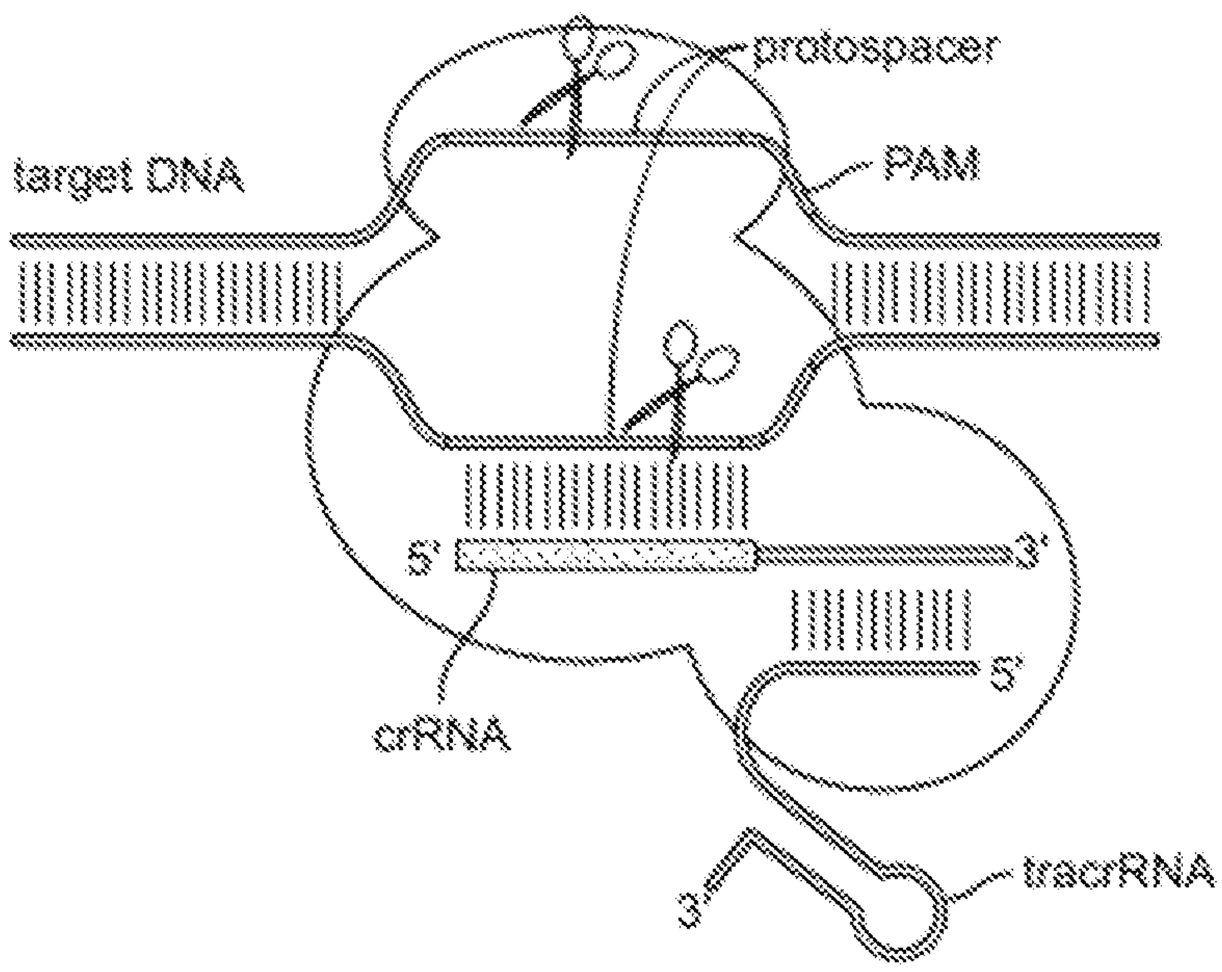

Cas9 programmed by crRNA:tracrRNA duplex
protospacer
target DNA
PAM
5'
3'
5'
crRNA
3'
tracrRNA
Cas9 programmed by single chimeric RNA
5'
20 nt
linker loop
3'
crRNA-tracrRNA
chimera

FIG. 2

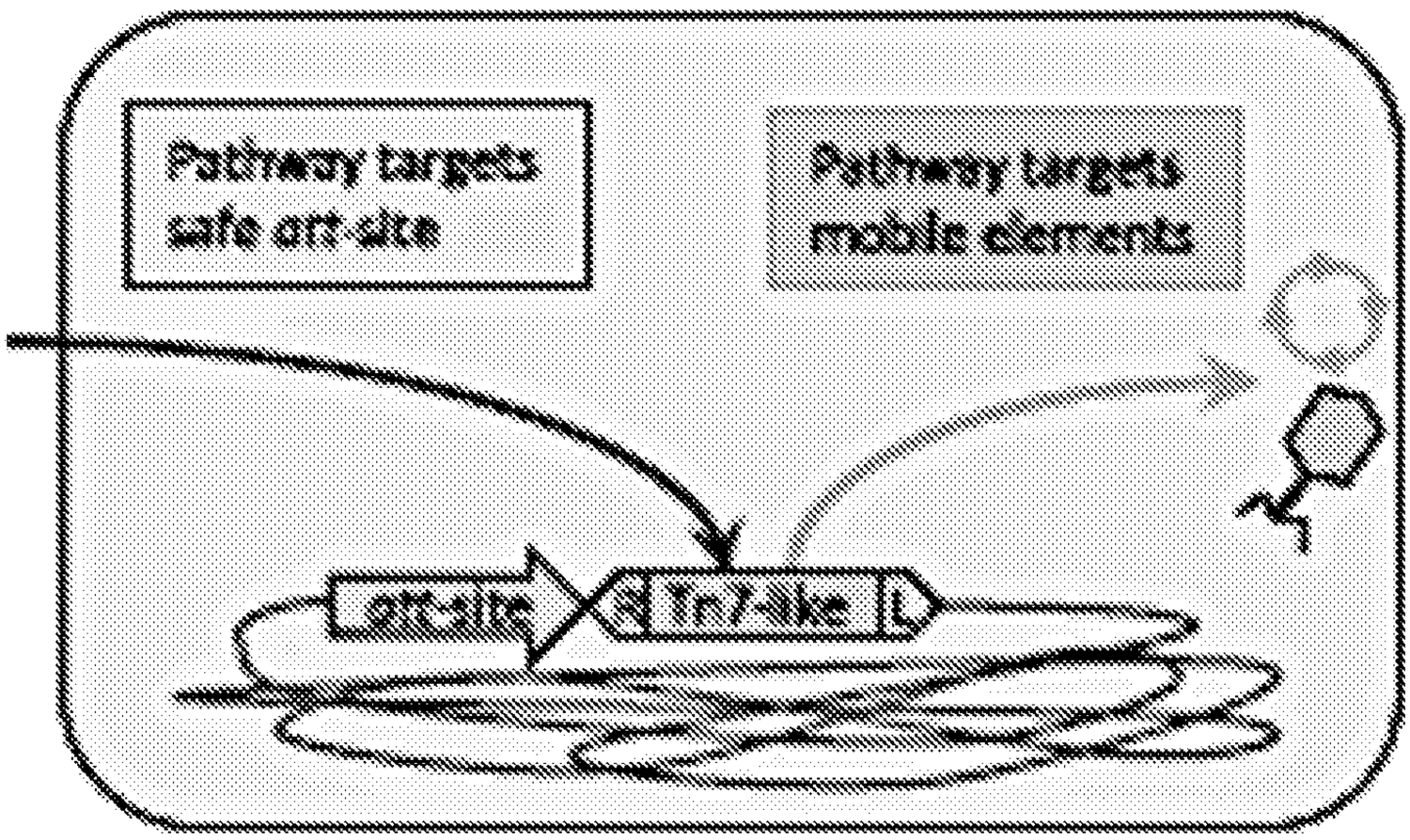

Pathway targets
safe art-site
Pathway targets
mobile elements

FIG. 3

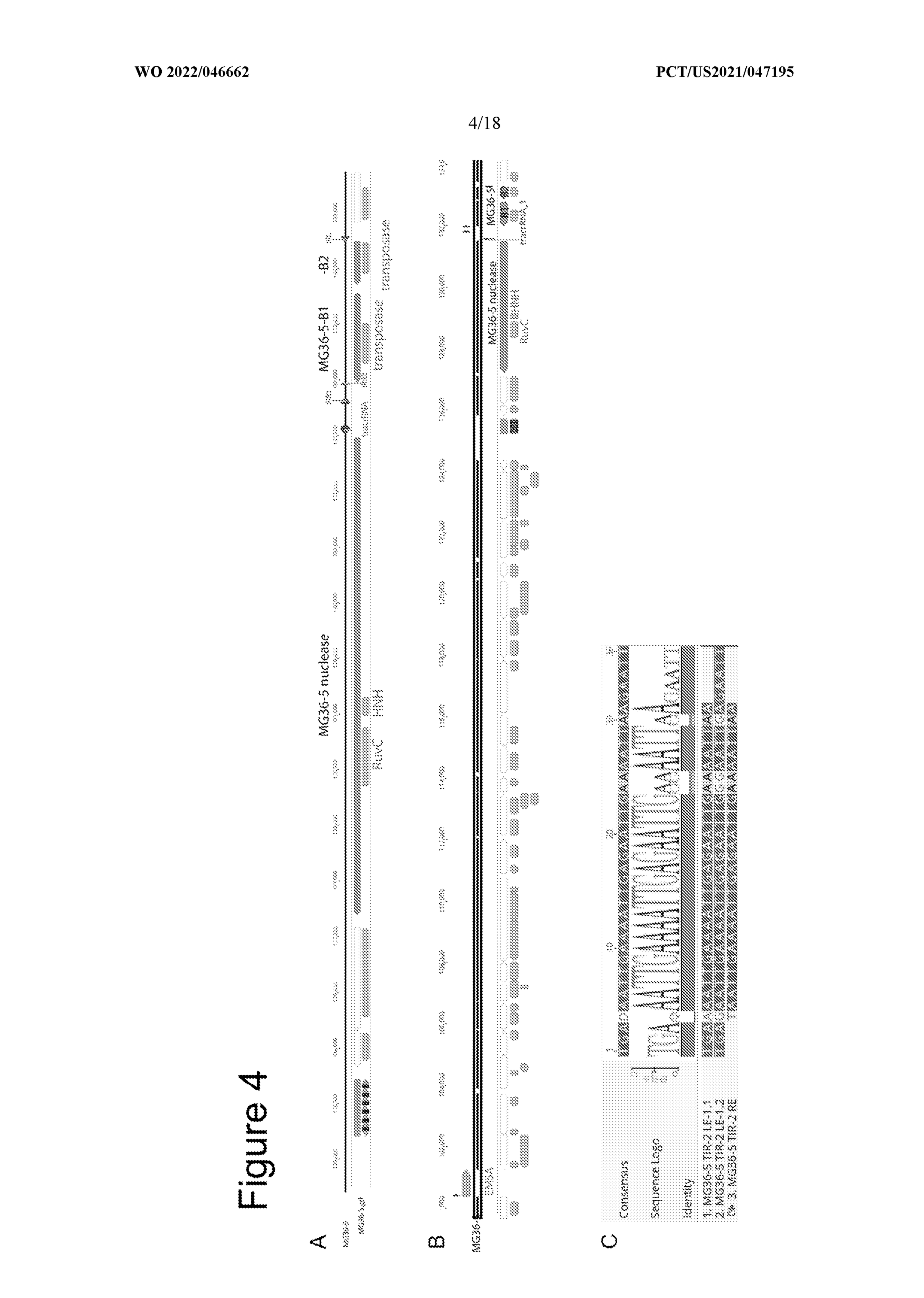

MG36-5-B1
-B2
transposase
MG36-5 nuclease
HNH
RuvC
MG36-5
transposase
MG36-5 nuclease
HNH
RuvC
A
B
C
Consensus
Sequence Logo
Identity
1. MG36-5 TIR-2 LE-1.1
2. MG36-5 TIR-2 LE-1.2
3. MG36-5 TIR-2 RE

Figure 4

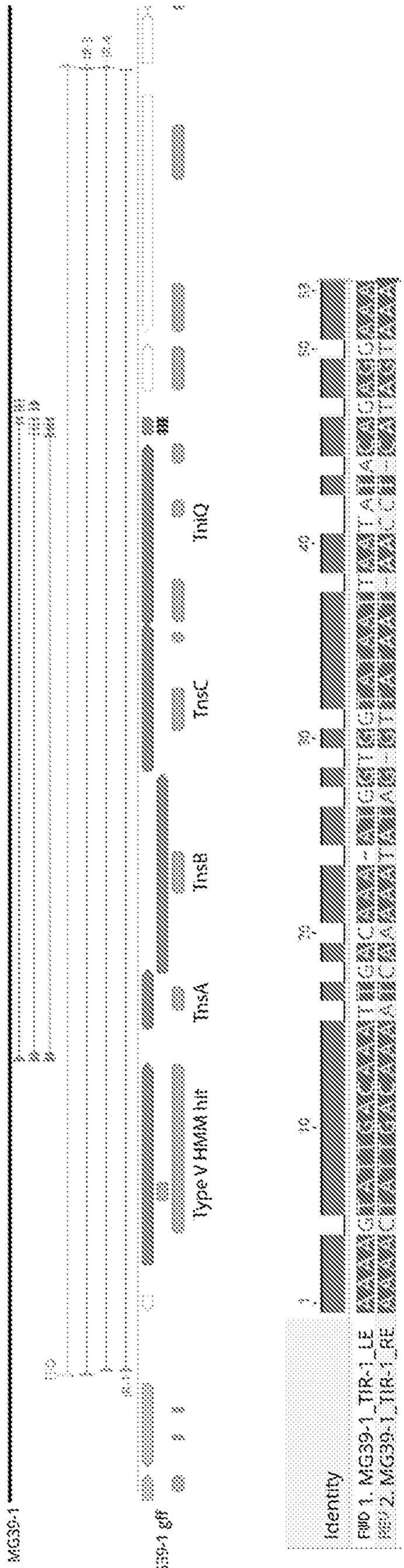

A
MG39-1
MG39-1 gff
TnsC
TniQ
TnsB
TnsA
Type V HMM hit
B
Identity
FIG 1. MG39-1_TIR-1_LE
FIG 2. MG39-1_TIR-1_RE

Figure 5

MG36-5 single guide RNA structure

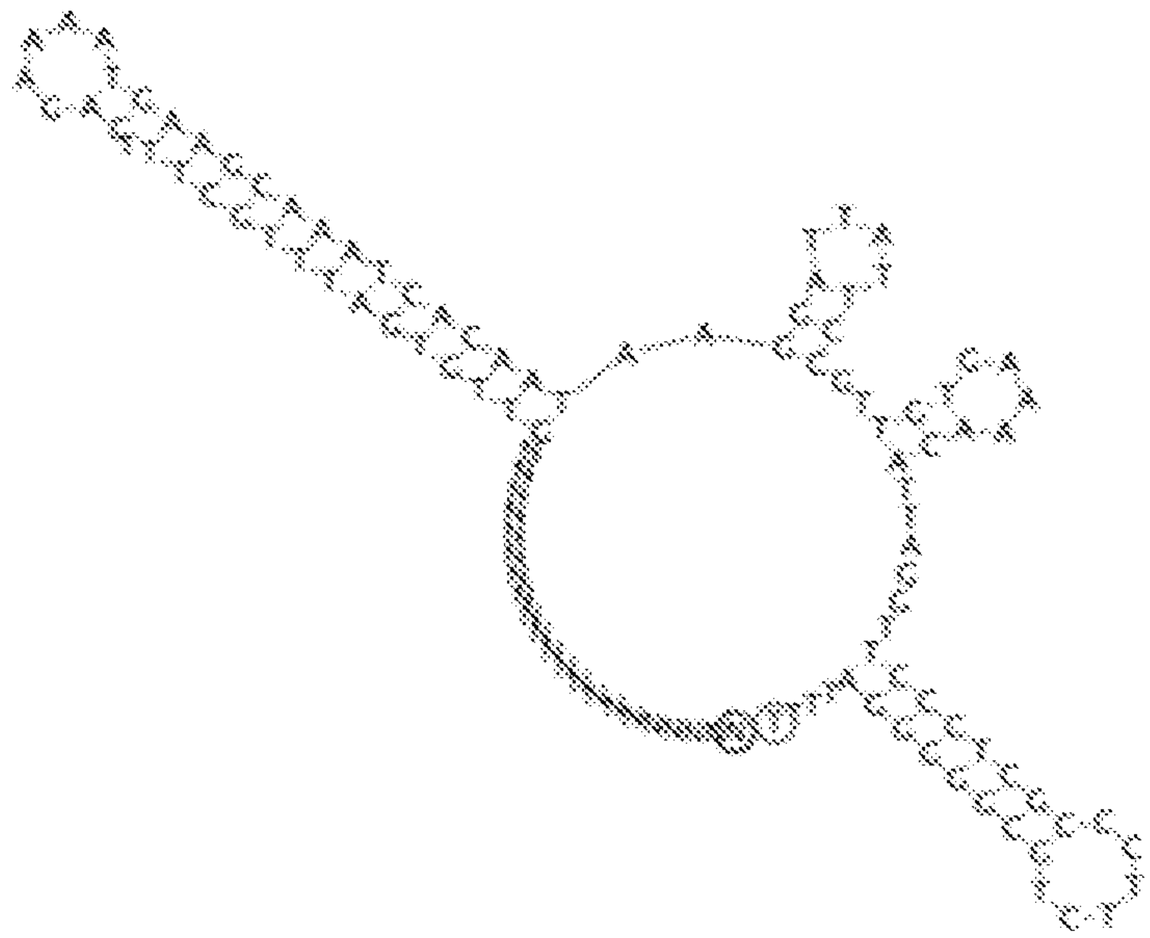
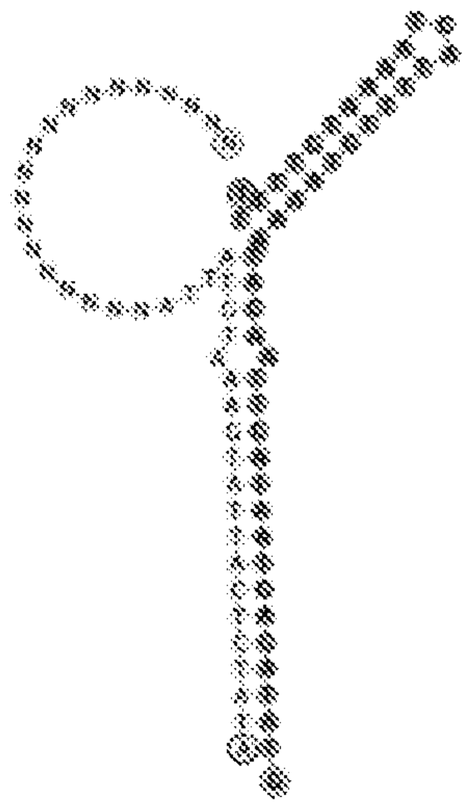
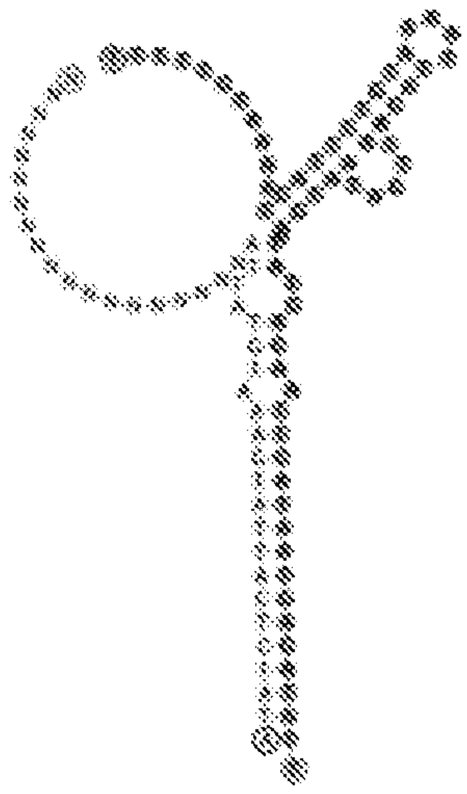


FIG. 6

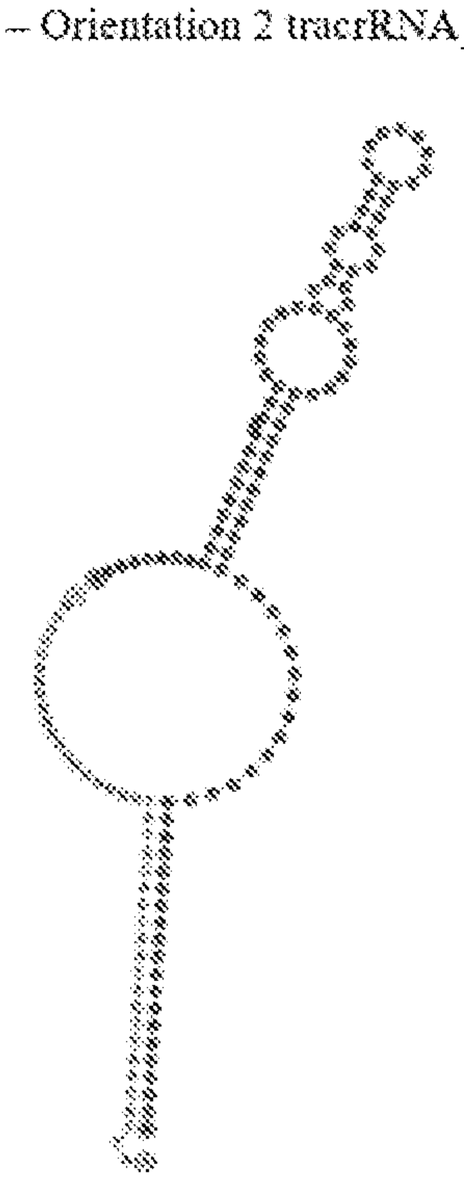
MG39-1 predicted double guide RNA structure – Orientation 1 tracrRNA_1



MG39-1 predicted double guide RNA structure – Orientation 1 tracrRNA_2



MG39-1 predicted double guide RNA structure – Orientation 2 tracrRNA_1



MG39-1 predicted double guide RNA structure – Orientation 2 tracrRNA_2

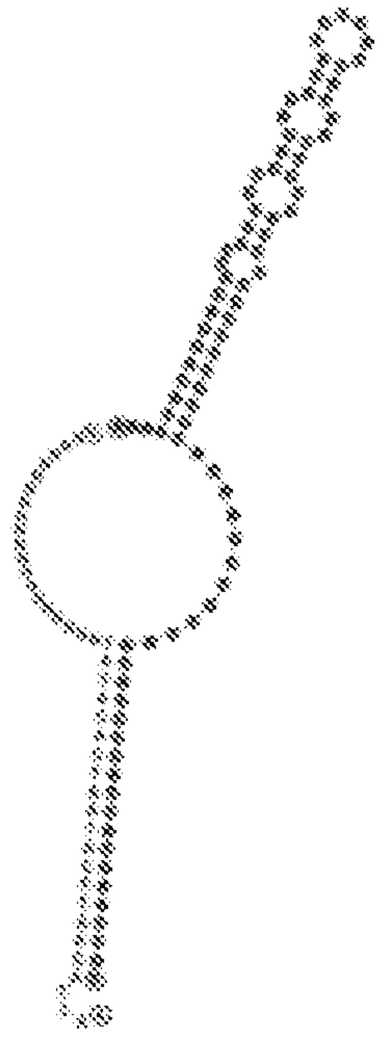


FIG. 7

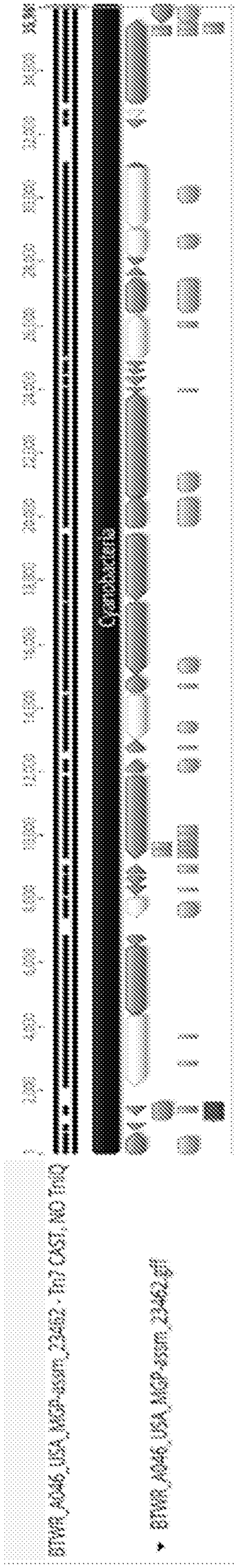

Candidate
BTN2A2 Q96A5 USA MSP-exon 2 2462 - fm3 CAST/NOD
BTN2A2 Q96A5 USA MSP-exon 2 2462 - fm3 CAST/NOD-exon 2 2462.gff

FIG. 8

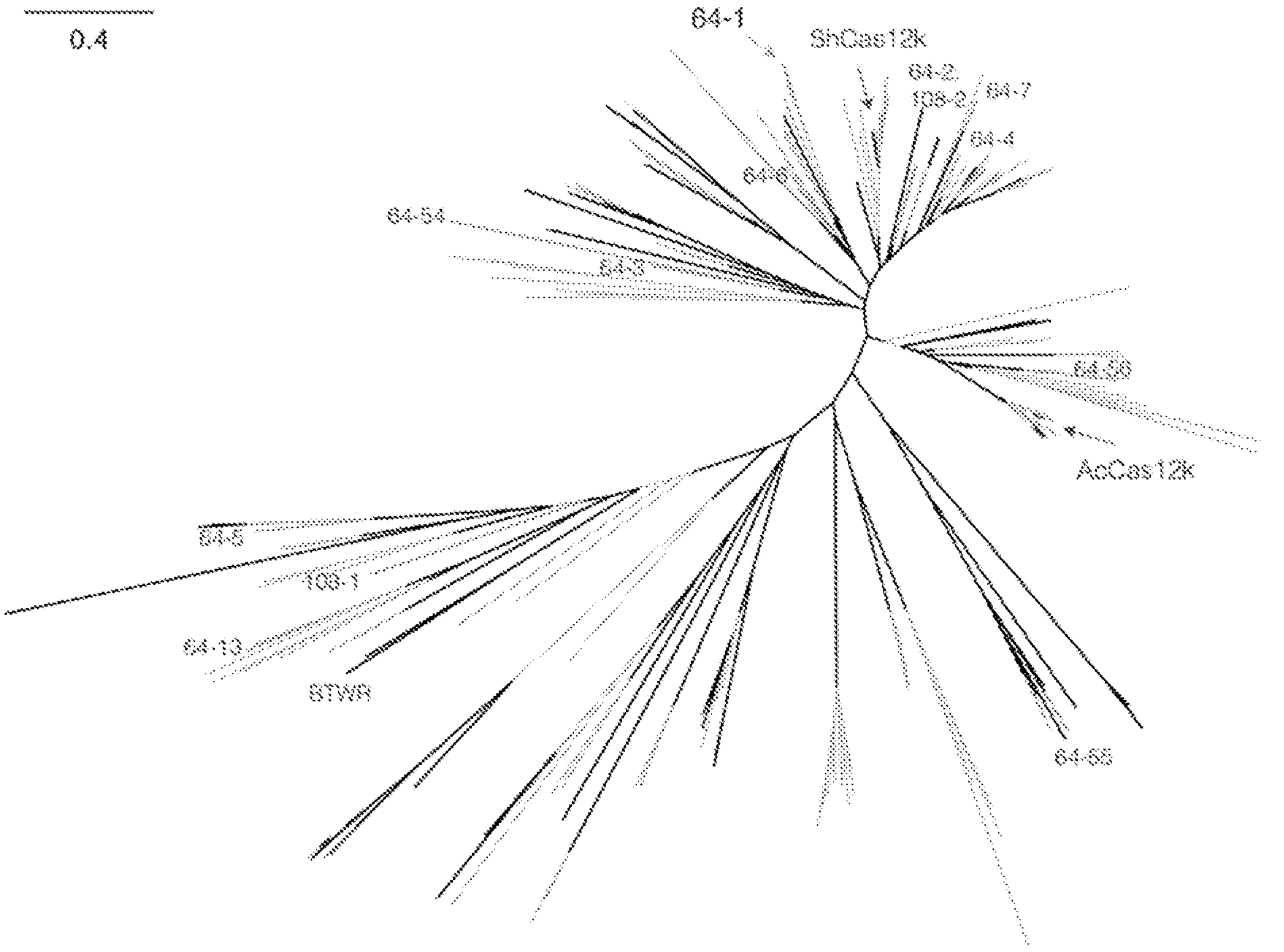

0.4
64-1
ShCas12k
64-2
108-2
64-7
64-4
64-6
64-54
64-3
64-56
AcCas12k
64-5
108-1
64-10
BTWR
64-65

FIG. 9

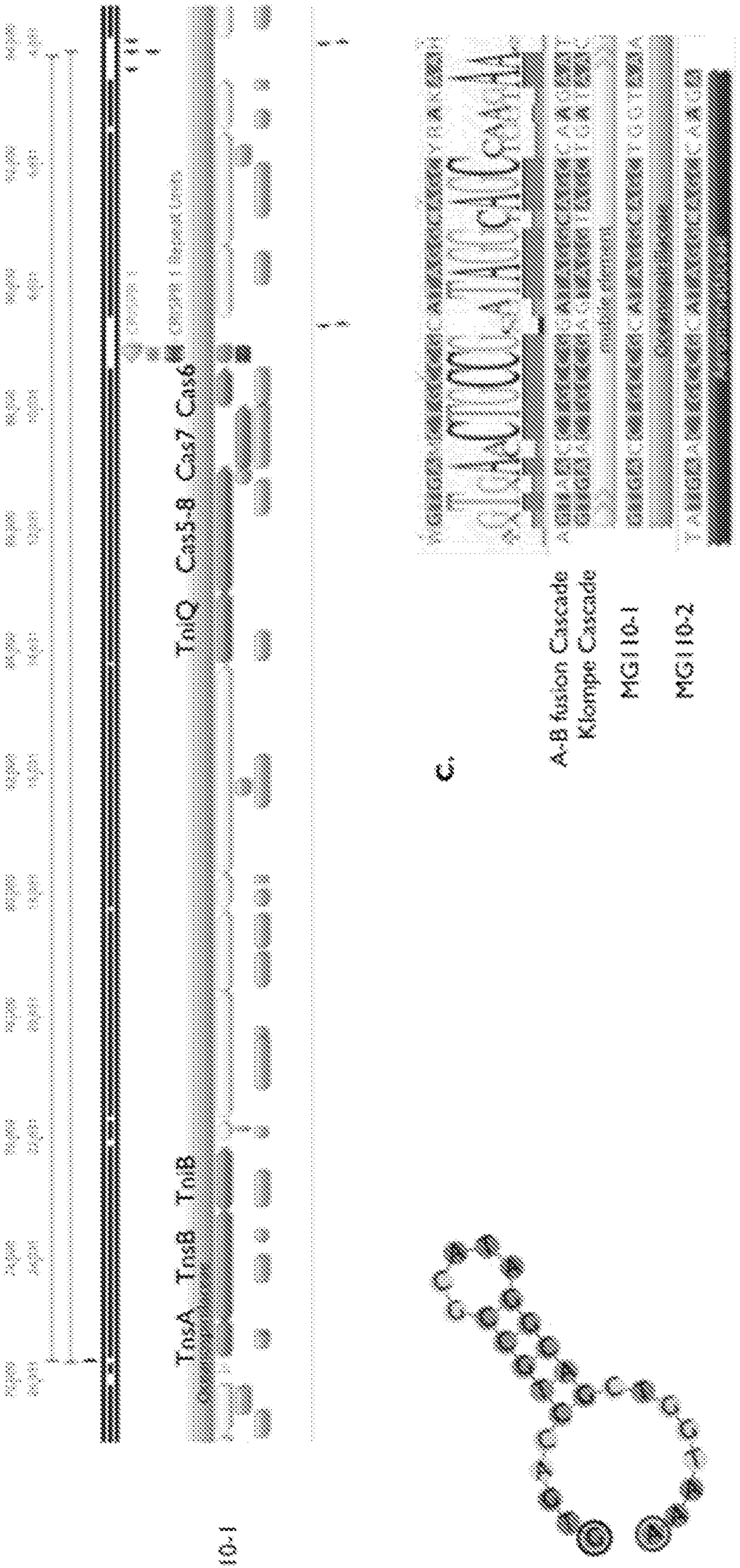

A.
MG110-1
TnsA TnsB TnsC TniQ Cas5-8 Cas7 Cas6
B.
C.
A-B fusion Cascade
Klampe Cascade
MG110-1
MG110-2

FIG. 10

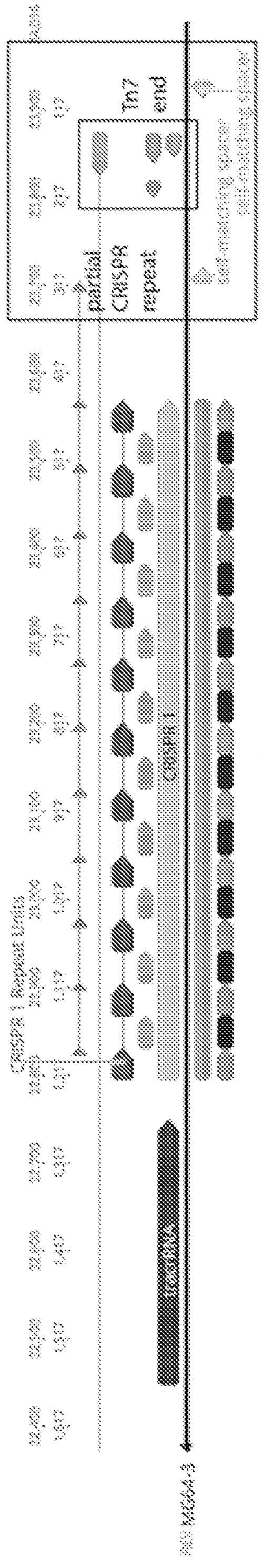

tracrRNA
partial CRISPR repeat
CRISPR 1 Repeat Units
CRISPR 1
tracrRNA-3

FIG. 11A

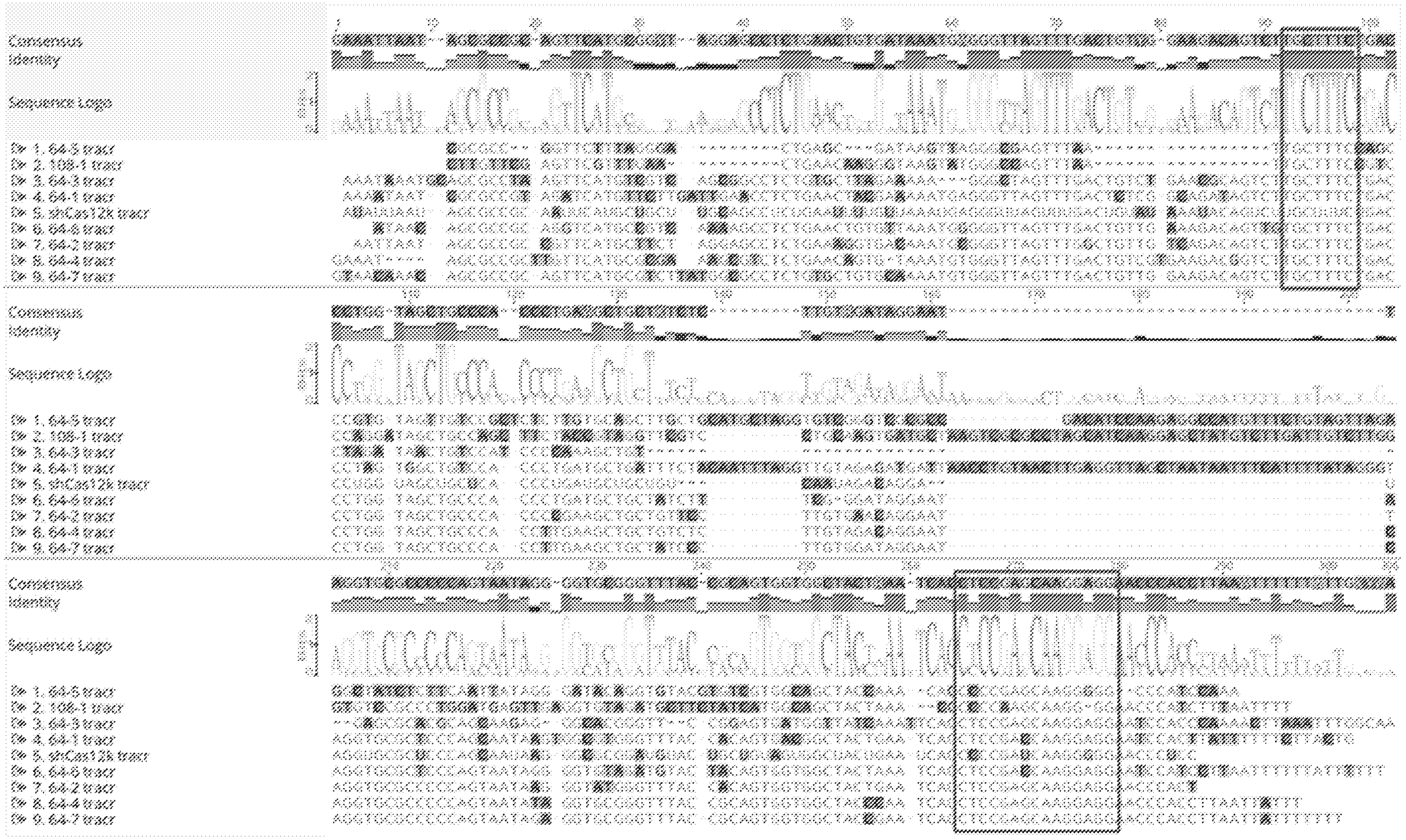
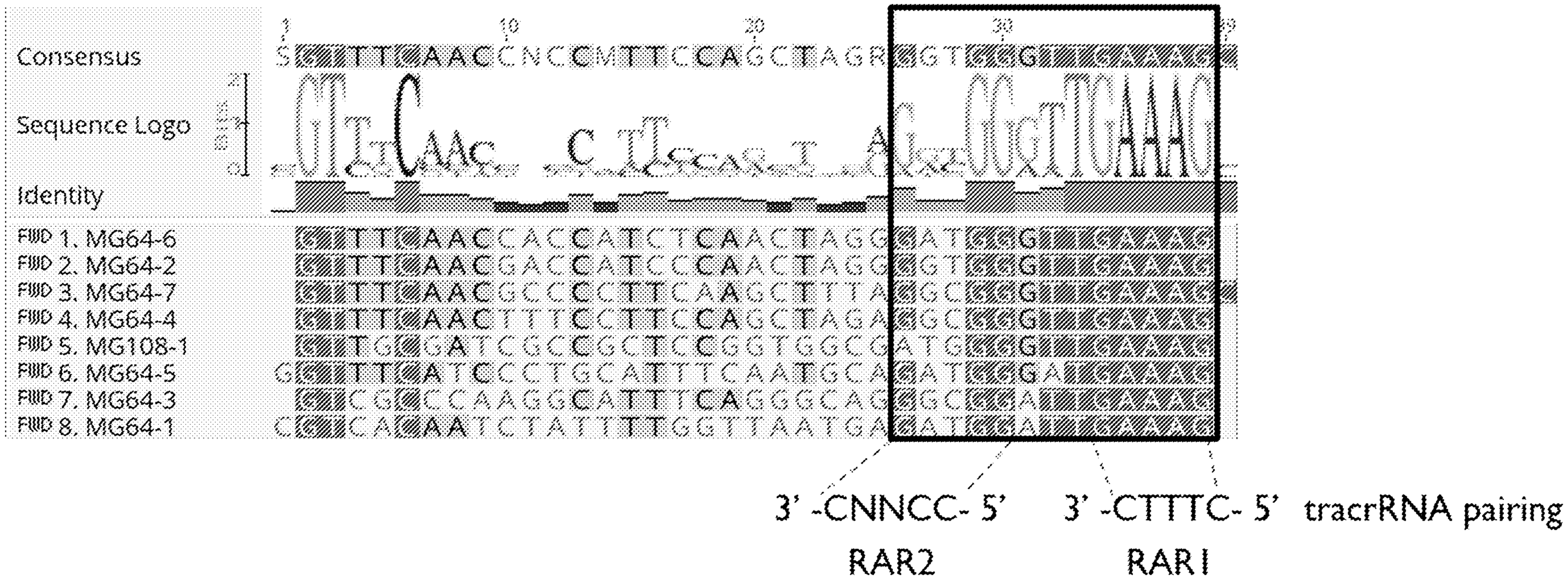

Consensus
Identity
Sequence Logo
tr 1. 64-5 tracr
tr 2. 108-1 tracr
tr 3. 64-3 tracr
tr 4. 64-1 tracr
tr 5. shCas12k tracr
tr 6. 64-8 tracr
tr 7. 64-2 tracr
tr 8. 64-4 tracr
tr 9. 64-7 tracr

FIG. 11B

FIG. 11C


Consensus
Sequence Logo
bits
Identity
FWD 1. MG64-6
FWD 2. MG64-2
FWD 3. MG64-7
FWD 4. MG64-4
FWD 5. MG108-1
FWD 6. MG64-5
FWD 7. MG64-3
FWD 8. MG64-1
3' -CNNCC- 5'
RAR2
3' -CTTTC- 5'
RAR1
tracrRNA pairing

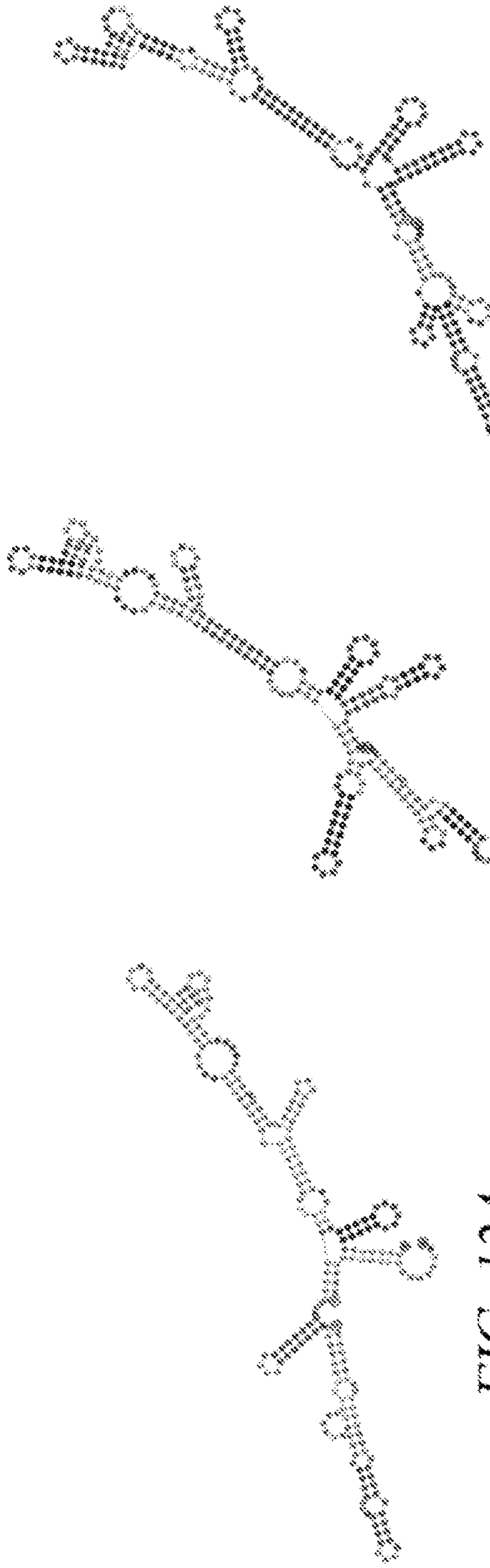


FIG. 12A

FIG. 12B

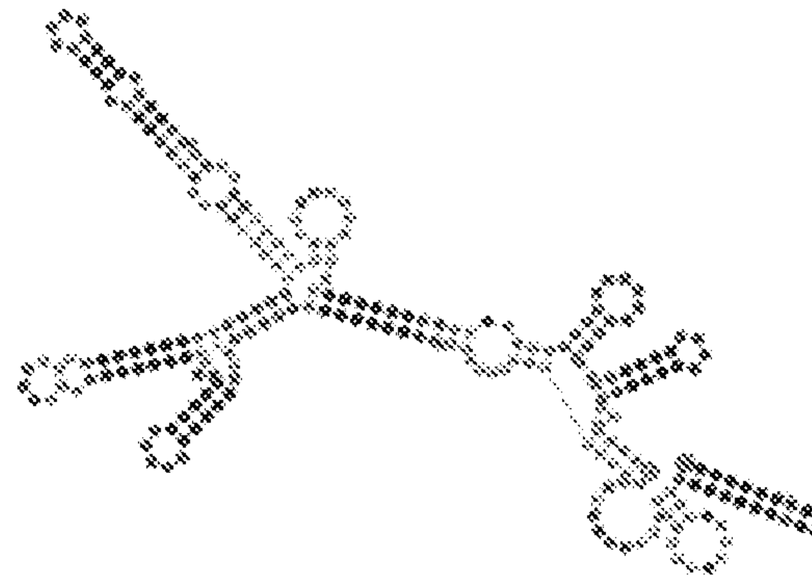


FIG. 12C

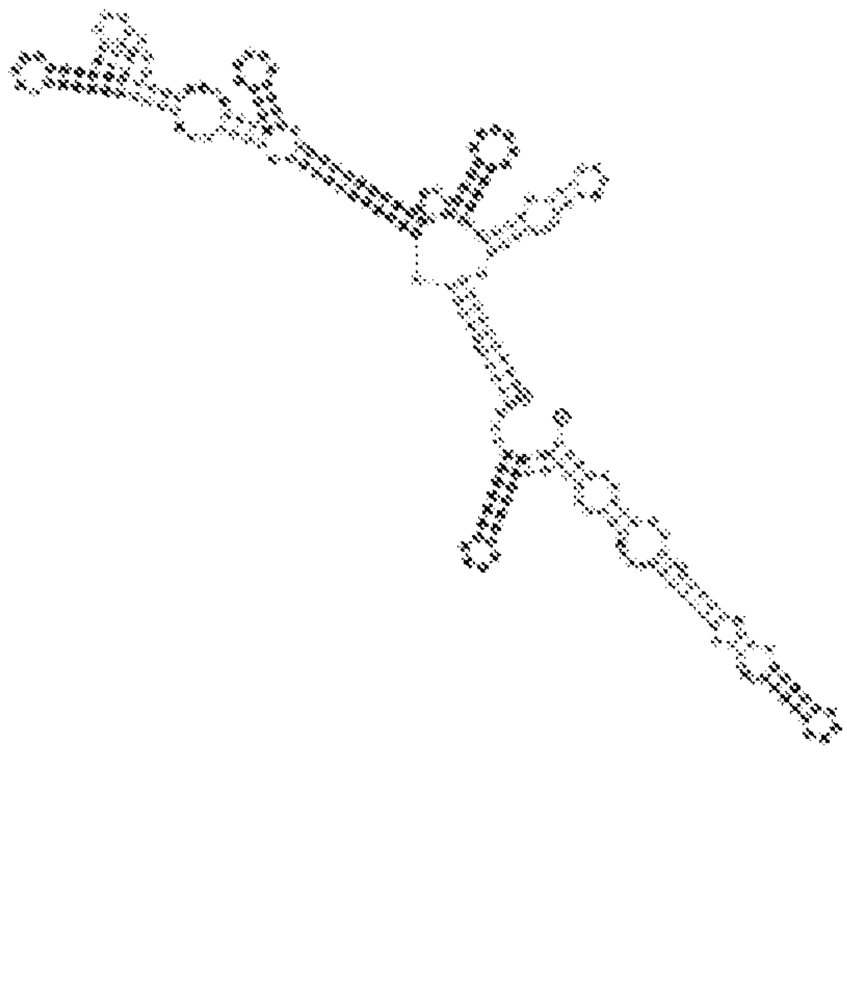


FIG. 12D

FIG. 12E

B.

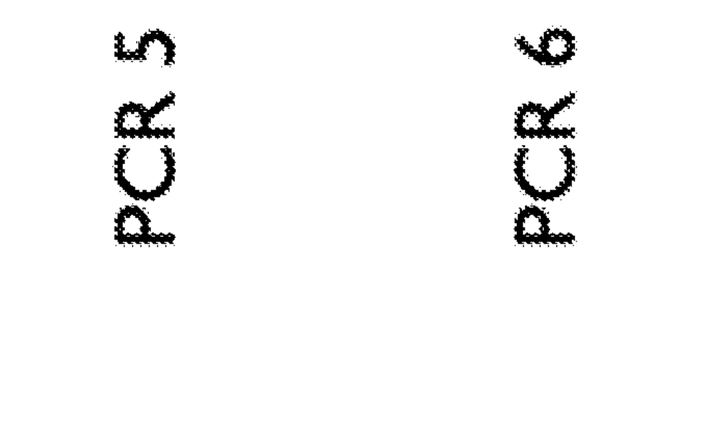


FIG. 13

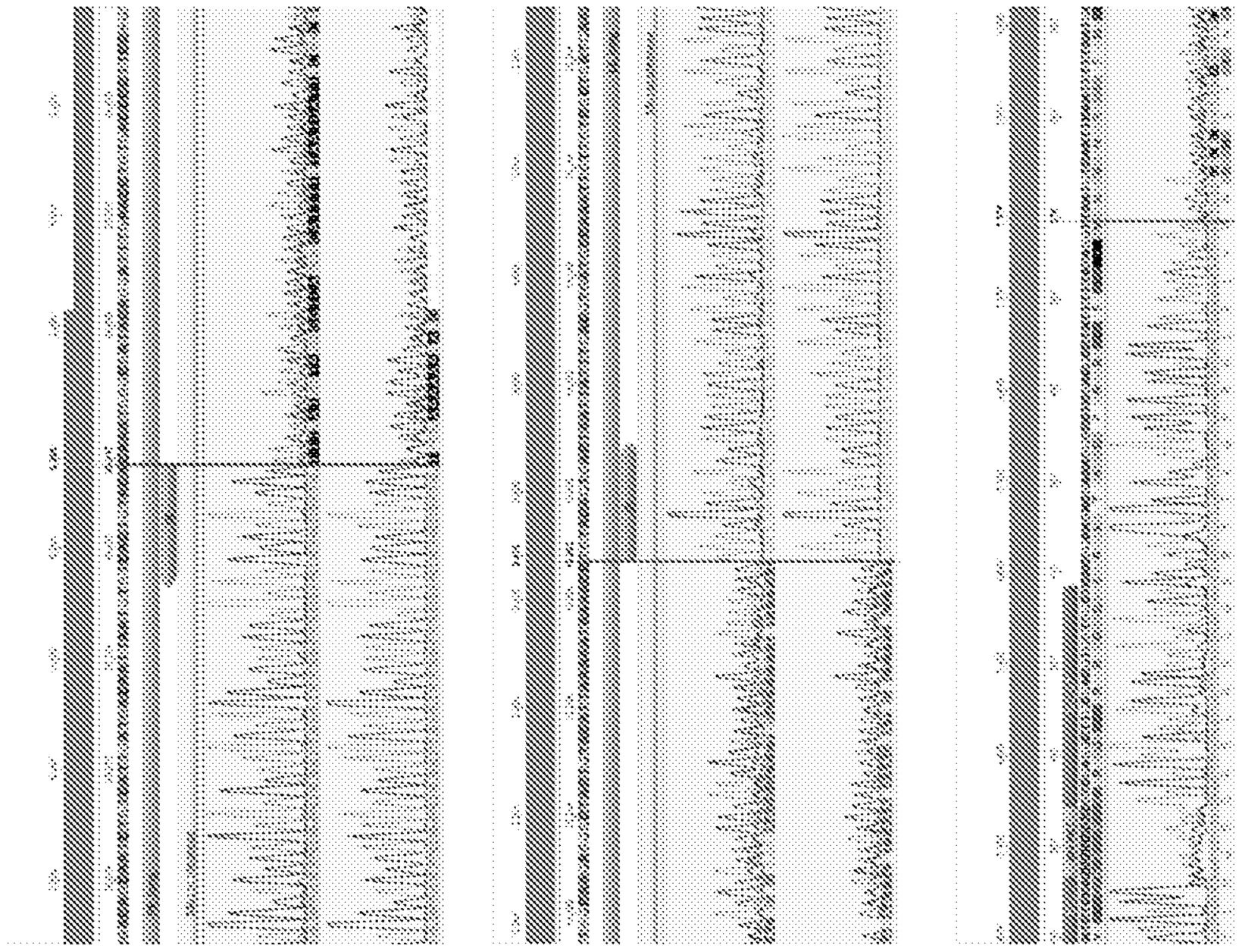

Reference sequence
PCR 4
PCR 6
Reference sequence
PCR 5
PCR 3
Reference sequence
PCR 6 showing PAM

Fig 13C

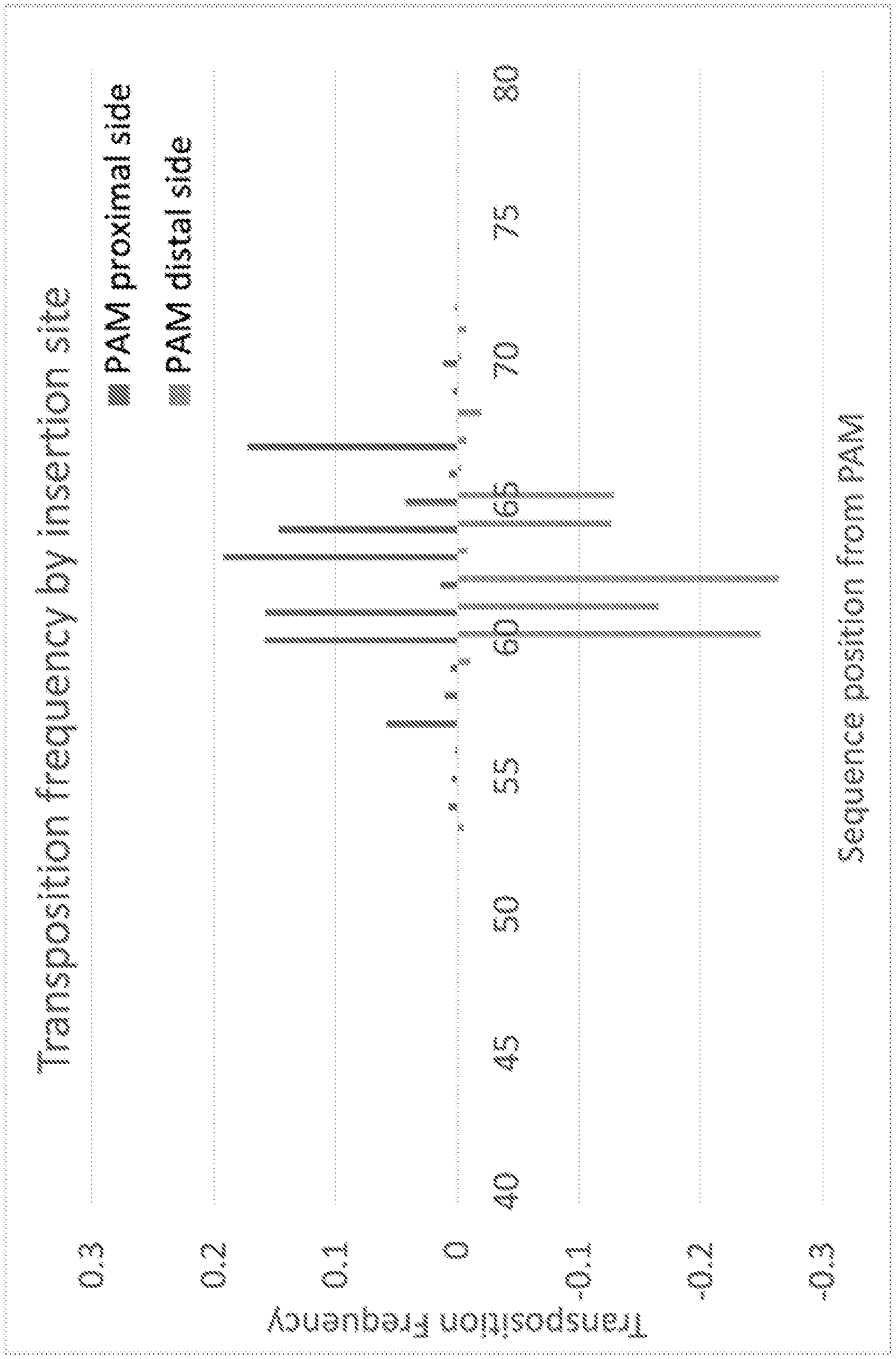

Transposition frequency by insertion site
PAM proximal side
PAM distal side
Transposition frequency
0.3
0.2
0.1
0
0.1
0.2
0.3
45
50
55
60
65
70
75
80
Sequence position from PAM

FIG. 14

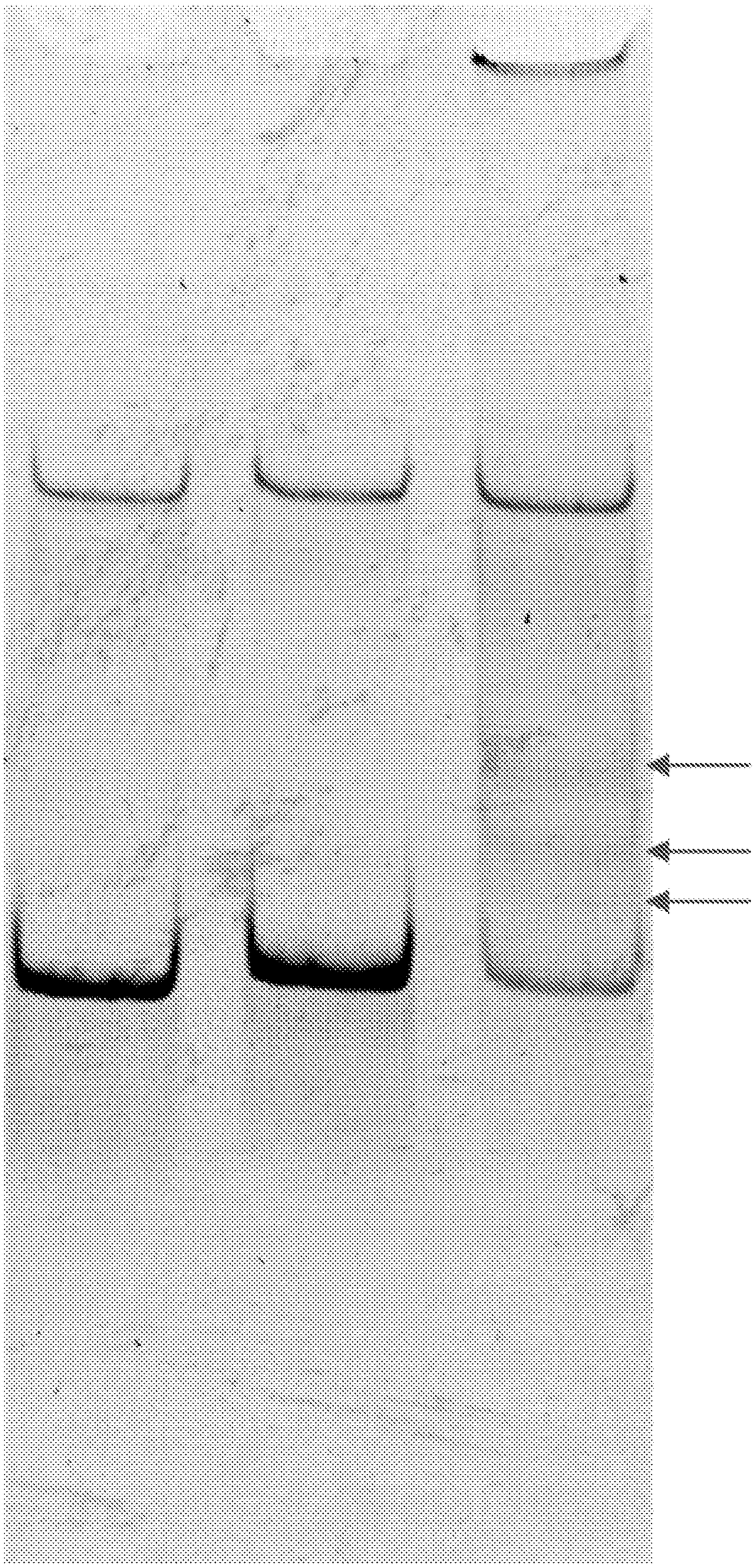


FIG. 15

SEQUENCE LISTING

<110> METAGENOMI IP TECHNOLOGIES, LLC

<120> SYSTEMS AND METHODS FOR TRANSPOSING CARGO NUCLEOTIDE SEQUENCES

<130> 55921-714.601

<140>
<141>

<150> 63/232,593
<151> 2021-08-12

<150> 63/186,698
<151> 2021-05-18

<150> 63/069,703
<151> 2020-08-24

<160> 261

<170> PatentIn version 3.5

<210> 1
<211> 1424
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG36 effector sequence

<220>
<223> MG36-5 effector

<400> 1
Met Lys Arg Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile Gly Trp
1 5 10 15

Ala Leu Val Asn Glu Ala Glu Thr Asp Glu Glu Val Ser Ser Ile Ile
20 25 30

Arg Leu Gly Val Arg Val Asn Pro Leu Thr Val Asp Glu Thr Gln Asn
35 40 45

Phe Glu Lys Gly Lys Ser Ile Thr Thr Asn Ala Glu Arg Thr Leu Lys
50 55 60

Arg Ser Met Arg Arg Asn Leu Gln Arg Tyr Lys Leu Arg Arg Asp Ala

```
65                  70                  75                  80

Leu Val Glu Phe Leu Lys Glu Leu Gly Phe Ile Thr Asp Glu Thr Ile
                85                  90                  95

Leu Ser Glu Gln Gly Asn Arg Thr Thr Phe Glu Thr Cys Arg Leu Arg
            100                 105                 110

Ala Lys Ala Ala Glu Glu Glu Ile Ser Leu Glu Glu Leu Ser Arg Val
        115                 120                 125

Leu Leu Met Ile Asn Lys Lys Arg Gly Tyr Lys Ser Ser Arg Lys Val
    130                 135                 140

Lys Lys Glu Glu Asp Gly Val Leu Ile Asp Gly Met Glu Ile Ala Lys
145                 150                 155                 160

Lys Leu Tyr Glu Glu Asp Leu Thr Pro Gly Gln Leu Cys Ser Gln Leu
                165                 170                 175

Leu Glu Ser Gly Lys Lys Val Leu Pro Asp Phe Tyr Arg Ser Asp Leu
            180                 185                 190

Gln Glu Glu Leu Asp Arg Ile Trp Ser Val Gln Lys Gln Phe His Pro
        195                 200                 205

Asp Val Phe Cys Asp Ala Ala Lys Glu Glu Ile Lys Gly Lys Asn Arg
    210                 215                 220

Ser Gln Thr Trp Ala Ile Leu Ala Asn Tyr Phe Val Trp Lys Glu Glu
225                 230                 235                 240

Val Val Gly Trp Asn Asp Arg Glu Ala Lys Asn Glu Thr Ile Glu Lys
                245                 250                 255

Glu Cys Lys Leu Val Gly Leu Lys Arg Thr Thr Lys Gly Tyr Glu Leu
            260                 265                 270

Lys Lys Glu Asn Tyr Arg Trp Arg Ala Gln Ala Leu Thr Glu Gln Leu
        275                 280                 285
```

```
Gly Leu Glu Glu Ile Ala Val Val Leu Gln Glu Ile Asn Gly Gln Ile
    290                 295                 300

Asn Ala Ser Ser Gly Tyr Leu Gly Ala Ile Ser Asp Arg Ser Lys Val
305                 310                 315                 320

Leu Tyr Phe Asn His Gln Thr Val Gly Gln Tyr Gln Met Ala Glu Leu
                325                 330                 335

Asp Lys Asn Pro His Ala Ser Leu Arg Asn Met Val Phe Tyr Arg Gln
            340                 345                 350

Asp Tyr Leu Asp Glu Phe Asp Arg Ile Trp Glu Lys Gln Ala Glu Phe
        355                 360                 365

His Lys Glu Leu Thr Ala Glu Leu Lys Lys Glu Ile Arg Asp Met Ile
    370                 375                 380

Ile Phe Tyr Gln Arg Arg Leu Lys Ser Gln Lys Gly Leu Ile Ser Phe
385                 390                 395                 400

Cys Glu Phe Glu Arg Leu Glu Ile Val Val Glu Lys Asp Gly Lys Lys
                405                 410                 415

Gln Thr Lys Val Ile Gly Cys Lys Val Ile Pro Arg Ser His Pro Leu
            420                 425                 430

Phe Gln Glu Phe Lys Val Trp Gln Thr Leu Asn Asp Ile Lys Val Leu
        435                 440                 445

Gly Arg Glu Lys Lys Arg Asn Ala Leu Ser Ala Ser Glu Ser Arg Ala
    450                 455                 460

Leu Tyr Pro Glu Glu Lys Glu Ile Leu Ala Arg Glu Leu Ala Ile Lys
465                 470                 475                 480

Glu Asn Met Lys Lys Ala Asp Val Leu Lys Leu Leu Phe Glu Asn Pro
                485                 490                 495

Gln Glu Leu Asp Leu Asn Phe Lys Gln Ile Asp Gly Asn Arg Thr Gly
```

```
            500                 505                 510

Phe Ser Leu Phe Ser Ala Tyr Ser Lys Met Ile Glu Lys Tyr Gly Tyr
        515                 520                 525

Glu Pro Leu Asp Phe Lys Lys Pro Ala Asp Glu Ile Ile Gly Gln Leu
    530                 535                 540

Arg Thr Ile Phe Thr Asp Leu Gly Trp Asn Thr Asn Leu Phe Thr Ile
545                 550                 555                 560

Asp Leu Thr Lys Glu Gly Lys Glu Leu Glu Met Gln Pro Tyr Phe Arg
                565                 570                 575

Leu Trp His Leu Leu Tyr Ser Phe Glu Gly Asp Asn Thr Ser Thr Gly
            580                 585                 590

Asn Gly Lys Leu Ile Glu Lys Ile Met Gln Leu Cys Gly Val Glu Lys
        595                 600                 605

Glu Tyr Ala Val Glu Leu Ala Ser Leu Ser Phe Gln Asp Asp Tyr Gly
    610                 615                 620

Ser Leu Ser Ala Lys Ala Ile Lys Arg Ile Leu Pro Tyr Leu Lys Glu
625                 630                 635                 640

Gly Asn Leu Tyr Asp Val Ala Cys Glu Tyr Ala Gly Tyr Arg His Ser
                645                 650                 655

Lys Ser Ser Leu Thr Lys Glu Glu Ile Glu Asn Lys Val Leu Lys Gly
            660                 665                 670

Lys Leu Glu Val Leu Pro Lys Asn Ser Leu Arg Asn Pro Val Val Glu
        675                 680                 685

Lys Ile Leu Ser Gln Met Val Asn Val Ile Asn Thr Ile Ile Asp Thr
    690                 695                 700

Tyr Gly Lys Pro Asp Glu Ile Arg Val Glu Leu Ala Arg Glu Leu Lys
705                 710                 715                 720
```

```
Lys Ser Ala Lys Glu Arg Glu Glu Leu Thr Lys Ala Ile Ala Lys Ser
                725                 730                 735

Thr Arg Glu His Glu Glu Ile Arg Lys Leu Leu Gln Asp Glu Phe Gly
            740                 745                 750

Met Met Asn Val Ser Arg Asn Asp Ile Ile Arg Tyr Lys Leu Tyr Glu
        755                 760                 765

Glu Leu Lys Asp Asn Gly Tyr Lys Thr Leu Tyr Ser Asn Gln Tyr Ile
    770                 775                 780

Pro Lys Glu Lys Ile Phe Ser Lys Glu Ile Asp Ile Glu His Ile Ile
785                 790                 795                 800

Pro Gln Ser Arg Leu Phe Asp Asp Ser Leu Ser Asn Lys Thr Leu Glu
                805                 810                 815

Tyr Lys Ala Ile Asn Ile Glu Lys Gly Asn Lys Thr Ala Tyr Asp Phe
            820                 825                 830

Val Lys Glu Lys Tyr Gly Glu Glu Gly Leu Gln Gln Tyr Ile Asn Arg
        835                 840                 845

Cys Glu Ser Leu Phe Asn Asp Lys Lys Ala Lys Leu Arg Lys Leu Lys
    850                 855                 860

Met Glu Gln Lys Asp Ile Pro Asp Gly Phe Ile Asp Arg Asp Leu Arg
865                 870                 875                 880

Asn Thr Gln Tyr Ile Ala Lys Lys Ala Leu Ala Met Leu Asn Glu Ile
                885                 890                 895

Cys Arg Arg Val Val Ala Thr Thr Gly Ser Ile Thr Asp Gln Leu Arg
            900                 905                 910

Glu Asp Trp Gln Leu Val Asp Val Met Lys Glu Leu Asn Leu Pro Lys
        915                 920                 925

Tyr Glu Ala Leu Gly Phe Val Glu Thr Tyr Glu Asp Lys Asp Gly Arg
```

```
    930                 935                 940

Lys Ile Lys Arg Ile Lys Asp Trp Thr Lys Arg Asn Asp His Arg His
945                 950                 955                 960

His Ala Met Asp Ala Leu Thr Val Ala Phe Thr Lys Asp Ala Phe Ile
                965                 970                 975

Gln Tyr Phe Asn Asn Lys Asn Ala Ala Gln Asp Ser Gly Asn Arg Asn
            980                 985                 990

His Thr Asn Ile Ile Gly Ile Lys  Thr Arg Tyr Phe Asp  Lys Gly Arg
        995                 1000                1005

Ala Leu  Pro Pro Ile Pro Leu  Glu Gln Phe Arg Ile  Glu Ala Lys
    1010                1015                1020

Gln His  Leu Glu Lys Leu Leu  Val Ser Ile Lys Ala  Lys Asn Lys
    1025                1030                1035

Val Val  Thr Val Asn Val Asn  Arg Thr Lys Lys Arg  Lys Gly Glu
    1040                1045                1050

Asn Thr  Lys Ile Gln Gln Thr  Pro Arg Gly Gln Leu  His Leu Glu
    1055                1060                1065

Thr Val  Tyr Gly Ser His Lys  Gln Tyr Val Thr Lys  Ile Glu Lys
    1070                1075                1080

Val Asn  Ala Ser Phe Asp Ala  Ala Lys Ile Ala Thr  Val Ser Lys
    1085                1090                1095

Arg Ala  Tyr Arg Asn Ala Leu  Leu Lys Arg Leu Glu  Thr Phe Gly
    1100                1105                1110

Asn Asp  Pro Lys Lys Ala Phe  Thr Gly Lys Asn Ala  Leu Glu Lys
    1115                1120                1125

Asn Pro  Leu Tyr Ala Asp Lys  Tyr Gln Thr Ile Arg  Val Pro Glu
    1130                1135                1140
```

```
Lys Val  Gln Thr Val Glu Phe  Glu Thr Ile Tyr Thr  Ile Arg Lys
    1145                1150                1155

Pro Val  Asp Pro Ala Leu Asn  Val Asp Lys Val Val  Asp Val Lys
    1160                1165                1170

Val Arg  Ala Ile Leu Glu Arg  Arg Leu Lys Glu Tyr  Gly Gly Asp
    1175                1180                1185

Pro Lys  Lys Ala Phe Val Asn  Phe Trp Glu Asn Pro  Ile Trp Leu
    1190                1195                1200

Asn Lys  Glu Lys Gly Ile Ser  Ile Lys Arg Val Ser  Ile Arg Gly
    1205                1210                1215

Ile Asn  Asn Ala Gln Ser Ile  His Val Lys Lys Asp  Lys Asn Gly
    1220                1225                1230

Asn Pro  Ile Trp Asp Glu Asn  Gly Lys Gln Ile Pro  Val Asp Phe
    1235                1240                1245

Val Asn  Thr Gly Asn Asn His  His Val Ala Ile Tyr  Arg Lys Pro
    1250                1255                1260

Val Leu  Asp Lys Lys Gly Gln  Val Thr Phe Asp Glu  Asp Gly Asn
    1265                1270                1275

Leu Met  Tyr Glu Leu Asp Glu  Val Val Val Pro Phe  Phe Glu Ala
    1280                1285                1290

Val Thr  Arg Ala Asn Leu Gly  Leu Pro Ile Ile Asp  Lys Asp Tyr
    1295                1300                1305

Arg Lys  Ser Glu Gly Trp Gln  Phe Leu Phe Ser Met  Lys Gln Asn
    1310                1315                1320

Glu Tyr  Phe Val Phe Pro Asn  Glu Lys Thr Gly Phe  Asn Pro Lys
    1325                1330                1335

Glu Val  Asp Leu Leu Asn Pro  Asp Asn Tyr Ala Met  Ile Ser Pro
```

```
    1340                1345                1350

Asn Leu  Phe Arg Val Gln Thr  Met Ser Lys Val Met  Tyr Gly Asn
    1355                1360                1365

Asn Val  Val Arg Asp Tyr Lys  Phe Arg His His Leu  Glu Thr Thr
    1370                1375                1380

Val Lys  Asp Met Lys Glu Leu  Lys Asp Ile Ala Tyr  Lys Gln Tyr
    1385                1390                1395

Lys Thr  Leu Ser Phe Gly Asn  Ser Val Val Lys Ile  Arg Ile Asn
    1400                1405                1410

His Ile  Gly Gln Ile Val Ser  Val Gly Glu Tyr
    1415                1420
```

<210> 2
<211> 264
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG36 transposition protein sequence

<220>
<223> MG36-5-B1 transposition protein

<400> 2

```
Met Lys Ala Asp Val Ser Glu Arg Gln Arg Val Lys Leu Leu Ala Asp
1               5                   10                  15

Met Ile Gly Tyr Tyr Arg Leu Cys Cys Pro Arg Ile Gly Gly Phe Lys
            20                  25                  30

Leu Phe His Leu Leu Glu Lys Asp Leu Gly His Ala Val Thr Leu Gly
        35                  40                  45

Arg Asp Ser Phe Leu Lys Val Tyr Glu Ser Lys Gly Phe Lys Leu Asn
    50                  55                  60

Pro Asn Lys Arg Arg Arg Thr Thr Asp Ser Asn His Val Tyr Lys Arg
65                  70                  75                  80
```

```
Tyr Pro Asn Leu Ile Lys Gly Lys Asp Ala Arg Tyr Ser Asn His Ile
                85                  90                  95

Trp Val Ser Asp Ile Thr Tyr Val Trp Ile Leu Gly Asp Val Leu Tyr
            100                 105                 110

Leu His Leu Val Thr Asp Ala Tyr Ser His Ala Val Leu Gly Trp Cys
        115                 120                 125

Leu Ser Asp Ser Leu Ser Ala Ser His Thr Thr Glu Ala Leu Arg Met
    130                 135                 140

Ala Ile Arg Ile Ala Gly Gly Gly Asn Leu Cys Gly Thr Ile His His
145                 150                 155                 160

Ser Asp Arg Gly Ser Gln Tyr Ala Ser Glu Ala Tyr Val Ser Cys Leu
                165                 170                 175

Met Glu His His Ile Arg Ile Ser Met Thr Glu Gly Tyr Glu Pro Thr
            180                 185                 190

Asp Asn Ala Met Ala Glu Arg Gln Asn Gly Ile Phe Lys Val Glu Trp
        195                 200                 205

Ile Tyr Glu Gln Glu Met Tyr Arg Asp Lys Glu Gln Ala Ile Asn Glu
    210                 215                 220

Ile Asn Arg Met Ile Asp Phe Tyr Asn Asn Arg Arg Pro His Met Ser
225                 230                 235                 240

Ile Gly Met Glu Cys Pro Met Glu Val Tyr Lys Gly Lys Leu Pro Gly
                245                 250                 255

Lys Asn Leu Trp Arg Lys Arg Pro
            260
```

<210> 3
<211> 126
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG36 transposition protein sequence

<220>
<223> MG36-5-B2 transposition protein

<400> 3
Met Ile Lys Arg Lys Pro Lys Arg Val Phe Thr Glu Pro Phe Lys Leu
1 5 10 15

Ser Val Leu Arg Glu Tyr Tyr Ser Ser Gly Met Ser Lys Tyr Ala Ile
20 25 30

Thr Lys Lys Tyr Ser Leu Ser Pro Pro Cys Ile Tyr Arg Trp Leu Lys
35 40 45

Glu Tyr Pro Val Gly Ser Asp Gln Leu Pro Leu Pro Ser Glu Thr Lys
50 55 60

Glu Arg Leu Gln Met Val Pro Lys Gln Ser Asp Leu Thr Asp Met Glu
65 70 75 80

Ser Leu Gln Lys Arg Ile Glu Glu Leu Arg Arg Ser Leu Glu Leu Glu
85 90 95

Lys Met Arg Ser Arg Ala Phe Glu Lys Met Ile Glu Ile Ala Glu Glu
100 105 110

Glu Glu Gly Ile Ser Ile Phe Arg Lys Asp Gly Ala Lys Gln
115 120 125

<210> 4
<211> 286
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG36 transposition protein sequence

<220>
<223> MG36-5-T1 transposition protein

<400> 4

```
Met Ala Arg Pro Asn Lys Glu Gly Leu Asp Tyr Phe Pro Phe Asp Val
1               5                  10                  15

Asp Phe Phe Ser Asp Glu Lys Ile Gly Ser Ile Ser Gly Glu Phe Gly
            20                  25                  30

Ile Lys Gly Glu Ile Thr Ala Ile Lys Leu Leu Cys Ala Ile Tyr Arg
        35                  40                  45

Asn Gly Tyr Phe Ile Leu Trp Asn Asp Ala Leu Lys Met Ser Leu Leu
    50                  55                  60

Arg Gly Leu Pro Gly Ile Ser Leu Glu Leu Leu Glu Gln Ile Val Thr
65                  70                  75                  80

Arg Leu Val Arg Trp Gly Phe Phe Glu Gln Thr Leu Phe Ser Thr Val
                85                  90                  95

Ser Val Leu Thr Ser Lys Gly Ile Gln Glu Arg Tyr Phe Lys Ala Ile
            100                 105                 110

Lys Arg Arg Lys Asp Ser Ser Asn Tyr Pro Tyr Leu Leu Val Asn Val
        115                 120                 125

Asp Asn Asn Lys Val Asn Val Ser Asn Asn Asp Ile Asn Val Asn Thr
    130                 135                 140

Asn Pro Ile Lys Glu Arg Lys Gly Asn Lys Asn Arg Glu Ser Leu Asn
145                 150                 155                 160

Thr Arg Glu Thr Leu Phe Asp Asn Phe Lys Asn Glu Leu Leu Arg Asp
                165                 170                 175

Glu Glu Trp Arg Arg Tyr Ala Cys Gln Ile Ser Gly Leu Ser Val Ala
            180                 185                 190

Phe Asn Asp Leu Ile Pro Gly Glu Leu Asp Asn Phe Leu Ala Trp Met
        195                 200                 205

Val Ser Thr Gly Glu Gly Asp Thr Leu Lys Thr Ile Asp Asp Val Lys
    210                 215                 220
```

```
Arg Arg Phe Thr Tyr Trp Trp Gln Gly Thr Gly Leu Arg Ala Tyr Asn
225                 230                 235                 240


Gln Arg Tyr Gly Gly Thr Arg Lys Glu Thr Phe Gly Gly Tyr Thr Ser
                245                 250                 255


His Ala Gly Ala Tyr Gly Lys Arg Glu Ala Pro Ala Lys Thr Gly Val
            260                 265                 270


Gln Pro Ser Glu Glu Ala Arg Lys Asp Tyr Thr Glu Arg Phe
        275                 280                 285


<210> 5
<211> 203
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
      MG36 transposition protein sequence

<220>
<223> MG36-5-C transposition protein

<400> 5
Met Ile Lys Ala Leu Gly Cys Asn Tyr Leu Gly Ile Glu Arg Arg Gln
1               5                   10                  15


Phe Glu Thr Asp Arg Gly Asn Asp Lys Val Leu Arg Phe Leu Leu Tyr
            20                  25                  30


Tyr Phe Asn Asp Cys Pro Leu Ala Glu Ser Val Phe Pro Glu Glu Asn
        35                  40                  45


Tyr Lys Leu His Lys Asn Leu Leu Ile Val Gly Asp Pro Gly Thr Gly
    50                  55                  60


Lys Thr Leu Met Met Gln Ile Phe Ala Asp Tyr Leu Lys Leu Thr Asp
65                  70                  75                  80


Asn Pro Lys Arg Phe Val Asn Leu Ser Val Thr Gln Met Met Asn Tyr
                85                  90                  95
```

Tyr Lys Ile His Gly His Ile Asp Arg Phe Thr Tyr Asn Glu Glu Ala
100 105 110

Gly Lys Gly Ser Met Glu Gly Asn Pro Phe Asp Ile Cys Leu Asn Asp
115 120 125

Ile Gly Leu Glu Thr Glu Asn Gln Lys Ser Tyr Gly Thr Ser Leu Asn
130 135 140

Ser Val Ile Asp Glu Phe Leu Tyr Ala Arg Tyr Glu Ile Tyr Gln Ser
145 150 155 160

His Gln Lys Lys Tyr His Ile Thr Ser Asn Leu Ser Val Thr Asp Phe
165 170 175

Lys Asn Arg Phe Gly Thr Arg Leu Val Asp Arg Phe Lys Ser Phe Asn
180 185 190

Val Ile Ile Leu Asn Gly Glu Ser Arg Arg Arg
195 200

<210> 6
<211> 724
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG39 effector sequence

<220>
<223> MG39-1 effector

<400> 6
Val Lys Ile Ile Lys Arg Ser Gly Ala Glu Asn Thr Phe Asp Lys Glu
1 5 10 15

Lys Ile Glu Asn Ala Val Ala Lys Ala Asn Ile Thr Val Glu Glu Lys
20 25 30

Asp Arg Leu Ser Glu Gly Glu Ile Gly Glu Ile Ala Gln Asn Ile Glu
35 40 45

```
Asp Lys Cys Ser Glu Met Asn Arg Ala Met Asp Val Glu Thr Ile Gln
    50                  55                  60

Asp Trp Val Glu Ala Asp Ile Met Arg His Gly Lys Tyr Thr Val Ala
65                  70                  75                  80

Lys His Tyr Ile Thr Tyr Arg Tyr Glu Arg Ser Ile Val Arg Gln Ala
                85                  90                  95

Asn Thr Thr Asp Lys Gln Ile Leu Ser Leu Leu Asn Phe Glu Asn Glu
            100                 105                 110

Glu Val Lys Gln Glu Asn Ser Asn Lys Asn Pro Thr Val Asn Ser Val
        115                 120                 125

Gln Arg Asp Tyr Met Ala Gly Glu Val Ser Lys Asp Ile Thr Arg Arg
    130                 135                 140

Phe Leu Leu Pro Asp Asp Ile Val Glu Ala His Glu Lys Gly Leu Ile
145                 150                 155                 160

His Phe His Asp Ala Asp Tyr Phe Ala Gln His Met His Asn Cys Cys
                165                 170                 175

Leu Val Asn Leu Glu Asp Met Leu Gln Asn Gly Thr Val Ile Ser Glu
            180                 185                 190

Val Met Ile Glu Lys Pro His Ser Phe Ser Thr Ala Cys Asn Ile Ala
        195                 200                 205

Thr Gln Ser Ile Ala Gln Ile Ala Ser Ser Gln Tyr Gly Gly Gln Ser
    210                 215                 220

Ile Thr Leu Ser His Leu Ala Pro Phe Val Gln Ile Ser Arg Asp Lys
225                 230                 235                 240

Tyr Arg Arg Glu Val Lys Lys Glu Phe Ala Glu Leu Asn Ile Pro Ala
                245                 250                 255

Asp Glu Asp Thr Ile Asn Lys Val Ala Glu Met Arg Val Lys Ala Glu
            260                 265                 270
```

```
Ile Val Gln Gly Val Gln Met Ile Gln Tyr Gln Val Ile Thr Leu Met
        275                 280                 285

Thr Thr Asn Gly Gln Ala Pro Phe Val Thr Val Phe Met Tyr Leu Asp
    290                 295                 300

Glu Val Pro Glu Gly Gln Thr Arg Asp Asp Leu Ala Ala Ile Ile Glu
305                 310                 315                 320

Glu Met Leu Arg Gln Arg Ile Gln Gly Val Lys Asn Glu Lys Gly Val
                325                 330                 335

Tyr Ile Thr Pro Ala Phe Pro Lys Leu Ile Tyr Val Leu Glu Glu Asp
            340                 345                 350

Asn Ile Arg Glu Gly Ser Lys Tyr Trp Glu Leu Thr Lys Leu Ala Ala
        355                 360                 365

Lys Cys Thr Ala Lys Arg Met Val Pro Asp Tyr Ile Ser Glu Lys Lys
    370                 375                 380

Met Lys Glu Leu Lys Val Asp Lys Asn Gly Asn Gly Gln Cys Tyr Pro
385                 390                 395                 400

Cys Met Gly Cys Arg Ser Phe Leu Thr Thr Tyr Leu Asp Glu Asn Gly
                405                 410                 415

Lys Pro Lys Tyr Tyr Gly Arg Phe Asn Gln Gly Val Val Thr Ile Asn
            420                 425                 430

Leu Val Asp Val Ala Cys Ser Ser Tyr Lys Asp Met Asp Lys Phe Trp
        435                 440                 445

Leu Ile Phe Asp Glu Arg Leu Glu Leu Cys Arg Arg Ala Leu Met Leu
    450                 455                 460

Arg His Glu Arg Leu Lys Gly Thr Pro Ser Asp Val Ala Pro Ile Leu
465                 470                 475                 480
```

```
Trp Gln Asn Gly Ala Leu Ala Arg Leu Lys Lys Gly Glu Thr Ile Asp
                485                 490                 495

Lys Leu Leu Phe Gly Gly Tyr Ser Thr Ile Ser Leu Gly Tyr Ala Gly
            500                 505                 510

Leu Cys Glu Cys Val Arg Tyr Met Thr Gly Lys Ser His Thr Asp Pro
        515                 520                 525

Ser Ala Thr Pro Phe Ala Leu Glu Val Met Gln His Leu Asn Asp Ala
    530                 535                 540

Cys Ala Lys Trp Arg Ala Glu Thr Asn Ile Asp Phe Ser Leu Tyr Gly
545                 550                 555                 560

Thr Pro Leu Glu Ser Thr Thr Tyr Lys Phe Ala Arg Cys Leu Gln Lys
                565                 570                 575

Arg Phe Gly Val Ile Glu Gly Val Thr Asp Arg Asn Tyr Ile Thr Asn
            580                 585                 590

Ser Tyr His Ile His Val Thr Glu Asn Ile Asp Ala Phe Asp Lys Leu
        595                 600                 605

Thr Phe Glu Ser Gln Phe Gln Ala Leu Ser Pro Gly Gly Ala Ile Ser
    610                 615                 620

Tyr Val Glu Val Pro Asn Met Gln Asn Asn Ile Glu Ala Val Leu Ala
625                 630                 635                 640

Val Met Gln His Ile Tyr Asp Asn Ile Met Tyr Ala Glu Leu Asn Thr
                645                 650                 655

Lys Ser Asp Tyr Cys Gln Lys Cys Gly Phe Asp Gly Glu Ile Lys Ile
            660                 665                 670

Val Glu Asp Asp Gly Lys Leu Val Trp Glu Cys Pro Asn Cys Gly Asn
        675                 680                 685

Arg Asp Gln Asn Thr Leu Asn Val Ala Arg Arg Thr Cys Gly Tyr Ile
    690                 695                 700
```

Gly Thr Gln Phe Trp Asn Gln Gly Arg Thr Gln Glu Ile Lys Glu Arg
705 710 715 720

Val Leu His Leu

<210> 7
<211> 214
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG39 transposition protein sequence

<220>
<223> MG39-1-A transposition protein

<400> 7
Met Ser Lys Arg Ile Ser Arg Ala Thr Lys Leu Ala Arg Gly Gln Gly
1 5 10 15

Ser Gly Glu Gly Ala Lys Tyr Lys Ser Tyr Val Met Thr Asn Glu Phe
20 25 30

Asn Ser Leu Gly Thr Thr Ser Val Ile Lys Asp Trp Lys Thr Gly Arg
35 40 45

Gly Val His Cys Leu Ser Gln Ala Glu Ala Leu Trp Phe Tyr Ile Leu
50 55 60

Arg Trp Asp Asp Asn Asn Ile Asp Ile Arg Glu Gln Tyr Pro Leu Asp
65 70 75 80

Arg Asn Ile Thr Ser Arg Ile Ala Asp Lys Tyr Gly Phe Lys His Pro
85 90 95

Gly Asn Ser Asp His Ile Met Thr Thr Asp Phe Leu Val Thr Lys Lys
100 105 110

Asn Asn Lys Leu His Ala Tyr Ser Val Lys Pro Asp Arg Asp Leu Ser
115 120 125

Lys Arg Thr Leu Glu Ile Leu Cys Ile Glu Lys Leu Tyr Trp Glu His
130 135 140

Asn Asn Ile Glu Phe Asp Met Leu Phe Lys Glu Asp Val Asn Thr Ile
145 150 155 160

Leu Ala Ser Asn Ile Arg Leu Val Thr Glu Tyr Tyr Asp Glu Ser Arg
165 170 175

Val Phe Asp Arg Tyr Ser Asn Ile Arg His Asn Ile Ala Thr Lys Lys
180 185 190

Ile Arg Cys Asp Met Glu His Lys Ile Leu Thr Asn Asp Asp Leu Asp
195 200 205

Arg Ile Trp Gly Glu Tyr
210

<210> 8
<211> 716
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG39 transposition protein sequence

<220>
<223> MG39-1-B transposition protein

<400> 8
Met Thr Ile Leu Ile Glu Tyr Gly Val Asn Thr Glu Met Asn Asp Tyr
1 5 10 15

Val Leu Ala Ile Gly Asp Ile Leu Tyr Asp Lys Asp Asn Gln Asn Ser
20 25 30

Tyr Arg Ile Ile Ser Leu Ile Asp Asp His Leu Ile Leu Cys Glu Met
35 40 45

Glu Thr Thr Lys Leu Glu Leu Gln Gln Ile Lys Tyr Thr Ile Ile Ala
50 55 60

```
Asp Leu Val Leu Ser Asn Lys Ile Glu Ile Lys Lys Asp Gln Ala Leu
65                  70                  75                  80

Val Tyr Asp Ile Asp Gln Leu Ser Glu Ser Val Arg Asn Arg Tyr Ile
                85                  90                  95

Leu Lys Val His Ile Met Asn Asp Val Lys Ile Ala Tyr Gly Pro Ser
            100                 105                 110

Tyr Leu Gly Leu Met Gly Lys Ser Ser Lys Ile Glu Leu Gln Lys Ile
        115                 120                 125

Leu Ala Lys Tyr Asn Tyr Pro Ile Ser Ser Phe Trp Arg Met Cys Thr
    130                 135                 140

Thr Tyr Phe Gln Ser Gly Met Lys Asn Tyr Ser Leu Ile Asn Ala Lys
145                 150                 155                 160

Ser Phe Lys Ser Asn Glu Val Lys Thr Tyr Thr Tyr Lys Ala Arg Pro
                165                 170                 175

Gly Ala Lys Ser Thr Tyr Asn Leu Asp Asn Val Val Lys Pro Asn Glu
            180                 185                 190

Tyr Val Ile Tyr Phe Glu Glu Ala Leu Asn Glu Tyr Lys Ala Gly Arg
        195                 200                 205

Glu Lys Thr Leu Lys Asn Ala Phe Ser Arg Met Asn Ile Leu His Phe
    210                 215                 220

Thr Gln Ala Glu Ile Ile Asp Gly Val Ala Thr Arg Leu Leu Leu Pro
225                 230                 235                 240

Glu Cys Gln Arg Pro Thr Tyr Lys Gln Phe Tyr Tyr Tyr Ala Gln Lys
                245                 250                 255

His Leu Thr Lys Glu Glu Lys Asp Leu Ile Lys Thr Ser Ala Ala Glu
            260                 265                 270

Gln Arg Asn Asn Lys Arg Leu Leu Ile Ser Asp Ser Leu Lys Asp Val
        275                 280                 285
```

```
Tyr Gly Pro Ala Asp Met Val Glu Ile Asp Ala Cys Glu Ala Asp Val
    290                 295                 300

Ser Leu Val Ser Glu Leu Asp Pro Asp Gln Ala Ile Gly Arg Pro Ile
305                 310                 315                 320

Val Tyr Phe Met Ile Asp Val Tyr Ser Arg Ile Ile Leu Ala Val Ser
                325                 330                 335

Val Ala Phe Asp Asn Asn Ser Ile Leu Gly Ile Thr Asn Leu Phe Leu
            340                 345                 350

Asn Leu Ala Asp Asp Lys Gln Glu Tyr Cys Lys Lys Tyr Gly Ile Glu
        355                 360                 365

Phe Asn Asp Lys Arg Leu Trp Pro Ser Gly Val Ile Pro Lys Arg Ile
    370                 375                 380

Arg Val Asp Arg Gly Ser Glu Phe Lys Ser Tyr Glu Phe Asp Arg Ile
385                 390                 395                 400

Cys Asn Glu Leu Gly Ile Glu Lys Gln Ile Val Ser Gly Ala Ser Gly
                405                 410                 415

Ser Leu Lys Gly Val Val Glu Gln Ala Phe His Gln Met His Ala Lys
            420                 425                 430

Gln Asn Val His Leu Glu Asn His Gly Leu Ile Glu Lys Arg Tyr Asp
        435                 440                 445

Ser Leu His His Lys Glu Ala Ser Leu Thr Ile His Asp Tyr Thr Arg
    450                 455                 460

Met Val Ile Asn Phe Val Leu Ala His Asn Gln Gln His Leu Glu Thr
465                 470                 475                 480

Tyr Pro Leu Thr Lys Glu Met Ile Glu Lys Asn Ile Ala Pro Val Pro
                485                 490                 495
```

```
Ala Ile Leu Trp Glu Tyr Gly Ser Lys Lys Tyr Gly Met Pro Gln Pro
            500                 505                 510

Ile Pro Val Leu Glu Gln Tyr Leu Phe Ser Leu Met Thr Pro Ile Lys
        515                 520                 525

Ala Lys Ile Ser Lys Arg Gly Ile Ser Tyr Lys Gly Leu Trp Tyr Phe
    530                 535                 540

Ala Pro Asn Asp Lys Arg Leu Met Ser Glu Met Tyr Ala Ala Gly Thr
545                 550                 555                 560

Arg Arg Met Pro Phe Glu Val Arg Met Asp Met Arg Asp Val Gly Ala
                565                 570                 575

Ile Tyr Tyr Ile Arg Asn Ser Lys Leu Val Lys Ile Pro Leu Asn Val
            580                 585                 590

Leu Ile Thr Gly Asn Ser Asp Tyr Lys Gly Leu Thr Met Lys Gln Tyr
        595                 600                 605

Glu Glu Tyr Tyr Ser Ala Lys Lys Lys Met Gln Ala Lys Gly Arg Ile
    610                 615                 620

Asp Asn Glu Lys Ile Asp Thr Ala Val Tyr Ala Asn Asn Glu Ser Ile
625                 630                 635                 640

Val Lys Ser Ala Lys Lys Asn Val His Ser Arg Thr Lys Asp Ile Arg
                645                 650                 655

Ser Ser Arg Glu Ile Asp Lys Gln Lys Val Ser Tyr Glu Gly Lys Ile
            660                 665                 670

Ser Ala Arg Leu Glu Asn Lys Asp Glu Thr Ser Lys Asn Ile Ser Ser
        675                 680                 685

Asp Glu Gln Lys Glu Asn Gly Val Thr Glu Tyr Arg Asp Tyr Ala Ser
    690                 695                 700

Phe Glu Glu Ala Leu Gln Asp Phe Tyr Asp Asn Asn
705                 710                 715
```

<210> 9
<211> 534
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG39 transposition protein sequence

<220>
<223> MG39-1-C transposition protein

<400> 9
Met Ala Lys Tyr Asn Tyr Val Thr Cys Glu Glu Val Val Ser Gln Tyr
1 5 10 15

Asp Asp Tyr Asp Leu Lys Glu Gln Arg Ile Val Pro Ala Arg Tyr Val
20 25 30

Glu Ala Lys Leu Ala Ile Asp Thr Gly Asn Pro Tyr Ile Glu Ala Leu
35 40 45

Pro Tyr Pro Arg Thr Gly Arg Asn Ile Ile Ser Ser Tyr Ser Gln Thr
50 55 60

Met Ala Asp Phe Asp Tyr Asp Lys Ile Lys Ser Met Ser Thr Ile Asp
65 70 75 80

Lys Ile Leu Gln Ile Arg Ser Leu Arg Ser Ile Arg Phe Pro Leu Pro
85 90 95

Phe His Ala Glu Leu Glu Leu Ser Phe Tyr Asn Ala Leu Ile Thr Ser
100 105 110

Tyr Arg Ser Arg His Ile Leu His Ser Asp Asn Asp Lys Val Ser Tyr
115 120 125

Ser Val Glu Asn Gln Glu Tyr Ala Ala Ser Asn Ile Leu Val Gly Asp
130 135 140

Ser Ser Ala Ser Thr Asp Ala Gly Phe Ser Leu Ile Gly Tyr Ser Gly
145 150 155 160

```
Cys Gly Lys Ser Ser Ala Ile Gln Met Leu Val Ser Tyr Tyr Pro Gln
                165                 170                 175

Val Ile Met His Thr Thr Glu Asn Gly Glu Tyr Phe Pro Gln Ile Thr
            180                 185                 190

Tyr Leu Val Val Asn Cys Ile Pro Asn Ser Asn Phe Ser Ala Leu Tyr
        195                 200                 205

Asp Gly Ile Gly Asp Ala Ile Asp Lys Ala Leu Gly Asn Ile Lys Pro
    210                 215                 220

Ile Tyr Ser Ala Glu Ile Met Lys Ile Arg Thr Leu Gly Ala Lys Ala
225                 230                 235                 240

Glu Arg Ile Arg Glu Tyr Val Glu Lys Phe Ala Ile Gly Ile Ile Ile
                245                 250                 255

Phe Asp Glu Ile Gln Leu Ile Asp Phe Ser His Thr Arg Glu Asn Ser
            260                 265                 270

Phe Asp Ser Leu Leu Thr Leu Ser Asn Arg Thr Lys Val Ala Thr Ala
        275                 280                 285

Ile Val Gly Thr Glu Asp Ala Lys Ala Lys Met Phe Lys Thr Leu Arg
    290                 295                 300

Thr Ala Arg Arg Val Gly Asn Val Ile Asn Gly Asn Met Tyr Cys Met
305                 310                 315                 320

Asp Lys Lys Phe Phe Tyr Phe Leu Val Asn Glu Leu Phe Arg Tyr Gln
                325                 330                 335

Trp Phe Asp Lys Pro Val Thr Val Ser Glu Glu Ile Thr Asp Ala Leu
            340                 345                 350

Tyr Asp Val Thr Lys Gly Ile Val Asp Gln Leu Ile Gly Ile Tyr Ile
        355                 360                 365

Tyr Met Asn Ile Asp Tyr Leu Asn Lys Lys Lys Lys Pro Glu Ile Asn
```

```
    370                 375                 380

Ser Lys Tyr Ile Tyr Ala Val Thr Asn Lys His Phe Ala Gly Leu Arg
385                 390                 395                 400

Asp Val Leu Glu Asn Leu Asp Ser Asp Phe Asn Asn Glu Leu Met Gln
                405                 410                 415

Glu Ile Asn Asp Asn Ala Lys Leu Glu Leu Asp Lys Ile Glu Asp Glu
            420                 425                 430

Ala Lys Gln Leu Lys Glu Ala Glu Lys Ile Leu Asn Asn Asn Ser Ser
        435                 440                 445

Glu Glu Lys Val Gln Ile Lys Asn Val Val Ala Asn Ile Thr Ala Ile
    450                 455                 460

Tyr Asp Glu Tyr Thr Pro Thr Gln Ile Glu Asp Ala Phe Lys Lys Val
465                 470                 475                 480

Ile Arg Lys Lys Ser Ser Val Gly Lys Ala Glu Arg Glu Ile Ser Lys
                485                 490                 495

Leu Val Ile Glu Gln Leu Thr Lys Gln Gln Arg Thr Arg Ser Asp Leu
            500                 505                 510

Cys Lys Lys Asn Thr Pro Ser Val Ile Gln Met Gln Asn Phe Leu Gly
        515                 520                 525

Ile Asp Lys Glu Asp Lys
    530
```

<210> 10
<211> 641
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG39 transposition protein sequence

<220>
<223> MG39-1-Q transposition protein

<400> 10
Met Pro Val Ile Pro Val Tyr Ile Asn Pro Tyr Pro Asp Glu Leu Met
1   5   10   15

Tyr Ser Trp Ile His Arg Leu Ala Lys Lys Asn Gly Leu Pro Ile Thr
            20   25   30

Thr Phe Ala Asp Ser Tyr Leu Asn Lys Phe Asn Ser Lys Ile Gly Ser
        35   40   45

Leu Glu Tyr Asp Ile Arg Tyr Gly Leu Leu Cys Leu Asn Glu Ser Phe
    50   55   60

Phe Ile Gln Lys Asp Leu Lys Glu Met Phe Leu Ser Met Ser Ile Phe
65   70   75   80

Pro Phe Glu Ala Met Phe Phe Ser Val Gly Gln Gln Thr Arg Tyr Val
                85   90   95

Asn Asn Met Phe Arg Lys Pro Asp Pro Leu Asn Ala Pro Val Asn Thr
            100   105   110

Met Ile Arg Glu Ile His Ile Cys Pro Gln Cys Ile Glu Asn Asp Ile
        115   120   125

Glu Ile Phe Gly Glu Pro Tyr Ile His Arg Ala His Gln Leu Ser Gly
    130   135   140

Val Cys Thr Cys His Lys His Lys Ala Pro Leu Tyr Lys Tyr Thr Gly
145   150   155   160

Ile Lys Gly His Glu Cys Glu Tyr Asp Leu Ser His Tyr Thr Glu Leu
                165   170   175

Glu Ile Lys Asp Ile Ala Ile Glu Asn Glu Tyr Thr Asp Tyr Ala Gln
            180   185   190

Ala Leu Phe Asn Ser Asn Thr Asn Cys Asn Ile Ala Asp Leu Lys His
        195   200   205

```
Met Ile Tyr Asn Lys Leu Lys Glu Leu Gly His Lys Ala Thr Asn Gly
    210                 215                 220

Tyr Glu Asp Phe Ile Ser Ala Phe Asn Asp Ser Lys Leu Val Thr Leu
225                 230                 235                 240

Phe Asp Ala Asp Leu Lys Asn Phe Leu Leu Val Ser Ile Ile Ser Thr
                245                 250                 255

Gln Ser Thr Ser Ala Arg Ser Met Leu Pro Leu Leu Met Tyr Leu Phe
            260                 265                 270

Pro Asn Val Gln Glu Ile Ile His Arg Phe Glu Asn Ala Pro Pro Ile
        275                 280                 285

Ile Gln Lys Tyr His Cys Ala Glu Cys Gly Lys Asn Phe Tyr Ala Thr
    290                 295                 300

Pro Thr Ser Leu Thr Glu Gly Trp Gly Cys Thr Tyr Cys Asp Ala Asn
305                 310                 315                 320

Lys Ser Ile Asp Glu Arg Tyr Lys Lys Leu Ile Asp Phe Ala Gly Lys
                325                 330                 335

Gly Asn Tyr Glu Pro Leu Glu Pro Phe Arg Ser Leu Asn Ile Lys Ser
            340                 345                 350

Lys Ile His His Lys Ile Cys Gly Glu Thr Ile Gln Ile Lys Pro Arg
        355                 360                 365

Lys Phe Ile Phe Asp His Val Arg Cys Ile Cys Glu Ser Leu Leu Asn
    370                 375                 380

Glu Trp Asp Val Arg Gln Lys Leu Glu Lys Phe Glu Asp Tyr Glu Phe
385                 390                 395                 400

Ile Ser Tyr Asp Ser Glu Ser Ser Lys Ile Thr Met Arg Ser Lys Ala
                405                 410                 415

Cys Gly His Val Phe Ser Cys Arg Phe His Lys Phe Ile Lys Tyr Pro
            420                 425                 430
```

```
Ser Cys Arg Val Cys Arg Pro Lys Asn Met Thr Thr Glu Leu Tyr Thr
        435                 440                 445

Glu Arg Val Tyr Asp Leu Val Gly Asp Asp Tyr Thr Val Leu Ser Glu
    450                 455                 460

Phe Val Asp Gln Arg Thr Lys Ile Ala Ile Lys His Asn Lys Cys Gly
465                 470                 475                 480

Gly Ile Gln Glu Tyr Lys Pro Ser Ala Phe Leu Asp Gly Gln Arg Cys
                485                 490                 495

Asn Ala Cys Asn Ser Leu Ile Val Lys Lys Ala Asn Asp Ser Trp Glu
            500                 505                 510

Lys Gly Tyr Ala Leu Leu Cys Glu Tyr Lys Glu Glu Cys Gly Thr Ala
        515                 520                 525

Asn Ile Pro Lys Arg Asp His Tyr Lys Gly Val Phe Leu Gly Asn Trp
    530                 535                 540

Leu Gln Ser Gln Arg Asp Lys Tyr Lys Ala Gly Lys Leu Thr Arg Ser
545                 550                 555                 560

Gln Glu Asp Ala Leu Val Ser Leu Gly Ile Thr Leu Asp Pro Leu Ala
                565                 570                 575

Ala Glu Trp Glu Arg Arg Tyr Glu Gln Tyr Lys Arg Tyr Ile Gln Gln
            580                 585                 590

Asn Asn Gly Ser Ser Asp Ile Thr Lys Arg Thr Ile Phe Glu Gly Glu
        595                 600                 605

Lys Leu Gly Val Trp Val Val Leu Gln Arg Arg Asn Tyr Asn Ile Gly
    610                 615                 620

Lys Leu Ser Glu Glu Arg Tyr Lys Lys Leu Cys Asp Ile Asn Met Lys
625                 630                 635                 640
```

Phe

<210> 11
<211> 118
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG36-5 effector tested sgRNA sequence

<220>
<221> modified_base
<222> (1)..(24)
<223> a, c, u, g, unknown or other

<400> 11
nnnnnnnnnn nnnnnnnnnn nnnnguugug auuugcuuuc agaaaaugaa gcaaaucaca 60

auaaggauua uuccguugug aaaacauuag guucccucgc ccuucugcgg gggauuuu 118

<210> 12
<211> 80
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG36-5 effector tracrRNA sequence

<400> 12
uacuaaaaug aagcaaauca caauaaggau uauuccguug ugaaaacauu agguucccuc 60

gcccuucugc gggggauuuu 80

<210> 13
<211> 46
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG39-1 effector tracrRNA sequence 1

<400> 13
uuuuuguaaa aucauuuugu aaaauuacaa uucauaauga gauaug 46

<210> 14
<211> 66
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG39-1 effector tracrRNA sequence 2

<400> 14
uuuuuguaaa aucauuuugu aaaauuacaa uucauaauga gauaugacaa uucauaauga 60

gauauu 66

<210> 15
<211> 109
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG39-1 effector tracrRNA sequence 3

<400> 15
aaaauauagg auauuaaugu cagagauauc ucauuauaau uguaauuuug uuuuguguac 60

aagcuguuaa uaucucauua ugaauuguca uaucucauua ugaauugua 109

<210> 16
<211> 88
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>

<223> MG39-1 effector tracrRNA sequence 4

<400> 16
aaaauauagg auauuaaugu cagagauauc ucauuauaau uguaauuuug uuuuguguac 60

aagcuguuaa uaucucauua ugaauugu 88

<210> 17
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> putative MG36-5 transposon end LE1

<400> 17
ttaattgaaa attgaaaatt gagaattgaa aattaa 36

<210> 18
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> putative MG36-5 transposon end LE2

<400> 18
tgagaattga aaattgagaa ttgggaattg agaatt 36

<210> 19
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> putative MG36-5 transposon end RE1

<400> 19
ttaattttca attctcaatt ttcaatta 28

<210> 20
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> putative MG39-1 transposon end LE1

<400> 20
aaaagtattg acaaattgac aaaatggtgg tataat 36

<210> 21
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> putative MG39-1 transposon end RE1

<400> 21
attataacct atatttttga ttttgtcaat agtttt 36

<210> 22
<211> 637
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-2 effector

<400> 22
Met Ser Gln Ile Thr Ile Gln Cys Gln Leu Val Ala Ser Ala Ser Thr
1 5 10 15

Arg Gln Gln Leu Trp Leu Leu Met Thr Gln Lys Asn Thr Pro Leu Ile
20 25 30

```
Asn Glu Leu Leu Gln Gln Val Gly Gln His Pro Glu Phe Glu Thr Trp
        35                  40                  45

Arg Gln Lys Gly Lys Leu Gln Ala Gly Ile Val Lys Ala Leu Cys Gln
    50                  55                  60

Pro Leu Lys Ala Asp Pro Arg Phe Met Gly Gln Pro Ala Arg Phe Tyr
65                  70                  75                  80

Ala Ser Ala Ile Ala Val Val Asp Tyr Ile Tyr Arg Ser Trp Leu Ala
                85                  90                  95

Leu Gln Lys Arg Leu Gln Tyr Gln Leu Glu Gly Gln Thr Arg Trp Tyr
            100                 105                 110

Gln Met Leu Arg Ser Asp Ala Glu Leu Ile Glu Leu Tyr Gly Gly Ser
        115                 120                 125

Leu Glu Thr Leu Arg Ser Lys Ala Ala Glu Ile Leu Ala Gln Phe Ala
    130                 135                 140

Pro Glu Ser Ala Ser Val Asn Pro Gln Pro Thr Lys Gly Lys Lys Ser
145                 150                 155                 160

Lys Lys Arg Lys Asn Ser Ser Asn Asn Pro Asn Leu Ser Thr Ala Leu
                165                 170                 175

Phe Glu Ala Tyr Arg Gln Thr Glu Asp Ile Leu Ser Ser Cys Ala Ile
            180                 185                 190

Asn Tyr Leu Leu Lys Asn Gly Cys Lys Val Ser Glu Lys Glu Glu Asp
        195                 200                 205

Pro Glu Lys Phe Ala Lys Arg Arg Arg Ser Val Glu Ile Arg Ile Glu
    210                 215                 220

Arg Leu Lys Glu Lys Leu Ala Ser Arg Met Pro Lys Gly Arg Asp Leu
225                 230                 235                 240

Thr Asp Glu Lys Trp Leu Glu Thr Leu Met Val Ala Ser Thr Thr Val
```

```
                245                 250                 255

Pro Ser Ser Glu Phe Gln Ala Lys Ser Trp Gln Asp Asn Leu Leu Arg
            260                 265                 270

Lys Ser Ser Ala Ile Pro Phe Pro Val Ala Tyr Glu Thr Asn Glu Asp
        275                 280                 285

Met Thr Trp Phe Lys Asn Ser Lys Gly Arg Ile Cys Val Lys Phe Asn
    290                 295                 300

Gly Leu Ser Glu Gln Thr Phe Glu Val Tyr Cys Asp Ser Arg Gln Leu
305                 310                 315                 320

Tyr Trp Phe Gln Arg Phe Leu Glu Asp Gln Gln Ile Lys Arg Asn Ser
                325                 330                 335

Lys Asp Gln His Ser Ser Ser Leu Phe Thr Leu Arg Ser Gly Arg Ile
            340                 345                 350

Ala Trp Ser Glu Gly Glu Gly Lys Gly Asp Pro Trp Asn Ile His Arg
        355                 360                 365

Leu Thr Leu Tyr Cys Cys Val Asp Thr Arg Leu Trp Thr Ser Glu Gly
    370                 375                 380

Thr Glu Gln Val Arg His Glu Lys Ala Asp Glu Ile Thr Gln Ile Ile
385                 390                 395                 400

Thr Lys Thr Lys Gln Lys Thr Asp Leu Asn Glu Lys Gln Gln Ala Phe
                405                 410                 415

Ile Lys Arg Lys Thr Ser Thr Leu Ala Arg Ile Asn Asn Pro Phe Pro
            420                 425                 430

Arg Pro Ser Lys Pro Val Tyr Gln Gly His Ser His Ile Leu Val Gly
        435                 440                 445

Val Ser Leu Gly Leu Asp Lys Pro Ala Thr Leu Ala Val Ile Asp Ala
    450                 455                 460
```

```
Ile Ala Asn Lys Val Ile Thr Tyr Arg Ser Ile Arg Gln Leu Leu Gly
465                 470                 475                 480

Asp Asn Tyr Gln Leu Leu Asn Arg Gln Arg Gln Gln Gln His Gln Asn
                485                 490                 495

Ala His Lys Arg Gln Thr Ala Gln Arg Gln Asp Ala Pro His Gln Phe
            500                 505                 510

Ser Glu Ser Glu Leu Gly Gln Tyr Ile Asp Arg Leu Leu Ala Lys Ala
        515                 520                 525

Ile Val Ala Val Ala Lys Thr Tyr Gln Ala Gly Ser Ile Val Leu Pro
    530                 535                 540

Gln Leu Gly Asp Val Arg Glu Ser Ile Glu Ser Glu Ile Lys Ala Arg
545                 550                 555                 560

Ala Glu Gln Lys Cys Pro Asp Leu Val Glu Val Gln Lys Gln Tyr Ala
                565                 570                 575

Lys Gln Tyr Arg Ser Ser Ile His Arg Trp Ser Tyr Ala Arg Leu Ile
            580                 585                 590

Asp Ser Ile Lys Ser Gln Ala Ser Gln Val Gly Ile Phe Ile Glu Glu
        595                 600                 605

Gly Lys Gln Pro Val Arg Gly Ser Pro Gln Asp Lys Ala Lys Asp Leu
    610                 615                 620

Ala Ile Ala Val Tyr His Ser Arg Leu Asn Thr Lys Ser
625                 630                 635
```

<210> 23
<211> 560
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-2-B transposition protein

<400> 23
Met Gln Asp Asp Arg Glu Ser Glu Val Pro Ile Pro Ala Glu Val Asn
1   5   10   15

Glu Ile Val Thr Asp Phe Ser Glu Asp Ala Lys Leu Met Gln Glu Val
            20   25   30

Ile Gln Ser Leu Leu Glu Pro Cys Asp Arg Ile Thr Tyr Gly Gln Arg
        35   40   45

Gln Arg Glu Ala Ala Ala Lys Leu Gly Lys Ser Val Arg Thr Ile Arg
    50   55   60

Arg Leu Val Lys Lys Trp Glu Thr Glu Gly Leu Ser Ala Leu Gln Pro
65   70   75   80

Asn Lys Arg Thr Asp Lys Gly Lys His Arg Ile Asp Gln Glu Trp Gln
                85   90   95

Glu Phe Ile Ile Lys Thr Tyr Lys Glu Gly Asn Lys Gly Ser Lys Arg
            100   105   110

Ile Thr Pro Gln Gln Val Ala Val Arg Val Ala Ala Lys Ala Ala Asp
        115   120   125

Leu Gly Gln Glu Lys Tyr Pro Ser Tyr Arg Thr Val Tyr Arg Val Leu
    130   135   140

Gln Pro Ile Ile Glu Lys Gln Glu Lys Thr Gln Ser Val Arg Ser Arg
145   150   155   160

Gly Trp Arg Gly Ser Arg Leu Ser Val Lys Thr Arg Asp Gly Gln Asp
                165   170   175

Leu Ser Val Glu Tyr Ser Asn His Val Trp Gln Cys Asp His Thr Arg
            180   185   190

Ala Asp Ile Leu Leu Val Asp Gln Asp Gly Gln Leu Leu Gly Arg Pro
        195   200   205

```
Trp Leu Thr Thr Val Ile Asp Thr Tyr Ser Arg Cys Ile Ile Gly Ile
    210                 215                 220

Asn Leu Gly Tyr Asp Ala Pro Ser Ala Tyr Val Val Ala Leu Ala Leu
225                 230                 235                 240

Arg His Ala Ile Leu Pro Lys Gln Tyr Ser Ser Glu Tyr Lys Leu His
                245                 250                 255

Cys Gln Trp Gly Thr Tyr Gly Lys Pro Glu His Phe Tyr Thr Asp Gly
            260                 265                 270

Gly Lys Asp Phe Arg Ser Asn His Leu Gln Gln Ile Gly Val Gln Leu
        275                 280                 285

Gly Phe Val Cys His Leu Arg Asp Arg Pro Ser Glu Gly Gly Ile Val
    290                 295                 300

Glu Arg Pro Phe Gly Thr Phe Asn Thr Asp Leu Phe Ser Thr Leu Pro
305                 310                 315                 320

Gly Tyr Thr Gly Ser Asn Val Gln Glu Arg Pro Glu Glu Ala Glu Lys
                325                 330                 335

Glu Ala Ser Leu Thr Leu Arg Glu Leu Glu Gln Leu Leu Val Arg Tyr
            340                 345                 350

Ile Val Asp Lys Tyr Asn Gln Ser Ile Asp Ala Arg Met Gly Asp Gln
        355                 360                 365

Thr Arg Phe Gln Arg Trp Glu Ala Gly Leu Ile Ala Ala Pro Asn Leu
    370                 375                 380

Ile Ser Glu Arg Asp Leu Asp Ile Cys Leu Met Lys Gln Thr Arg Arg
385                 390                 395                 400

Thr Ile Tyr Arg Gly Gly Tyr Leu Gln Phe Glu Asn Leu Thr Tyr Arg
                405                 410                 415
```

Gly Asp Tyr Leu Glu Gly Tyr Ser Gly Glu Ser Val Val Leu Arg Tyr
420 425 430

Asp Pro Lys Asp Ile Thr Thr Ile Leu Val Tyr Arg Lys Glu Gly Asp
435 440 445

Lys Glu Val Phe Leu Ala Arg Ala Tyr Glu Gln Asp Leu Gln Thr Glu
450 455 460

Gln Leu Ser Phe Asp Glu Val Lys Ala Ile Ser Arg Lys Leu Arg Glu
465 470 475 480

Ala Lys Lys Ala Val Asn Ser Arg Ser Ile Leu Ala Glu Val Arg Asp
485 490 495

Arg Glu Thr Phe Leu Thr Gln Lys Lys Ser Lys Lys Gln Arg Gln Lys
500 505 510

Ser Glu Gln Ala Gln Ile His Arg Lys Lys Glu Leu Phe Pro Ile Glu
515 520 525

Ala Glu Ala Thr Glu Phe Glu Ser Pro Val Asn Glu Leu Glu Thr Glu
530 535 540

Ile Ile Glu Val Phe Asp Tyr Glu Gln Met Arg Glu Asp Tyr Gly Phe
545 550 555 560

<210> 24
<211> 287
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-2-C transposition protein

<400> 24
Met Met Asn Gln Glu Lys Glu Ala Lys Ala Ile Ala Gln Lys Leu Gly
1 5 10 15

Asn Ile Pro Leu Asn Asp Glu Lys Ile Gln Ala Glu Ile Gln Arg Leu

```
            20                  25                  30

Asn Arg Lys Asn Phe Val Pro Leu Glu Gln Val Lys Ala Leu His Asp
        35                  40                  45

Trp Leu Glu Ser Lys Arg Gln Ala Arg Gln Cys Cys Arg Val Ile Gly
    50                  55                  60

Glu Ser Arg Thr Gly Lys Thr Met Ala Cys Asn Ala Tyr Arg Leu Arg
65                  70                  75                  80

Asn Lys Ala Ile Gln Ser Leu Gly Gln Pro Pro Thr Val Pro Val Val
                85                  90                  95

Tyr Ile Gln Ile Pro Gln Glu Cys Thr Pro Lys Glu Leu Phe Ser Val
            100                 105                 110

Leu Leu Glu His Leu Asn His Gln Met Thr Lys Gly Thr Thr Ala Glu
        115                 120                 125

Met Arg Asn Arg Thr Leu Arg Val Leu Lys Ala Cys Arg Val Glu Met
    130                 135                 140

Leu Ile Ile Asp Glu Ala Asp Arg Leu Lys Pro Lys Thr Phe Ala Asp
145                 150                 155                 160

Val Arg Asp Ile Phe Asp Asn Leu Glu Ile Ser Val Val Leu Val Gly
                165                 170                 175

Thr Val Arg Leu Glu Lys Val Met Thr Asp Asp Glu Gln Val Cys Asn
            180                 185                 190

Arg Phe Ser Ala Cys Tyr Arg Tyr Gly Lys Leu Ser Val Glu Glu Phe
        195                 200                 205

Lys Lys Thr Val Asn Ile Trp Glu Asn Gln Val Leu Lys Leu Pro Val
    210                 215                 220

Ser Ser Asn Leu Thr Gln Ser Lys Met Leu Asp Ile Leu Lys Asp Lys
225                 230                 235                 240
```

```
Thr Gln Cys Tyr Ile Gly Leu Met Asp Met Ile Leu Arg Asp Ala Ala
                245                 250                 255


Ile Arg Ala Leu Asn Lys Gly Met Pro Lys Ile Asp Leu Glu Thr Leu
            260                 265                 270


Lys Glu Val Thr Asn Glu Tyr Thr Ala Pro Pro Lys Arg Ser Lys
        275                 280                 285
```

<210> 25
<211> 168
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-2-Q transposition protein

<400> 25

```
Met Lys Ala Ser Asp Ile Gln Pro Trp Leu Phe Arg Val Glu Pro Phe
1               5                   10                  15


Glu Gly Glu Ser Ile Ser His Phe Leu Gly Arg Phe Gln Arg Ala Asn
            20                  25                  30


Glu Leu Thr Pro Ser Gly Leu Ala Lys Ala Ala Gly Leu Gly Gly Ala
        35                  40                  45


Ile Ala Arg Trp Glu Lys Phe Arg Phe Asn Pro Pro Pro Ser Pro Gln
    50                  55                  60


Gln Leu Glu Ala Leu Ala Leu Val Val Glu Val Glu Ala Glu Arg Leu
65                  70                  75                  80


Val Gln Met Leu Pro Pro Ala Gly Val Gly Met Lys Met Glu Pro Ile
                85                  90                  95


Arg Leu Cys Gly Ala Cys Tyr Ala Glu Val Pro Cys His Lys Ile Glu
            100                 105                 110
```

Trp Gln Phe Lys Thr Thr Gln Gly Cys Asp Arg His Gln Leu Ser Leu
        115   120   125

Leu Ser Glu Cys Pro Asn Cys Gly Ala Arg Tyr Lys Val Pro Ala Leu
    130   135   140

Trp Pro Asp Gly Trp Cys Ser Arg Cys Phe Leu Pro Phe Ala Asp Met
145   150   155   160

Val Lys Trp Gln Lys Leu Ile Val
                165

<210> 26
<211> 637
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
      MG64 effector sequence

<220>
<223> MG64-4 effector

<400> 26
Met Ser Gln Ile Thr Ile Gln Cys Arg Leu Val Cys Cys Glu Ser Thr
1   5   10   15

Arg Ala Ser Leu Trp Lys Leu Met Ala Glu Lys Asn Thr Pro Leu Thr
            20   25   30

Asn Glu Leu Leu His Gln Val Gly His His Cys Glu Phe Glu Ser Trp
        35   40   45

Arg Lys Lys Gly Lys Ile Pro Ser Gly Ile Val Lys Glu Leu Cys Gln
    50   55   60

Pro Leu Lys Thr Asp Pro Arg Phe Thr Gly Gln Pro Gly Arg Phe Tyr
65   70   75   80

Thr Ser Ala Ile Ala Leu Val Ser Tyr Ile Tyr Lys Ser Trp Leu Val
                85   90   95

Leu Met Lys Arg Ser His Tyr Lys Leu Glu Gly Lys Thr Arg Trp Leu

```
            100                 105                 110

Glu Met Leu Ser Ser Asp Ala Glu Leu Val Glu Arg Cys Gly Ile Ser
        115                 120                 125

Val Asp Gly Ile Arg Ala Arg Ala Ser Glu Ile Leu Ala Gln Leu Ala
    130                 135                 140

Leu Pro Asp Ser Ala Asp Thr Gly Gln Val Lys Gly Arg Lys Ala Lys
145                 150                 155                 160

Lys Ser Lys Lys Phe Gln Gln Gln Gln Asp Arg Asp Arg Asn Leu Ser
                165                 170                 175

Arg Thr Leu Phe Glu Ala Tyr Arg Asp Thr Glu Asp Ile Leu Thr Arg
            180                 185                 190

Cys Ala Ile Cys Tyr Leu Leu Lys Asn Gly Cys Gln Val Ser Asp Arg
        195                 200                 205

Glu Glu Asn Ser Glu Lys Phe Thr Gln Arg Arg Arg Gln Val Glu Ile
    210                 215                 220

Gln Ile Glu Arg Leu Lys Glu Lys Leu Ala Ala Arg Ile Pro Lys Gly
225                 230                 235                 240

Arg Asp Leu Thr Asp Ala Lys Trp Leu Glu Thr Leu Phe Leu Ala Thr
                245                 250                 255

Ser His Ala Pro Glu Ser Glu Ala Glu Ala Lys Leu Trp Gln Asp Cys
            260                 265                 270

Leu Leu Arg Gln Pro Ser Ser Val Pro Phe Pro Val Ala Tyr Glu Thr
        275                 280                 285

Ser Glu Asp Met Thr Trp Phe Lys Asn His Lys Gly Arg Ile Ser Val
    290                 295                 300

Lys Phe Asn Gly Leu Ser Glu His Thr Phe Glu Val Tyr Cys Asp Ser
305                 310                 315                 320
```

```
Arg Gln Leu His Trp Phe Glu Arg Phe Leu Glu Asp Gln Gln Ile Lys
                325                 330                 335

Arg Asn Asn Gln Asn Gln Tyr Ser Ser Ser Leu Phe Thr Leu Arg Ser
            340                 345                 350

Ser Arg Ile Ala Trp Gln Glu Gly Lys Gly Lys Gly Glu Pro Trp Asn
        355                 360                 365

Phe Asn His Leu Thr Leu Tyr Cys Ser Val Asp Thr Arg Leu Trp Thr
    370                 375                 380

Ala Glu Gly Thr Gln Leu Val Ile Glu Glu Lys Val Glu Glu Ile Val
385                 390                 395                 400

Lys Asn Ile Thr Asn Thr Lys Asp Lys Gly Asn Leu Asn Glu Lys Gln
                405                 410                 415

Leu Ala His Ile Gln Arg Lys Asn Ser Thr Leu Ala Arg Ile Thr Asn
            420                 425                 430

Pro Phe Pro Arg Pro Ser Lys Ser Ile Tyr Arg Gly Gln Ser His Ile
        435                 440                 445

Leu Val Gly Val Ser Leu Gly Leu Glu Lys Pro Ala Thr Val Ala Val
    450                 455                 460

Val Asp Ala Thr Thr Ser Asn Val Ile Ala Tyr Arg Ser Ile Lys Gln
465                 470                 475                 480

Leu Leu Gly Asp Asn Tyr Arg Leu Leu Asn Arg Gln Arg Gln Gln Lys
                485                 490                 495

Gln Val Leu Ser His Gln Arg His Leu Ala Gln Arg Leu Asp Ala Leu
            500                 505                 510

Asn Ser Leu Gly Glu Ser Glu Leu Gly Gln Tyr Val Asp Arg Leu Leu
        515                 520                 525

Ala Lys Glu Ile Val Ala Ile Ala Gln Thr Tyr Cys Ala Gly Ser Ile
```

530 535 540

Val Leu Pro Ser Leu Gly Asp Met Arg Glu Gln Val Gln Ser Glu Ile
545 550 555 560

Gln Ala Lys Ala Glu Gln Lys Ser Asp Leu Ile Glu Val Gln Gln Lys
565 570 575

Tyr Ala Lys Gln Tyr Arg Ala Ser Val His Gln Trp Ser Tyr Gly Arg
580 585 590

Leu Ile Ser Ser Ile Gln Ser Gln Ala Lys Lys Ala Gly Ile Ala Ile
595 600 605

Glu Glu Ala Lys Gln Pro Ile Arg Gly Ser Pro Gln Gln Lys Ala Lys
610 615 620

Glu Leu Val Ile Ala Ala Tyr Asn Ala Arg Lys Thr Ser
625 630 635

<210> 27
<211> 563
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-4-B transposition protein

<400> 27
Met Ser Lys Leu Ser Glu Asp Asn His Ile Asp Asn Gln Lys Pro Ala
1 5 10 15

Val Gly Glu Ile Val Pro Glu Ile Ala Asp Asp Asn Lys Gln Leu Leu
20 25 30

Glu Val Ile Gln Lys Leu Leu Glu Pro Cys Asp Arg Ile Thr Tyr Gly
35 40 45

Gln Arg Gln Arg Glu Ala Ala Ala Gln Leu Gly Lys Ser Val Arg Thr
50 55 60

```
Val Arg Arg Leu Val Lys Lys Trp Glu Glu Glu Gly Leu Ala Ala Leu
65                  70                  75                  80

Ser Gln Thr Thr Arg Ala Asp Lys Gly Lys His Arg Ile Glu Gln Asp
                85                  90                  95

Trp Gln Asp Phe Ile Ile Lys Thr Tyr Lys Glu Gly Asn Lys Gly Ser
            100                 105                 110

Lys Arg Ile Thr Pro Lys Gln Val Ala Val Arg Val Gln Ala Lys Ala
        115                 120                 125

Ala Glu Leu Gly Gln Asp Arg Tyr Pro Ser Tyr Arg Thr Val Tyr Arg
    130                 135                 140

Val Leu Gln Pro Ile Ile Glu Arg Gln Glu Gln Gln Ala Ser Val Arg
145                 150                 155                 160

Ser Arg Gly Trp Arg Gly Ser Arg Leu Ser Val Lys Thr Arg Asp Gly
                165                 170                 175

Lys Asp Leu Ser Val Glu Tyr Ser Asn His Val Trp Gln Cys Asp His
            180                 185                 190

Thr Arg Val Asp Val Leu Leu Val Asp Arg Asn Gly Ala Ile Leu Ser
        195                 200                 205

Arg Pro Trp Leu Thr Thr Val Val Asp Thr Tyr Ser Arg Cys Ile Met
    210                 215                 220

Gly Phe Asn Leu Gly Tyr Asp Ala Pro Ser Ser Gln Val Val Ala Leu
225                 230                 235                 240

Ala Leu Arg His Ala Ile Leu Pro Lys Gln Tyr Asp Trp Glu Tyr Gln
                245                 250                 255

Leu Asp Cys Asp Trp Gly Thr Tyr Gly Lys Pro Glu His Phe Tyr Thr
            260                 265                 270
```

```
Asp Gly Gly Lys Asp Phe Arg Ser Asn His Leu Gln Gln Ile Gly Val
        275                 280                 285

Gln Leu Gly Phe Ala Ser His Leu Arg Asp Arg Pro Ser Glu Gly Gly
    290                 295                 300

Ile Val Glu Arg Pro Phe Gly Thr Phe Asn Thr Glu Phe Phe Ser Thr
305                 310                 315                 320

Leu Pro Gly Tyr Thr Gly Ser Asn Val Gln Glu Arg Pro Glu Gln Ala
                325                 330                 335

Glu Lys Glu Ala Arg Leu Thr Leu Arg Glu Leu Glu Arg Leu Leu Val
            340                 345                 350

Arg Tyr Ile Val Asp Lys Tyr Asn Gln Ser Ile Asp Ala Arg Leu Gly
        355                 360                 365

Asp Gln Thr Arg Phe Gln Arg Trp Glu Ala Gly Leu Ile Ala Ala Pro
    370                 375                 380

Asn Pro Ile Ala Glu Arg Asp Leu Asp Ile Cys Leu Met Lys Gln Thr
385                 390                 395                 400

Arg Arg Ser Ile Tyr Arg Gly Gly Tyr Leu Gln Phe Glu Asn Leu Thr
                405                 410                 415

Tyr Arg Gly Glu Asn Leu Ala Gly Tyr Ala Gly Glu Ser Val Val Leu
            420                 425                 430

Arg Tyr Asp Pro Arg Asp Ile Thr Thr Val Leu Val Tyr Arg Gln Glu
        435                 440                 445

Ala Gly Lys Glu Val Phe Leu Ala Arg Ala Phe Ala Gln Asp Leu Glu
    450                 455                 460

Thr Glu Gln Met Ser Leu Asp Glu Ala Lys Ala Ser Ser Arg Lys Leu
465                 470                 475                 480

Arg Glu Thr Gly Lys Thr Ile Ser Asn Arg Ser Ile Leu Ala Glu Val
                485                 490                 495
```

Arg Asp Arg Glu Thr Phe Leu Thr Gln Lys Lys Thr Lys Lys Glu Arg
500 505 510

Gln Lys Ala Glu Gln Ala Glu Val Lys Arg Ala Lys Gln Pro Phe Ser
515 520 525

Val Glu Arg Glu Glu Glu Ile Glu Ala Ala Ser Ile Pro Asn Gln Pro
530 535 540

Glu Pro Glu Met Pro Asp Val Phe Asp Tyr Glu Gln Met Arg Glu Asp
545 550 555 560

Tyr Gly Phe

<210> 28
<211> 278
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-4-C transposition protein

<400> 28
Met Ser Leu Lys Asp Ala Gln Ala Ile Ala Gln Gln Leu Gly Asn Ile
1 5 10 15

Gln Pro Asn Asp Ala Arg Leu Gln Thr Glu Ile Gln Arg Leu Asn Gly
20 25 30

Lys Ser Phe Val Pro Leu Glu Gln Val Lys Ile Leu His Asp Trp Leu
35 40 45

Asp Gly Lys Arg Gln Ala Arg Gln Gly Cys Arg Val Val Gly Glu Ser
50 55 60

Arg Thr Gly Lys Thr Ile Ala Cys Asp Ala Tyr Arg Leu Arg His Lys
65 70 75 80

```
Pro Ile Gln Glu Leu Gly Lys Pro Pro Ile Val Pro Val Ala Tyr Ile
                85                  90                  95

Leu Val Pro Pro Asp Cys Gly Ser Lys Asp Leu Phe Gly Leu Ile Val
            100                 105                 110

Glu His Leu Lys Tyr Gln Met Thr Lys Gly Thr Val Ala Glu Ile Arg
        115                 120                 125

Glu Arg Thr Arg Arg Val Leu Lys Gly Cys Gly Val Glu Met Leu Ile
    130                 135                 140

Ile Asp Glu Ala Asp Arg Leu Lys Pro Asn Thr Phe Arg Asp Val Arg
145                 150                 155                 160

Asp Ile Gly Glu Asn Leu Gly Ile Thr Val Val Leu Val Gly Thr Asp
                165                 170                 175

Arg Leu Asp Ala Val Ile Lys Lys Asp Pro Gln Val Tyr Asn Arg Phe
            180                 185                 190

Arg Ala Cys His Arg Phe Gly Asn Leu Ser Gly Asp Asp Phe Lys Arg
        195                 200                 205

Thr Val Glu Ile Trp Glu Lys Lys Val Leu Gln Leu Pro Val Ala Ser
    210                 215                 220

Asn Leu Ser Gly Lys Thr Met Leu Lys Thr Leu Gly Glu Ala Thr Gly
225                 230                 235                 240

Gly Tyr Ile Gly Leu Leu Asp Met Ile Leu Arg Glu Ala Ala Ile Arg
                245                 250                 255

Ala Leu Lys Lys Gly Leu Gln Lys Ile Asp Leu Glu Thr Leu Lys Glu
            260                 265                 270

Val Ala Gly Glu Tyr Arg
        275
```

<210> 29

<211> 167
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-4-Q transposition protein

<400> 29
Met Glu Ala Ser Asp Ile Gln Pro Trp Leu Phe Arg Ala Glu Pro Tyr
1 5 10 15

Glu Gly Glu Ser Ile Ser His Tyr Leu Gly Arg Phe Arg Arg Ala Asn
20 25 30

Asp Leu Thr Pro Thr Gly Leu Gly Lys Ala Thr Gly Leu Gly Gly Ala
35 40 45

Ile Ala Arg Trp Glu Lys Phe Arg Phe Asn Pro Pro Pro Ser Arg Lys
50 55 60

Gln Leu Glu Ala Leu Ala Val Val Val Gly Val Glu Ala Asp Arg Leu
65 70 75 80

Ala Gln Met Ile Pro Pro Ala Gly Val Gly Met Lys Leu Glu Pro Ile
85 90 95

Arg Leu Cys Ala Ala Cys Tyr Ala Gln Ala Pro Tyr His Arg Ile Glu
100 105 110

Trp Gln Phe Lys Val Thr Ala Gly Cys Asp His His Gln Leu Arg Leu
115 120 125

Leu Ser Glu Cys Pro Asn Cys Lys Ala Arg Phe Lys Ile Pro Ala Leu
130 135 140

Trp Val Asp Gly Trp Cys Gln Arg Cys Phe Thr Lys Phe Glu Glu Met
145 150 155 160

Thr Ser Asn Gln Ala Ala Leu
165

<210> 30
<211> 632
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-6 effector

<400> 30
Met Ser Gln Ile Thr Ile Gln Cys Arg Leu Ile Ala Lys Glu Ser Ser
1 5 10 15

Arg Gln Ala Leu Trp Arg Leu Met Ala Glu Leu Asn Thr Pro Leu Ile
20 25 30

Asn Asp Ile Leu Asn Gln Ile Ala Asn His Pro Asp Phe Glu Thr Trp
35 40 45

Arg Glu Lys Gly Lys Leu Pro Ala Gly Ile Val Lys Gln Leu Ser Asp
50 55 60

Ser Leu Lys Thr Asp Pro Arg Tyr Ile Gly Gln Pro Gly Arg Phe Tyr
65 70 75 80

Thr Ser Ala Ile Thr Leu Ile Ser Tyr Ile Tyr Lys Ser Trp Phe Lys
85 90 95

Val Gln Gln Arg Leu Gln Gln Arg Leu Val Gly Gln Thr Arg Trp Leu
100 105 110

Gly Ile Leu Lys Ser Asp Glu Glu Leu Val Ala Glu Ser Asp Arg Thr
115 120 125

Leu Glu Glu Ile Arg Ala Lys Ala Ile Gly Leu Leu Ala Ser Leu Thr
130 135 140

Pro Glu Asn Pro Ser Pro Glu Pro Lys Pro Ala Lys Lys Thr Lys Lys
145 150 155 160

```
Ala Lys Thr Ser Thr Asn Lys Pro Leu Leu Gly Ile Leu Phe Asp Asn
                165                 170                 175

Tyr Glu Lys Thr Glu Asp Ile Leu Thr His Ala Ala Ile Cys Tyr Leu
            180                 185                 190

Leu Lys Asn Gly Cys Lys Ile Pro Thr Lys Leu Glu Glu Pro Gln Glu
        195                 200                 205

Phe Ala Lys Lys Arg Arg Lys Ala Glu Ile Lys Val Glu Arg Leu Gln
    210                 215                 220

Glu Gln Leu Asn Ser Arg Lys Pro Lys Gly Arg Asp Leu Thr Gly Glu
225                 230                 235                 240

Lys Trp Leu Gln Thr Leu Ile Thr Ala Ser Thr Thr Ala Pro Glu Asn
                245                 250                 255

Glu Ala Gln Ala Lys Ser Trp Gln Asn Ile Leu Leu Thr Lys Ser Lys
            260                 265                 270

Ser Ile Pro Phe Pro Val Ala Tyr Glu Thr Asn Glu Asp Leu Thr Trp
        275                 280                 285

Ser Lys Asn Asp Lys Gly Arg Leu Cys Val His Phe Asn Gly Leu Gly
    290                 295                 300

Glu His Glu Phe Glu Ile Tyr Cys Asp Gln Arg Gln Leu Lys Trp Leu
305                 310                 315                 320

Gln Arg Phe Tyr Glu Asp Gln Glu Thr Lys Arg Ala Ser Lys Asp Gln
                325                 330                 335

His Ser Ser Ala Leu Phe Thr Leu Arg Ser Gly Arg Ile Gly Trp Gln
            340                 345                 350

Glu Gly Lys Gly Lys Gly Glu Pro Trp Asn Ile His Arg Leu Asn Leu
        355                 360                 365

Phe Cys Thr Ile Asp Thr Arg Phe Trp Thr Ala Glu Gly Thr Glu Gln
```

```
    370                 375                 380

Val Arg Gln Glu Lys Ala Thr Glu Ile Ala Gln Thr Ile Thr Lys Met
385                 390                 395                 400

Glu Gln Lys Gly Asp Leu Asn Asp Lys Gln Gln Ala Phe Ile His Arg
                405                 410                 415

Lys His Ser Thr Leu Ala Arg Ile Asn Asn Pro Phe Pro Arg Pro Ser
            420                 425                 430

Gln Pro Leu Tyr Gln Gly Lys Ser His Ile Leu Ile Gly Ile Ala Met
        435                 440                 445

Gly Leu Glu Asn Pro Ala Thr Ala Ala Ile Ile Asn Gly Thr Thr Gly
    450                 455                 460

Glu Ala Leu Ala Tyr Arg Ser Ile Lys Gln Leu Leu Gly Asp Asn Tyr
465                 470                 475                 480

Gln Leu Leu Thr Arg Gln Gln Lys Gln Lys Gln Arg Gln Ser His Gln
                485                 490                 495

Arg His Lys Ala Gln Lys Asn Ala Ala Pro Asn Gln Phe Gly Glu Ser
            500                 505                 510

Glu Leu Gly Glu Tyr Val Asp Arg Leu Leu Ala Lys Ala Ile Val Ala
        515                 520                 525

Leu Ala Lys Thr Tyr Gln Ala Gly Ser Ile Val Val Pro Glu Leu Gly
    530                 535                 540

Asp Met Arg Glu Leu Val Gln Ala Glu Val Lys Ala Lys Ala Glu Ala
545                 550                 555                 560

Lys Ile Pro Gly Cys Ile Glu Ala Gln Glu Lys Tyr Ala Lys Gln Tyr
                565                 570                 575

Arg Val Asn Thr His Gln Trp Ser Tyr Gly Arg Leu Ile Asp Asn Ile
            580                 585                 590
```

Gln Ala Gln Ala Ser Lys Ile Gly Ile Val Ile Glu Gln Gly Gln Gln
595 600 605

Pro Ile Arg Gly Ser Pro Gln Glu Lys Ala Lys Glu Met Ala Leu Leu
610 615 620

Ala Tyr His Ser Arg Ser Lys Ser
625 630

<210> 31
<211> 578
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-6-B transposition protein

<400> 31
Met Lys Lys Leu Phe Ala Gln Asp Val Asn Ile Asp Thr Glu Val Ile
1 5 10 15

Ser Asn Gln Ile Pro Thr Ser Asp Pro Ser Gln Ser Asn Leu Ile Ala
20 25 30

Ser Glu Leu Pro Glu Glu Ala Arg Pro Lys Leu Glu Val Ile Gln Ser
35 40 45

Leu Leu Glu Pro Cys Asp Arg Val Thr Tyr Gly Glu Arg Leu Arg Glu
50 55 60

Gly Ala Glu Lys Leu Gly Leu Ser Val Arg Ser Val Gln Arg Leu Phe
65 70 75 80

Lys Lys Tyr Gln Glu Lys Gly Leu Ile Ala Leu Leu Ser Gly Ser Arg
85 90 95

Thr Asp Lys Gly Glu His Arg Ile Ser Glu Leu Trp Gln Asn Phe Ile
100 105 110

```
Val Lys Thr Tyr Gln Glu Gly Asn Lys Gly Ser Lys Arg Met Ser Pro
        115                 120                 125

Lys Gln Val Ala Leu Lys Val Gln Ala Lys Ala Gly Ala Ile Ala Asp
    130                 135                 140

Asp Asn Pro Pro Ser Tyr Arg Thr Val Leu Arg Val Leu Lys Pro Ile
145                 150                 155                 160

Leu Glu Lys Gln Glu Lys Ala Lys Ser Ile Arg Ser Pro Gly Trp Arg
                165                 170                 175

Gly Ser Thr Leu Ser Val Lys Thr Arg Asp Gly Asp Asp Leu Asp Ile
            180                 185                 190

Ser Tyr Ser Asn Gln Val Trp Gln Cys Asp His Thr Arg Ala Asp Val
        195                 200                 205

Leu Leu Val Asp Gln His Gly Lys Leu Leu Val Arg Pro Trp Leu Thr
    210                 215                 220

Thr Val Ile Asp Ser Tyr Ser Arg Cys Ile Met Gly Ile Asn Leu Gly
225                 230                 235                 240

Phe Asp Ala Pro Ser Ser Gln Val Val Ala Leu Ala Leu Arg His Ala
                245                 250                 255

Ile Leu Pro Lys Arg Tyr Gly Thr Glu Tyr Lys Leu Asn Cys Asp Trp
            260                 265                 270

Gly Thr Tyr Gly Thr Pro Glu Tyr Leu Phe Thr Asp Gly Gly Lys Asp
        275                 280                 285

Phe Arg Ser Asn His Leu Ala Glu Ile Gly Leu Gln Leu Gly Phe Val
    290                 295                 300

Cys Lys Leu Arg Asp Arg Pro Ser Glu Gly Gly Ile Val Glu Arg Pro
305                 310                 315                 320

Phe Lys Thr Leu Asn Gln Ser Leu Phe Ser Thr Leu Pro Gly Tyr Thr
                325                 330                 335
```

```
Gly Ser Asn Val Gln Glu Arg Pro Glu Asp Ala Glu Lys Asp Ala Gln
            340                 345                 350

Leu Thr Leu Arg Asp Leu Glu Gln Leu Ile Val Arg Phe Ile Val Asp
        355                 360                 365

Arg Tyr Asn Gln Ser Ile Asp Ala Arg Met Gly Asp Gln Thr Arg Tyr
    370                 375                 380

Gln Arg Trp Glu Ala Gly Leu Gln Lys Glu Pro Asp Val Ile Ser Glu
385                 390                 395                 400

Arg Asp Leu Asp Ile Cys Leu Met Lys Met Ser Arg Arg Thr Val Gln
                405                 410                 415

Arg Gly Gly His Leu Gln Phe Glu Asn Val Met Tyr Leu Gly Glu Tyr
            420                 425                 430

Leu Ala Gly Tyr Ala Gly Glu Val Val Ser Phe Arg Tyr Asp Pro Arg
        435                 440                 445

Asp Ile Thr Thr Ile Trp Val Tyr Arg Gln Glu Asn Asp Arg Glu Val
    450                 455                 460

Phe Leu Thr Arg Ala His Ala Gln Gly Leu Glu Thr Glu Gln Leu Ser
465                 470                 475                 480

Val Asp Asp Ala Lys Ala Ser Ala Lys Arg Leu Arg Ala Ala Gly Lys
                485                 490                 495

Thr Ile Ser Asn Gln Ser Ile Leu Gln Glu Thr Ile Glu Arg Glu Val
            500                 505                 510

Leu Ala Glu Arg Thr Lys Ser Arg Lys His Arg Gln Lys Glu Glu Gln
        515                 520                 525

Ser Tyr Lys Arg Ser Pro Ser Ala Ala Val Met Val Glu Val Glu Ser
    530                 535                 540
```

Glu Gln Leu Glu Ile Glu Ser Ser Asn Glu Ala Asn Ala Asn Ser Val
545 550 555 560

Ser Ala Asp Ile Glu Val Trp Asp Tyr Asp Glu Met Arg Glu Gly Leu
565 570 575

Gly Trp

<210> 32
<211> 284
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-6-C transposition protein

<400> 32
Met Ile Lys Glu Asn Ser Ser Gln Glu Lys Pro Ala Ser Gln Ile Ala
1 5 10 15

Glu Glu Leu Gly Asp Phe Lys Val Asp Ser Gln Leu Leu Gln Ile Glu
20 25 30

Ile Ala Arg Leu Asn Lys Lys Ser Ile Val Pro Leu Glu His Ile Lys
35 40 45

Asp Leu His Asp Trp Leu Asp Glu Lys Arg Lys Ala Arg Gln Ser Cys
50 55 60

Arg Leu Val Gly Glu Ser Arg Thr Gly Lys Thr Val Ala Cys Glu Ala
65 70 75 80

Tyr Thr Phe Arg Asn Lys Pro Lys Gln Glu Gly Lys Gln Ala Pro Thr
85 90 95

Val Pro Val Val Tyr Ile Met Pro Pro Ala Lys Cys Gly Ala Lys Glu
100 105 110

Leu Phe Arg Glu Ile Ile Glu Tyr Leu Lys Tyr Arg Ala Val Arg Gly

```
        115                 120                 125

Thr Val Ala Asp Phe Arg Ser Arg Ala Met Glu Val Leu Lys Gly Cys
    130                 135                 140

Glu Val Glu Met Ile Ile Ile Asp Glu Ala Asp Arg Leu Lys Pro Glu
145                 150                 155                 160

Thr Phe Ser Asp Val Arg Asp Ile Asn Asp Lys Leu Gly Ile Ala Val
                165                 170                 175

Val Leu Val Gly Thr Asp Arg Leu Asp Ala Val Ile Lys Arg Asp Glu
            180                 185                 190

Gln Val Tyr Asn Arg Phe Arg Ala Ser Arg Arg Phe Gly Lys Leu Thr
        195                 200                 205

Gly Glu Asp Phe Lys Arg Thr Val Glu Ile Trp Glu Asp Lys Val Leu
    210                 215                 220

Lys Met Pro Val Ala Ser Asn Leu Thr Asn Lys Glu Met Leu Lys Ile
225                 230                 235                 240

Leu Leu Lys Ala Thr Glu Gly Tyr Ile Gly Arg Leu Asp Glu Ile Leu
                245                 250                 255

Arg Glu Ala Ala Ile Lys Ser Leu Ser Arg Gly Phe Arg Lys Val Glu
            260                 265                 270

Lys Ala Val Leu Gln Glu Val Ala Arg Glu Tyr Ser
        275                 280


<210> 33
<211> 167
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
      MG64 transposition protein sequence

<220>
<223> MG64-6-Q transposition protein
```

<400> 33
Met Thr Glu Asn Asp Ile Gln Pro Trp Leu Phe Ala Ile Ala Pro Leu
1   5   10   15

Pro Gly Glu Ser Leu Ser His Phe Leu Gly Arg Phe Arg Arg Arg Asn
            20   25   30

His Leu Thr Pro Ser Ser Leu Gly Gln Ile Ala Lys Ile Gly Ala Val
        35   40   45

Val Ala Arg Trp Glu Arg Phe His Phe Asn Pro Tyr Pro Thr Gln Gln
    50   55   60

Glu Phe Glu Ala Leu Ala Glu Val Val Gly Val Glu Val Glu Arg Val
65   70   75   80

Trp Glu Met Leu Pro Pro Met Gly Glu Gly Met Lys Cys Glu Pro Ile
                85   90   95

Arg Leu Cys Cys Ala Cys Tyr Ala Glu Ser Pro Cys His Arg Ile Glu
            100   105   110

Trp Gln Phe Lys Ser Val Trp Lys Cys Asp Arg His Gln Leu Lys Leu
        115   120   125

Leu Ala Lys Cys Pro Gln Cys Glu Ala Arg Phe Lys Ile Pro Ala Leu
    130   135   140

Trp Glu Asn Gly Arg Cys Asp Arg Cys Gln Ile Thr Phe Gly Glu Leu
145   150   155   160

Ala Ala Tyr Gln Lys Ser Ala
                165

<210> 34
<211> 713
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
      MG64 effector sequence

<220>
<223> MG64-7 effector

<400> 34
Met Ser Gln Ile Thr Ile Gln Cys Arg Leu Val Ala Ser Ala Ser Thr
1   5   10   15

Arg Gln Lys Leu Trp Lys Leu Met Ala Glu Phe Asn Thr Pro Leu Ile
            20   25   30

Asn Glu Leu Leu Ile Leu Ala Tyr Gln His Pro Asp Phe Lys Thr Trp
        35   40   45

Gln His Lys Gly Ala Ile Pro Ala Gly Ile Ile Lys Gln Leu Cys Glu
    50   55   60

Pro Leu Lys Thr Asp Ala Arg Phe Val Gly Gln Pro Gly Arg Phe Phe
65   70   75   80

Ala Ser Ala Ile Ala Thr Val Ser Tyr Ile Tyr Lys Ser Trp Val Lys
                85   90   95

Val Gln Lys Arg Leu Gln Leu Gln Ile Asp Gly Lys Thr Arg Trp Leu
            100   105   110

Glu Met Leu Asn Ser Asp Thr Glu Leu Val Glu Met Ala Gly Val Ala
        115   120   125

Leu Asp Thr Leu Arg Ala Thr Ala Thr Glu Leu Leu Asn Gln Leu Asn
    130   135   140

Pro Gln Pro Thr Thr Glu Glu Ser Pro Lys Lys Lys Gly Lys Lys Ala
145   150   155   160

Lys Lys Thr Gln Gln Pro Gln Gly Glu Arg Ser Leu Ser Lys Ile Leu
                165   170   175

Phe Asp Thr Tyr Arg Asp Ala Glu Asp Ile Gln Thr Arg Cys Ala Ile
            180   185   190

Ser Tyr Leu Leu Lys Asn Gly Cys Lys Ile Arg Ser Gln Glu Glu Asp

```
        195                 200                 205

Ser Lys Lys Phe Ala Gln Arg Arg Arg Lys Val Glu Ile Gln Ile Gln
    210                 215                 220

Arg Leu Thr Asp Gln Leu Ala Ser Arg Val Pro Lys Gly Arg Asp Leu
225                 230                 235                 240

Thr Ala Thr Lys Trp Leu Glu Ser Leu Leu Glu Ala Thr Arg Lys Val
                245                 250                 255

Pro Lys Asn Glu Ala Glu Ala Lys Ser Trp Gln Asp Ser Leu Asn Arg
            260                 265                 270

Gln Ser Ser Thr Leu Pro Phe Pro Val Ala Phe Glu Ser Ser Glu Asp
        275                 280                 285

Met Ile Trp Phe Arg Lys Leu Lys Leu Asn Asn Ile Pro Ile Lys Leu
    290                 295                 300

Trp Thr Ile Leu Leu Tyr Ile Asp Tyr Leu Ile Val Ile Leu Phe Val
305                 310                 315                 320

Arg Asp Ser Leu Gln Asn Glu Ala Leu Trp Phe Lys Asn Phe Lys Ile
                325                 330                 335

Asn Asn Ile His Val Leu Thr Lys Leu Trp Met Ile Leu Leu Asn Ile
            340                 345                 350

Asn Ser Phe Ala Gly Val Leu Tyr Leu Gly Gly Val Leu Lys Lys Tyr
        355                 360                 365

Gln Gln Arg Ile Tyr Val His Phe Asn Gly Leu Ser Asp Cys Thr Phe
    370                 375                 380

Glu Ile Tyr Cys Asp Ser Arg His Ile His Trp Phe Lys Arg Phe Leu
385                 390                 395                 400

Glu Asp Gln Gln Ile Lys Arg Ser Ser Lys Asn Gln His Ser Ser Ser
                405                 410                 415
```

```
Leu Phe Thr Ile Arg Ser Gly Arg Ile Ala Trp Lys Ser Ala Gln Gly
            420                 425                 430

Lys Gly Lys Pro Trp Asn Val Asn Arg Leu Met Leu Tyr Cys Cys Val
        435                 440                 445

Asp Thr Arg Leu Trp Thr Ala Glu Gly Thr Lys Leu Val Val Glu Glu
    450                 455                 460

Lys Ala Glu Glu Ile Ala Lys Thr Ile Thr Arg Thr Lys Glu Lys Glu
465                 470                 475                 480

Thr Lys Glu Lys Val Gln Leu Asn Asp Lys Gln Leu Ala Tyr Ile Lys
                485                 490                 495

Arg Lys Asn Ala Thr Leu Thr Arg Ile Asn Asn Pro Phe Pro Arg Pro
            500                 505                 510

Ser Lys Pro Leu Tyr Asn Gly Gln Ser His Ile Leu Val Ser Val Ser
        515                 520                 525

Leu Gly Leu Glu Lys Pro Ala Thr Leu Ala Val Leu Asn Ala Met Thr
    530                 535                 540

Gly Lys Ile Ile Ala Tyr Arg Ser Val Lys Gln Leu Leu Gly Lys Asn
545                 550                 555                 560

Tyr Lys Leu Leu Asn Gln Gln Arg His Gln Lys Gln Ala Leu Ser His
                565                 570                 575

Gln Arg Lys Ile Ala Gln Thr Leu Ala Ala Pro Asn Gln Phe Gly Asp
            580                 585                 590

Ser Glu Leu Gly Glu His Ile Asp Arg Leu Leu Ala Lys Glu Ile Ile
        595                 600                 605

Ala Val Ala Gln Lys Phe Asn Ala Gly Ser Ile Val Val Pro Asn Leu
    610                 615                 620

Asp Asn Met Arg Glu Gln Val Asn Ser Glu Ile Gln Ala Lys Ala Glu
```

625 630 635 640

Glu Lys Cys Pro Glu Ser Ile Glu Ala Gln Lys Lys Tyr Ala Ser Ser
645 650 655

Tyr Arg Arg Ser Val Asn Gln Trp Ser Tyr Arg Arg Leu Ile Asp Cys
660 665 670

Ile Thr Asn Gln Ala Ala Lys Ala Gly Ile Val Ile Glu Glu Asn Lys
675 680 685

Gln Pro Ile Arg Ala Ser Pro Gln Asp Lys Ala Lys Gln Leu Ala Leu
690 695 700

Ser Ala Tyr His Ala Arg Lys Lys Ser
705 710

<210> 35
<211> 587
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-7-B transposition protein

<400> 35
Met Lys Asp Ala Glu Ser Ala Thr Asn Ser Pro Met Thr His Ala Ser
1 5 10 15

Ile Ala Asp Ala Glu Asn Gly Lys Ala Glu Ala Asn Ile Ile Val Ser
20 25 30

Glu Leu Ser Asp Glu Ala Leu Leu Lys Met Glu Val Ile Gln Ser Leu
35 40 45

Leu Lys Asn Ser Asp Arg Ser Thr Tyr Gly Glu Leu Leu Lys Gln Ser
50 55 60

Ala Glu Lys Leu Gly Lys Ser Val Arg Thr Val Arg Arg Leu Val Asp
65 70 75 80

```
Lys Trp Glu Lys Glu Gly Leu Ala Gly Leu Val Gln Asn Gln Arg Asp
                85                  90                  95

Asp Lys Gly Lys His Arg Val Asp Lys Tyr Trp Gln Glu Phe Val Leu
            100                 105                 110

Thr Thr Tyr Lys Glu Asn Asn Lys Gly Ser Lys Arg Met Thr Arg Gln
        115                 120                 125

Gln Val Phe Ile Arg Ala Lys Ala Arg Ala Asp Glu Leu Glu Ile Glu
    130                 135                 140

Pro Pro Ser Arg Met Thr Val Tyr Arg Ile Leu Lys Pro Ile Ile Asp
145                 150                 155                 160

Lys Gln Glu Gln Ala Lys Ser Ile Arg Ser Pro Gly Trp Arg Gly Ser
                165                 170                 175

Arg Leu Ser Val Lys Thr Arg Asp Gly Lys Asp Leu Gln Val Glu His
            180                 185                 190

Ser Asn Gln Val Trp Gln Cys Asp His Thr Arg Val Asp Val Leu Leu
        195                 200                 205

Val Asp Gln His Gly Lys Ile Leu Ser Arg Pro Trp Leu Thr Thr Val
    210                 215                 220

Ile Asp Ser Tyr Ser Arg Cys Ile Met Gly Ile Asn Leu Gly Tyr Asp
225                 230                 235                 240

Thr Pro Ser Ser Thr Val Val Ala Leu Ala Leu Arg His Ala Ile Leu
                245                 250                 255

Pro Lys Gln Tyr Ser Leu Glu Tyr Gly Leu His Glu Glu Trp Gly Thr
            260                 265                 270

Ser Gly Leu Pro Gln Asn Phe Tyr Thr Asp Gly Gly Lys Asp Phe Arg
        275                 280                 285
```

```
Ser Asn His Leu Gln Gln Ile Gly Val Gln Leu Gly Phe Val Cys His
    290                 295                 300

Leu Arg Asp Arg Pro Ser Glu Gly Gly Ser Val Glu Arg Pro Phe Lys
305                 310                 315                 320

Thr Leu Asn Thr Glu Leu Phe Ser Thr Leu Ala Gly Tyr Thr Gly Ser
                325                 330                 335

Asn Val Gln Glu Arg Pro Glu Glu Ala Glu Lys Glu Ala Ser Phe Thr
            340                 345                 350

Leu Arg Gln Leu Glu Lys Met Leu Val Arg Tyr Ile Val Asp Asn Tyr
        355                 360                 365

Asn Gln Arg Ile Asp Ala Arg Met Gly Asp Gln Thr Arg Phe Gln Arg
    370                 375                 380

Trp Glu Ser Gly Leu Ile Ala Met Pro Asp Leu Leu Ser Glu Arg Asp
385                 390                 395                 400

Leu Asp Ile Cys Leu Met Lys Gln Thr Arg Arg Gln Val Gln Arg Gly
                405                 410                 415

Gly Tyr Leu Gln Phe Glu Asn Leu Met Tyr Arg Gly Glu Leu Leu Ala
            420                 425                 430

Gly Tyr Ala Gly Glu Ser Val Val Leu Arg Tyr Asp Pro Lys Asp Ile
        435                 440                 445

Thr Thr Ile Leu Val Tyr Arg Ile Glu Glu Gly Lys Glu Ile Phe Leu
    450                 455                 460

Ala Arg Ala Tyr Ala Gln Asp Leu Glu Thr Glu Glu Leu Ser Leu Asp
465                 470                 475                 480

Glu Ala Lys Ala Ser Ser Arg Lys Val Arg Glu Ala Gly Lys Ala Ile
                485                 490                 495

Ser Asn Arg Ser Ile Leu Ala Glu Ile Arg Glu Arg Glu Thr Phe Pro
            500                 505                 510
```

```
Thr Gln Lys Lys Thr Arg Lys Glu Arg Gln Lys Leu Glu Gln Ala Glu
        515                 520                 525

Val Lys Lys Ala Lys Gln Leu Thr Pro Ala Glu Thr Glu Glu Glu Ile
    530                 535                 540

Ile Val Val Ser Ile Asp Ala Lys Pro Thr Ala Lys Asn Pro Leu Glu
545                 550                 555                 560

Ser Glu Leu Cys Thr Glu Ser Gly Glu Pro Asp Met Pro Glu Val Leu
                565                 570                 575

Asp Tyr Glu Gln Met Arg Glu Asp Tyr Gly Trp
            580                 585
```

<210> 36
<211> 278
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-7-C transposition protein

<400> 36

```
Met Val Ala Lys Glu Ala Gln Glu Val Ala Lys Gln Leu Gly Asp Ile
1               5                   10                  15

Pro Val Asn Asp Glu Lys Leu Gln Ala Glu Ile His Arg Leu Asn Arg
            20                  25                  30

Lys Gly Phe Val Pro Leu Glu Gln Val Gln Thr Leu His Asp Trp Leu
        35                  40                  45

Glu Gly Lys Arg Gln Ser Arg Gln Ser Gly Arg Val Val Gly Glu Ser
    50                  55                  60

Arg Thr Gly Lys Thr Met Gly Cys Asp Ala Tyr Arg Leu Arg Asn Lys
65                  70                  75                  80
```

```
Pro Lys Gln Glu Ala Gly Lys Pro Pro Thr Val Pro Ile Ala Tyr Ile
                85                  90                  95

Gln Ile Pro Gln Glu Cys Gly Ala Lys Glu Phe Phe Gly Val Ile Leu
            100                 105                 110

Glu His Leu Lys Tyr Gln Val Thr Lys Gly Thr Val Ala Glu Val Arg
        115                 120                 125

Asp Arg Ala Leu Arg Val Leu Lys Gly Cys Gly Val Glu Met Leu Ile
    130                 135                 140

Ile Asp Glu Ala Asp Arg Phe Lys Pro Lys Thr Phe Ala Glu Val Arg
145                 150                 155                 160

Asp Ile Phe Asp Lys Leu Glu Ile Pro Val Ile Leu Val Gly Thr Asp
                165                 170                 175

Arg Leu Asp Ala Val Ile Lys Arg Asp Glu Gln Val Tyr Asn Arg Phe
            180                 185                 190

Arg Ser Cys His Arg Phe Gly Lys Leu Ser Gly Glu Glu Phe Lys Arg
        195                 200                 205

Thr Val Asp Ile Trp Glu Lys Lys Val Leu Gln Leu Pro Val Thr Ser
    210                 215                 220

Asn Leu Ser Ser Lys Thr Met Leu Lys Thr Leu Gly Glu Ala Thr Gly
225                 230                 235                 240

Gly Tyr Ile Gly Leu Met Asp Met Ile Leu Arg Glu Ser Ala Ile Arg
                245                 250                 255

Ala Leu Lys Lys Gly Leu Gln Lys Ile Asp Leu Asn Thr Leu Lys Glu
            260                 265                 270

Val Thr Ala Glu Tyr Arg
        275
```

<210> 37

<211> 167
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-7-Q transposition protein

<400> 37
Met Glu Ser Glu His Ile Lys Ala Trp Leu Phe Gln Val Glu Pro Phe
1 5 10 15

Glu Gly Glu Ser Leu Ser His Phe Leu Gly Arg Phe Arg Arg Ala Asn
20 25 30

Asp Leu Thr Pro Gly Gly Leu Gly Ser Gln Ala Gly Leu Gly Gly Ala
35 40 45

Ile Ala Arg Trp Glu Lys Phe Arg Phe Asn Pro Pro Pro Ser His Gln
50 55 60

Gln Leu Glu Lys Leu Ala Val Val Ala Gly Ile Asp Ala Gly Arg Leu
65 70 75 80

Val Gln Met Leu Ala Pro Ala Gly Val Ser Ile Lys Leu Glu Pro Ile
85 90 95

Arg Leu Cys Ala Ala Cys Tyr Ala Glu Ser Pro Cys His Lys Ile Glu
100 105 110

Trp Gln Phe Lys Glu Thr Arg Gly Cys Lys His His Lys Leu Arg Leu
115 120 125

Leu Ser Glu Cys Pro Asn Cys Gly Ala Ser Phe Lys Ile Pro Ala Leu
130 135 140

Trp Val Asp Gly Trp Cys His Arg Cys Phe Thr Leu Phe Gly Glu Met
145 150 155 160

Val Asn His Gln Lys Pro Cys
165

<210> 38
<211> 739
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG108 effector sequence

<220>
<223> MG108-1 effector

<400> 38
Met Ser Phe Gln Thr Val Arg Cys Val Leu Cys Ala Ser Glu Glu Met
1 5 10 15

Arg Gln Lys Val Trp Leu Trp Thr Glu Lys Tyr Thr Leu Leu Val Asn
20 25 30

Glu Leu Leu Asp Gln Val Ala Ser His Glu Lys Phe Ser Glu Trp Gln
35 40 45

Ala Gln Gly Ser Ile Ala Arg Lys Glu Val Glu Ile Leu Leu Lys Ser
50 55 60

Leu Lys Ala Ser Asp Gln Phe Lys Gly Leu Pro Gly Arg Phe Tyr Thr
65 70 75 80

Ser Ala Val Leu Met Val Gln Tyr Thr Tyr Leu Ala Trp Leu Ala Leu
85 90 95

Gln Lys Lys Arg His Phe Lys Ile Ile Gly Lys Gln Arg Trp Leu Glu
100 105 110

Val Leu Glu Ala Asp Leu Glu Met Ala Ser Asn Thr Asp Phe Glu Trp
115 120 125

Gln Glu Ile Gln Ser Arg Ala Phe Glu Leu Leu Ala Gln Thr Lys Lys
130 135 140

Ser Ser Asp Thr Val Ser Ala Lys Pro Val Lys Lys Gly Ala Val Asn
145 150 155 160

```
Gln Glu Ala Ala Lys Ile Gln Lys Gln Thr Pro Phe Asn Ala Leu Leu
               165                 170                 175

Glu Arg Phe Ser Ser Glu Lys Asp Leu Leu Asn Arg Arg Ala Ile Ala
            180                 185                 190

His Leu Leu Arg Asn Asn Leu Gln Val Glu Glu Asn Glu Glu Ser Leu
        195                 200                 205

Asp Lys Leu Leu Phe Arg Phe Glu Lys Lys Arg Ile Glu Ile Glu Arg
    210                 215                 220

Ile Glu Lys Gln Leu Glu Ser Arg Arg Pro Lys Gly Arg Asp Pro Val
225                 230                 235                 240

Gly Thr Arg Phe Met Glu Ser Leu Glu Glu Ala Thr Ala Leu Pro Glu
               245                 250                 255

His Thr Lys Ser Glu Glu Ile Glu Ala Glu Leu Asn Gly Trp Lys His
            260                 265                 270

Gln Lys Gln Val Arg Ser Leu Asn Asn Leu Pro Tyr Pro Val Leu Phe
        275                 280                 285

Val Ser Ile Thr Asp Leu Ile Trp Ser Leu Gln Ser Asn Glu Asn Glu
    290                 295                 300

Gln Glu Ser Val Lys Gly Asn Gln Val Gln Pro Ala Arg Arg Gly Lys
305                 310                 315                 320

Cys Lys Lys Lys Gln Lys Arg Ser Ser Asp Arg Ile Cys Val Asn Phe
               325                 330                 335

Asn Gly Phe Lys Lys Phe Leu Gly Glu Thr Gln Gly Val Phe Lys Ile
            340                 345                 350

Gln Cys Asp Arg Arg Gln Leu Pro Ile Phe Arg Gln Ile Val Ser Asp
        355                 360                 365

Trp Arg Ala Tyr Gln Gln Leu Pro Asp Asp Glu Lys Phe Gly Leu Gly
```

```
    370                 375                 380

Leu Met Leu Leu Lys Ser Ala Gln Leu Ile Trp Lys Lys Asp Gln Gln
385                 390                 395                 400

Lys Leu Tyr Lys Lys Lys Ala Lys Asn Asn Ser Thr Gln His Glu Thr
                405                 410                 415

Ala Ser Gly Gly Gln Leu Val His Pro Trp Gln Thr His Arg Leu Tyr
            420                 425                 430

Leu His Cys Cys Ile Asn Pro Asp Leu Leu Thr Ala Glu Gly Thr Glu
        435                 440                 445

Val Val Arg Gln Lys Lys Leu Pro Ala Thr Ala Lys Lys Leu Glu Gly
    450                 455                 460

Ser Lys Lys Arg Gln Ala Glu Val Ala Glu Arg Ile Gln Thr Ser Asp
465                 470                 475                 480

Leu Gly Glu Asp Ser Thr Met Leu Lys Lys Leu His Thr Asn Leu Val
                485                 490                 495

Asn Arg Ile Ser Ala Asn Ala Thr Ser Leu Asn Arg Leu Glu Asn Ala
            500                 505                 510

Thr Pro Gln Arg Pro Ser Lys Pro Ala Tyr His Gly Gln Pro His Ile
        515                 520                 525

Ile Val Gly Val Ser Phe Ser Arg Glu Gln Leu Val Gly Val Ala Val
    530                 535                 540

Val Asp Thr Arg Thr Gly Arg Val Ile Glu Tyr Gln Ser Thr Arg Leu
545                 550                 555                 560

Leu Leu Ser Asn Pro Ala Val Lys Ala Lys Arg Gly Gly Arg Thr Val
                565                 570                 575

Thr Gln Leu Arg Leu Glu Lys Tyr Arg Leu Val Asn Arg Arg Gln Arg
            580                 585                 590
```

```
Gln Gln Gln Gln Asn Ala Ile Arg Arg Ser Lys Glu Gln Lys Gln Asn
        595                 600                 605

Arg Tyr Ala Arg Ser Lys Ser Glu Ser Asn Leu Gly Gln Tyr Leu Asp
    610                 615                 620

Arg Leu Ile Ala Ser Lys Val Ile Lys Leu Ala Ile Lys Trp Gln Ala
625                 630                 635                 640

Ser Ser Ile Val Val Pro Asn Leu Gly Asp Ile Arg Glu Ser Val Glu
                645                 650                 655

Ala Ser Leu Lys Ala Glu Ala Glu Arg Lys Phe Pro Asn Glu Phe Glu
            660                 665                 670

Arg Gln Lys Lys Tyr Ala Lys His Ile Arg Thr Ser Phe His Thr Trp
        675                 680                 685

Ser Tyr Gly Arg Leu Thr Glu Cys Ile Arg Ser Cys Ala Ala Lys Ala
    690                 695                 700

Gln Ile Pro Ile Glu Gln Gly His Gln Pro Ala Gln Gly Asp Leu Lys
705                 710                 715                 720

Glu Lys Ala Met Gln Val Ala Leu His Ala Tyr Tyr Thr Arg Gln Ser
                725                 730                 735

Pro Lys Gly
```

<210> 39
<211> 302
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG108 transposition protein sequence

<220>
<223> MG108-1-C transposition protein

<400> 39

```
Met Ser Gln Leu Gln Phe Ala Glu Pro Asp Arg Pro His Ser Ser Asn
1               5                  10                  15

Glu Glu Thr Leu Phe Leu Glu Glu Pro Arg Arg Ser Ile Glu Ala Gln
            20                  25                  30

Ala Glu Val Glu Arg Ile Gly Lys Ala Asp Thr Tyr Leu Val Leu Asp
        35                  40                  45

Arg Asp Ile Ala Leu Phe Asp Trp Leu Asp Asp Gln Arg Asp Thr Lys
    50                  55                  60

Leu Cys Gly Tyr Ile Ile Gly Ser Gln Gly Ser Gly Leu Pro Lys Ala
65                  70                  75                  80

Cys Gln Phe Tyr Arg Met Lys His Val Lys Arg Arg Gly Arg Leu Leu
                85                  90                  95

Gln Val Pro Ala Ser Val Ile Tyr Ala Glu Val Lys Gln His Gly Gly
            100                 105                 110

Pro Thr Asp Leu Tyr His Ala Ile Leu Glu Glu Ile Gly His Pro Leu
        115                 120                 125

Ala Lys Ala Gly Ala Leu Arg Asp Leu Arg Ser Arg Ala Trp Gly Thr
    130                 135                 140

Leu Lys Asp Tyr Gly Val Lys Leu Leu Ile Val Gly Asn Ala Asp Tyr
145                 150                 155                 160

Leu Lys Leu Glu Ser Phe Asn Glu Leu Ile Asp Leu Phe Gly Lys Leu
                165                 170                 175

Arg Ile Ala Ile Ala Leu Val Gly Thr Asp Thr Leu Ala Glu Thr Leu
            180                 185                 190

Asp Arg Arg Gly Ser Ala Tyr Arg Arg Ile His Asp Ala Phe Leu Glu
        195                 200                 205

Ser Phe Asp Phe Leu Asn Leu Thr Pro Glu Asp Ile Arg Glu Val Ile
    210                 215                 220
```

```
Glu His Trp Glu Asn Gln Phe Leu Pro Thr His Ser Arg Leu Asn Phe
225                 230                 235                 240


Ala Arg Ile Pro Asp Val Arg Lys Phe Leu Glu Leu Lys Ser Lys Gly
                245                 250                 255


Arg Ile Glu Pro Leu Tyr Asp Leu Leu Arg Lys Ile Ala Ile Leu Lys
            260                 265                 270


Ile Asp Glu Pro Lys Leu Glu Leu Asn Ser Thr Asn Leu Ala Asp Arg
        275                 280                 285


Leu Gly Lys Arg Asn Ala Pro Lys Asp Arg Ile Ser Thr Arg
    290                 295                 300


<210> 40
<211> 650
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
      MG108 transposition protein sequence

<220>
<223> MG108-1-B transposition protein

<400> 40
Met Phe Arg Thr Ser Ala Asp Ser Asn Glu Asn Ser Ala Ser Ser Pro
1               5                   10                  15


Ile Ser Glu Leu Ser Ser Pro Gly Leu Val Glu Ala Gly Ser Phe Lys
            20                  25                  30


Leu Leu Gln Asp Ile Ser Pro Glu Val Gln Gln Glu Met Asp Leu Ile
        35                  40                  45


Asp Ala Ile Val Gln Ala Ser Asn Lys Cys Asp Arg Lys Ala Ala Ile
    50                  55                  60


Asp Arg Ala Cys Ala Val Leu Asn Arg His Pro Arg Ser Ile Met Arg
65                  70                  75                  80
```

```
Lys Val Gln Arg Leu Arg Gln Glu Gly Ile Thr Ser Leu Ala Lys Gly
                85                  90                  95

Arg Lys Asp Asn Gly Gln Phe Arg Val Ser Glu Gln Trp Val Asn Phe
            100                 105                 110

Ile Thr Lys Met Tyr Ala Arg Gly Gln Lys Asp Ser Leu Arg Phe Ser
        115                 120                 125

Pro His Lys Val Phe Val Arg Leu Thr Val Phe Ala Ala Gln Gly Asn
    130                 135                 140

Arg Leu Arg Glu Ser Lys Tyr Ser Asp Leu Phe Lys Asp Tyr Pro Glu
145                 150                 155                 160

Val Leu Lys Asp Leu Ile Gln Gly Lys His Pro Ser His Pro Thr Val
                165                 170                 175

Tyr Lys Ile Ile Asp Asn Tyr Leu Leu Gln Lys Asn Thr Lys Ala Arg
            180                 185                 190

His Pro Gly Ala Phe Pro Gly Lys Leu Val Leu Gln Thr Thr Glu Gly
        195                 200                 205

Leu Ile Ser Leu Thr His Ser Asn Gln Val Trp Gln Ile Asp His Thr
    210                 215                 220

Lys Leu Asp Ile Leu Leu Val Ile Lys Val Glu Glu Asp Asp Lys Val
225                 230                 235                 240

Glu His Glu Gly Lys Lys Ile Thr Ala Arg Pro Tyr Leu Thr Leu Val
                245                 250                 255

Val Asp Ser Tyr Ser Ser Cys Val Ala Gly Tyr Tyr Leu Gly Phe Glu
            260                 265                 270

Pro Ala Gly Ser His Glu Val Ser Leu Ala Leu Arg Asn Ala Ile Leu
        275                 280                 285

Pro Lys Gln Tyr Gly Gln Glu Tyr Lys Leu Gln Ala Lys Trp Asn Ile
```

```
    290                 295                 300

Tyr Gly Val Pro Asp Tyr Val Met Thr Asp Arg Ala Lys Glu Phe Lys
305                 310                 315                 320

Ser Glu His Leu Lys Gln Ile Ser Phe Gln Leu Gly Phe Gln Arg Arg
                325                 330                 335

Leu Arg Ala Phe Pro Ser Ala Gly Gly Leu Val Glu Ser His Phe Asp
            340                 345                 350

Lys Ile Asn Lys Glu Tyr Leu Ser Glu Lys Pro Gly Tyr Thr Gly Ser
        355                 360                 365

Ser Val Glu Glu Arg Pro Lys Glu Ala Glu Asn Ala Ala Cys Leu Thr
    370                 375                 380

Leu Asp Asp Leu Glu Arg Glu Leu Val Arg Tyr Phe Val Asp His Asn
385                 390                 395                 400

Asn Gln His Phe Tyr Ser Glu Asp Lys Leu Thr Val Ser Glu Arg Met
                405                 410                 415

Ser Lys Pro Lys Arg Phe Glu Arg Trp Gln Asp Gly Leu Leu Val Glu
            420                 425                 430

Pro Glu Ile Leu Asn Glu Arg Glu Leu Asp Ile Cys Leu Met Lys Ser
        435                 440                 445

Ile Pro Arg Lys Val Glu Lys Tyr Gly Cys Val Arg Phe Glu Thr Leu
    450                 455                 460

Arg Tyr Gln Gly Glu Cys Leu Ala Asn Lys Ala Phe Ile Gly Arg Asp
465                 470                 475                 480

Ile Thr Leu Arg Tyr Asp Gln Arg Asn Ile Ala Thr Leu Leu Val Tyr
                485                 490                 495

Ser Tyr Ser Thr Thr Ser Gln Ile Gln Glu Phe Ile Gly Thr Val Gln
            500                 505                 510
```

Ala Ile Gly Leu Glu Arg Glu Arg Phe Ser Leu Gly Glu Trp Glu Ala
515 520 525

Leu Lys Gln Lys Met Arg Asp Arg Asp Lys Glu Ile Asp Asn Ser Ile
530 535 540

Leu Met Ala Glu Arg Met Gly Leu Ile Glu Ser Thr Lys Glu Gln Ile
545 550 555 560

Lys Ser Ser Arg Ser Lys Arg Gln Arg Ile Lys Lys Ala His Thr Glu
565 570 575

Gln Glu Arg Lys Ala Asn Arg Ser Lys Val Val Glu Leu Tyr Pro Glu
580 585 590

Arg Ala Ile Glu Asp Thr Thr Val Ile Gln Glu Asn Val Glu Ser Thr
595 600 605

Pro Asn Ser Ser Ile Gln Lys Arg Val Glu Ser Thr Gln Glu Gln Gln
610 615 620

Ile Gln Gln Ala Ala Pro Lys Thr Arg Arg Val Arg Thr Val Ala Pro
625 630 635 640

Asp Trp Asn Gln Ile Lys Gln Asp Asn Trp
645 650

<210> 41
<211> 200
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG110 effector sequence

<220>
<223> MG110-1-6 effector

<400> 41
Met Lys Leu Phe Ser Val Lys Ile Thr Tyr Leu Pro Glu Lys Cys Asp
1 5 10 15

```
His Ala Leu Leu Ala Gly Arg Cys Ile Lys Val Leu His Ala Phe Met
            20                  25                  30

Ser Arg Asn Gly Gln Phe Asn Ile Ala Val Ala Phe Pro Arg Trp Ser
        35                  40                  45

Glu Asn Thr Ile Gly Asn Gln Leu Val Phe Val Ser Pro Asp Tyr Lys
    50                  55                  60

Phe Leu Asp Met Leu Leu Glu Gln Pro Tyr Phe Arg Met Met Ile Glu
65                  70                  75                  80

Asn Gly Leu Phe Glu Thr Gly Thr Val Val Asp Leu Pro Val Ser Asp
                85                  90                  95

Ser Tyr Val Lys Phe Val Arg Asn Gln Ser Ile Asp Lys Met Thr Pro
            100                 105                 110

Ala Ala Lys Ala Arg Arg Leu Arg Arg Ala Lys Lys Arg Ala Leu Ala
        115                 120                 125

Arg Gly Glu Val Phe Asp Pro Ile Ala Pro Arg Ser Lys Asp Val Asp
    130                 135                 140

Phe Phe His Ser Ile Pro Met Glu Ser Ser Glu Ser Glu Met Ser Tyr
145                 150                 155                 160

Leu Leu Arg Val Gln Arg Tyr Glu Val Gln Gln Ala Asn Pro Val Ala
                165                 170                 175

Ser Phe Glu Val Cys Ser Tyr Gly Leu Ser Thr Asn Glu Ser His Gln
            180                 185                 190

Ala Leu Ile Pro Ser Ser Val Thr
        195                 200
```

<210> 42
<211> 342
<212> PRT
<213> Unknown

<220>

<223> Description of Unknown:
MG110 effector sequence

<220>
<223> MG110-1-7 effector

<400> 42

```
Met Gln Leu Pro Asn Gln Leu Ser Tyr Lys Arg Ser Ile Asn Pro Ser
1               5                  10                  15


Lys Ala Ile Phe Tyr Tyr Arg Arg Asp Glu Asp Leu Tyr Pro Leu Pro
            20                  25                  30


Val Glu Arg Ile Lys Ile Arg Gly Ser Lys Ser Gly Phe Ser Glu Ala
        35                  40                  45


Tyr Thr Ala Lys Gly Ile Lys Glu Ser Ala Thr Ile His Ser Leu Ala
    50                  55                  60


Thr Gly Asn Pro His Thr Ile Asp Thr Cys Tyr Leu Pro Pro Ala Ala
65                  70                  75                  80


Asp Ser Leu Val Cys Arg Phe Ser Leu Arg Val Ser Ala Asn Ser Leu
                85                  90                  95


Phe Pro Asp Arg Cys Ser Glu Ile Ala Phe Lys Asp Thr Val Ser Gln
            100                 105                 110


Phe Leu Asn Ser Tyr Ile Ala Lys Asp Gly Phe Lys Glu Leu Ala Ile
        115                 120                 125


Arg Tyr Ser Lys Asn Ile Ala Met Gly Thr Trp Leu Trp Arg Asn Lys
    130                 135                 140


Glu Gly Asn Thr Phe Asp Val Leu Val Arg Thr Ser Gln Gly Asn Glu
145                 150                 155                 160


Tyr Lys Phe Lys Asn Ala His Met Leu Phe Trp Asp Ser Val Trp Pro
                165                 170                 175


Asp Glu Ser Ser Glu Leu Leu Ala Leu Leu Ser Asp Glu Leu Ala Val
            180                 185                 190
```

```
Ala Leu Thr Lys Pro Arg Tyr Val Trp His Cys Asp Val Trp Ala Glu
        195                 200                 205

Val Lys Met Pro Phe Cys Ser Glu Val Phe Pro Ser Gln Cys Phe Val
    210                 215                 220

Asp His Asp Asp Lys Gln Ser Ala Ser Lys Val Leu Leu Thr Thr Asp
225                 230                 235                 240

Ile Asp Gly Val Met Thr Ala Cys Tyr Asn Ala Asp Lys Val Gly Ala
                245                 250                 255

Ala Ile Gln Met Ile Asp Asp Trp Trp Asp Glu Ser Cys Asp Phe Pro
            260                 265                 270

Leu Arg Val Asn Glu Tyr Ala Ala Asp His Glu Asn Leu Ile Ala Arg
        275                 280                 285

Arg His Pro Ser Thr Glu Arg Asp Phe Tyr Gln Cys Leu Gln Asn Leu
    290                 295                 300

Arg Gly Tyr Thr Asp Lys Leu Asn Lys Val Lys Ser Val Asp Asp Ile
305                 310                 315                 320

Glu Pro Asp Ala His Phe Val Ala Ser Val Leu Val Lys Gly Gly Met
                325                 330                 335

Phe Gln Gly Gly Lys Ser
            340
```

<210> 43
<211> 682
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG110 effector sequence

<220>
<223> MG110-1-8 effector

<400> 43

```
Met Ser Leu Gln Asp Thr Leu Gln Ala Val Lys Ala Leu Lys Ala Lys
1               5                  10                  15

Glu Arg Tyr Asp Ala Leu Lys Lys Ser Phe Glu Leu Tyr Ser Glu His
            20                  25                  30

Val Asp Val Thr Gly Ser Glu Leu Ala Cys Leu Thr Ile Leu Val Asn
        35                  40                  45

Leu Val Ser Lys Asp Ala Val Glu Leu Glu Gly Leu Arg His Ala Lys
    50                  55                  60

His Val Leu Lys Ser Asp Glu Phe Trp Gly Lys Phe Ser Lys Val Ala
65                  70                  75                  80

Ser Gln Leu His Thr His Asn Leu Lys Trp Pro Asp Ser Arg Val Asn
                85                  90                  95

Leu Lys His His Ile Arg Ala Leu Pro Glu Ile Gly Val Leu Pro Lys
            100                 105                 110

Phe Gly Trp Ser Gly Asn Ser Gly Asp Tyr Arg Phe Gly Arg Leu Leu
        115                 120                 125

Thr Ser Thr Phe Ile Trp Gln Gly Leu Glu His Ser Leu Ile Ser Val
    130                 135                 140

Trp Leu Gly Asp Phe Val Ala Trp Arg Lys Ala Ala Tyr Lys Leu Gly
145                 150                 155                 160

Ile Thr Lys Val Phe Trp Tyr Gln Ile Lys Arg Asp Leu Glu Asp Leu
                165                 170                 175

Phe Gln Glu Ser His Phe Pro Asp Val Val Asp Ser Tyr Ser Pro Glu
            180                 185                 190

Leu Gln Phe Pro Tyr Lys Asp Gln Tyr Leu Ser Val Thr Pro Val Ala
        195                 200                 205

Ser His Glu Thr Gln Leu Ala Ile Gln His Val Thr Gly Leu Pro Met
    210                 215                 220
```

```
His Ser Leu Ser Phe Pro His Pro Ser Ala Leu Gly Val Leu Cys Gly
225                 230                 235                 240

Ser Leu Gly Gly His Val Arg Leu Ile Arg Leu Ser Pro Leu Gly Asn
                245                 250                 255

Gln Asn Val Gln Arg Ser Ser Leu Gly Ser Leu Ser Leu Lys Asn Tyr
            260                 265                 270

Leu Asn Pro Tyr Val Leu Thr Ala Ile Pro Ala Ile Asn Leu Tyr Arg
        275                 280                 285

Arg Ile Val Asp Val Arg Thr Tyr Ser Ser Leu Arg Leu Lys Arg Arg
    290                 295                 300

Ala Arg Leu Asn Ala Leu Ser Thr Leu Asp Gly Ile Leu Ala Glu Trp
305                 310                 315                 320

Val Ala Pro Leu Val Gln Val Lys Leu Ser Ser Asn Val Asp Gly Arg
                325                 330                 335

Ile Asp Gly Leu Gly Ala Glu Glu Gln Glu Phe Val Glu Ser Ser Ser
            340                 345                 350

Val Asp Ile Glu Glu Phe Ser Arg Tyr Leu Asn Arg Lys Leu His Gly
        355                 360                 365

Met Leu Glu Leu Gly Lys Tyr Ser Arg Lys Phe Ser Tyr His Gln Arg
    370                 375                 380

Leu Val Gly Val Thr Gln Lys Arg Leu Ala Ser Leu Leu Arg Arg Leu
385                 390                 395                 400

Leu Leu Ala Glu Thr Gln Ser Glu Ser Asn Asn Thr Thr Phe Ile Ile
                405                 410                 415

Leu Lys Ser Leu Arg Ile Asn Glu Ala Asn Gly Leu Asn Asn Pro Tyr
            420                 425                 430
```

```
Val Val Gly Met Pro Ser Ile Ile Gly Leu Tyr Gly Phe Leu His Arg
        435                 440                 445

Phe Glu Cys Gln Leu Arg Glu Ile Tyr Ala Asp Ile Ser Val Asp Ser
    450                 455                 460

Phe Ala Leu His Cys Gly Glu Tyr Ser Tyr His Ala Ser Asn Asn Leu
465                 470                 475                 480

Pro Ala Pro Ser Ile Pro Asp Lys Glu Met Arg Ile Lys Arg Ser Gly
                485                 490                 495

Ile Thr Pro Glu Phe Lys Phe Asp Gly Lys Phe Ser Ile Ile Val Lys
            500                 505                 510

Leu His Cys Leu Thr Asp Asn Ala Val Leu Asp Val Glu Gln Ile Lys
        515                 520                 525

Ala Ala Val Pro Glu Arg Leu Trp Gly Gly Ser Val His Pro Pro Tyr
    530                 535                 540

Leu Tyr Glu Asn Thr Glu Trp Ala Ala Ile Ala Tyr Gly Ser Ala Asp
545                 550                 555                 560

Leu Glu Arg Tyr Leu Ile Arg Asn Leu Phe Phe Gly Asn Trp Ile Thr
                565                 570                 575

Pro Glu Glu Glu Asp Gly Phe Glu Leu Arg Lys His Ile Glu Lys Leu
            580                 585                 590

Asn Gly Lys Asn Glu Leu Ser Leu Cys Leu Val Gly Tyr Lys Leu Leu
        595                 600                 605

Glu Lys Val Lys Pro Arg Ser Val Val Ser Gly Ile His Ala Phe Cys
    610                 615                 620

Glu Pro Val Val Asp Leu Cys Cys Leu Lys Gln Thr His Lys Val Ile
625                 630                 635                 640

Lys Pro Thr Lys Ser Ile Glu Gln Glu Leu Phe Trp Arg Tyr Val Pro
                645                 650                 655
```

Val Ser Gln Asn Cys Thr Thr Leu Arg Val Ser Pro Val Cys Gly Glu
660 665 670

Thr His Ala Ala Ser Gln Ser Ala Glu Leu
675 680

<210> 44
<211> 383
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG110 transposition protein sequence

<220>
<223> MG110-1-Q transposition protein

<400> 44
Met Gly Phe Leu Pro Asn Ser Ile His Leu Tyr Pro Asp Glu Thr Val
1 5 10 15

Glu Ser Ala Leu Leu Arg Leu Cys Lys Ala Asn His Phe Glu Arg Tyr
20 25 30

Ile Asp Leu Ser Ile Glu Ile Arg Ser Trp Leu Glu Glu His His Pro
35 40 45

Thr Ile Ala Gly Ala Phe Pro Val Ala Leu Asp Ala Val Asn Val Tyr
50 55 60

His Ala Lys Gln Ser Ser Ala Lys Arg Val Gln Ser Leu Gln Leu Leu
65 70 75 80

Glu Gln Leu Val Gly Leu Pro Arg Phe Ser Leu Leu Asp Ile Ser Phe
85 90 95

Lys His Thr Asn Ala Val Asp Ala Gly His Phe Ala Glu Val Arg Tyr
100 105 110

Lys Gln Ile Thr Ile Pro Lys Ser Phe Val Arg Ala Ser Ser Val Pro
115 120 125

```
Val Cys Ile Ala Cys Leu Arg Glu Ser Asn Tyr Val Arg Phe Asp Trp
    130                 135                 140

His Ile Ser Lys Val Thr Cys Cys Glu Lys His Lys Val Lys Leu Leu
145                 150                 155                 160

Ser Asn Cys Pro Ala Cys Asn Ala Pro Leu Asn Tyr Met Ile Ser Glu
                165                 170                 175

Asp Pro Ser His Cys Val Cys Gly Phe Asn Leu Leu Glu Thr Thr Asn
            180                 185                 190

Ser Glu Lys Gly Ala Asp Asp Trp Arg Arg Arg Ile Ser Phe Asp Gln
        195                 200                 205

Gln Gly Ser Leu Ser Ile Ser Glu Gln Leu Ala Leu Leu Leu Phe Leu
    210                 215                 220

Glu Arg Tyr Phe Ser His Val Glu Leu Gly Asp Phe Ile Gly Asn Tyr
225                 230                 235                 240

Lys Ser Leu Ile Asp Lys His Leu Gln Ile Leu Thr Ser Lys Asn Ile
                245                 250                 255

Leu Leu Ala Thr Glu Lys Ile Asn Arg Val Ser Phe Ser Ser Leu Thr
            260                 265                 270

Asp Asn Phe Leu Gly Asp Val Ala Gln Ile Gly Ser Leu Pro Asn Phe
        275                 280                 285

Ile Lys Gln His Ile Ala Asn Val Val Ile Glu Leu Ala Leu Glu Thr
    290                 295                 300

Pro Arg Ser Thr Ile Ala Asn Leu Gly Asp Ser Leu Val Ser Met Arg
305                 310                 315                 320

Asp Ala Ala Leu Ile Thr Gly Ser Thr Ile Asp Asp Ile Phe Arg Leu
                325                 330                 335

Tyr Glu Ser Gly Ile Leu Val Val Gly Lys Arg Ile Arg Asn Glu Gly
```

```
            340                 345                 350

Arg Ile Glu Ser Phe Asn Pro Leu Phe Arg Leu Arg Asp Val Ala Ala
        355                 360                 365

Ile Ala Leu Ser Tyr Ser Arg Tyr Gly Tyr Ser Gln Ser Ala Trp
    370                 375                 380

<210> 45
<211> 334
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
      MG110 transposition protein sequence

<220>
<223> MG110-1-C transposition protein

<400> 45
Met Gly Leu Thr Asp Ala Asp Lys Ala Lys Leu Arg Glu Phe Lys Asp
1               5                   10                  15

Cys Phe Cys Pro Tyr Thr Pro Val Thr Thr Val Leu Asn Asp Leu Glu
            20                  25                  30

Ser Leu Tyr Gln Ser Ser Glu Ile Gly Gly Glu Gln Leu Ser Met Leu
        35                  40                  45

Leu Arg Gly Asp Thr Gly Thr Gly Lys Ser Ala Ile Ile His His Phe
    50                  55                  60

Cys Asn Ile Lys Asn Gly Ile Asn Leu Glu Ser Ile Pro Val Leu Leu
65                  70                  75                  80

Ser Arg Val Pro Ser Lys Leu Thr Val Glu Asp Met Thr Arg Gln Leu
                85                  90                  95

Leu Ser Asp Leu Gly Val Phe Gly Ser Ser Thr His Arg Ala Lys Asn
            100                 105                 110

Ala Gln Ser Asp Ala His Leu Thr Asn Arg Leu Leu Asp Ala Leu Lys
        115                 120                 125
```

```
Val Lys Asn Thr Lys Met Ile Ile Ile Asn Glu Phe Gln Glu Leu Ile
    130                 135                 140

Glu Phe Lys Gly Ala Arg Asp Arg Gln Ala Ile Gly Asn Arg Leu Lys
145                 150                 155                 160

Leu Ile Ser Glu Glu Ala Ala Val Pro Ile Val Leu Val Gly Met Pro
                165                 170                 175

Trp Ile Asp Glu Ile Leu Asn Asp Ser Gln Trp Ala Ser Arg Leu Ala
            180                 185                 190

Thr Arg Thr His Thr Leu Gln Tyr Phe Ser Leu Ser Lys Arg Ile Arg
        195                 200                 205

Glu Tyr Phe Glu Phe Leu Glu Ala Ile Glu Thr Tyr Val Pro Cys Glu
    210                 215                 220

Ile Glu Thr Ser Ile Ile Asp Phe Glu Ile Ser Ile Ala Leu Phe Ala
225                 230                 235                 240

Ala Ser Cys Gly Glu Met Arg Gln Leu Lys Ala Ile Leu Thr Glu Ala
                245                 250                 255

Ile Lys Leu Cys Leu Ile Ser Gly Glu Pro Phe Ser Lys Gln Ala Leu
            260                 265                 270

Ser His Ser Phe Ala Asn Leu Tyr Ala Gly Ile Glu Asn Pro Phe Asp
        275                 280                 285

Met Pro Lys Glu Lys Ile Lys Val Arg Glu Val Glu Met His Ser Arg
    290                 295                 300

Tyr Ile Arg Gly Asp Ser Thr His Arg Ala Ala Ile Glu Pro Arg Arg
305                 310                 315                 320

Leu Ser Asp Phe Met Ser Leu Thr Gln Ile Leu Ser Lys Lys
                325                 330
```

<210> 46
<211> 601
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG110 transposition protein sequence

<220>
<223> MG110-1-B transposition protein

<400> 46
Met Phe Asp Asp Glu Phe Asp His Asp Pro Leu Lys Glu Asp Val Glu
1 5 10 15

Ser Ser Asn Arg Gln Ser Asp Glu Ala Ser Tyr Leu Ser Pro Asp Leu
20 25 30

Asp Ser Tyr Pro Ile Lys Ser Ser Glu Glu Ala Leu Val Arg Tyr Glu
35 40 45

Leu Val Gln Phe Ile Arg Glu Arg Thr Ser Lys Gly Trp Thr Gln Arg
50 55 60

Asn Leu Asn Pro Leu Ile Ser Glu Tyr Phe Ser Thr His Arg Ser Leu
65 70 75 80

Glu Arg Pro Ser Trp Arg Thr Val Val Arg Trp His Lys Lys Leu Leu
85 90 95

Glu Gln Gly Asp Thr Pro Val Ser Leu Ile Glu Arg His His Asn Lys
100 105 110

Gly Asn Arg Asn Arg Lys Leu Leu Leu Asp His Glu Arg Phe Phe Glu
115 120 125

Ala Ala Ile Asp Ser Phe Leu Lys Ala Lys Arg Pro Ser Val Ala Thr
130 135 140

Ala Tyr Arg Tyr Tyr Lys Asp Gln Cys Leu Leu Gln Gly Asn Asn Ile
145 150 155 160

Asn Pro Met Ser Gln Arg Ala Phe Tyr Asp Arg Ile Asp Lys Leu Asn

```
                165                 170                 175

Ser Tyr Glu Val Ala Val Lys Arg Phe Gly Lys Tyr Lys Ala Asp Ile
            180                 185                 190

Met Tyr Gly His Lys Gly Ser Thr Leu Lys Pro Glu Arg Val Met Gln
        195                 200                 205

Arg Val Glu Ile Asp His Thr Val Leu Asp Ile Thr Leu Leu Asp Asp
    210                 215                 220

Asp Thr Gly Met Pro Ile Gly Lys Pro Ser Leu Thr Leu Leu Lys Cys
225                 230                 235                 240

Ala Leu Ser Gly Cys Ile Val Gly Tyr Asn Leu Thr Phe Lys Ala Pro
                245                 250                 255

Cys Phe Ala Ser Val Ala Lys Ala Ile Cys His Thr Leu Leu Pro Lys
            260                 265                 270

Thr Ser Ser Leu Asp Ile Trp Gly Ile Asp Trp Pro Cys Tyr Gly Lys
        275                 280                 285

Ile Glu Val Leu Val Val Asp Asn Gly Val Glu Phe Trp Ile Pro Ser
    290                 295                 300

Met Glu Lys Met Cys Phe Glu Leu Gly Ile Asn Val Gln Tyr Asn Pro
305                 310                 315                 320

Val Lys Lys Pro Trp Leu Lys Pro Phe Ile Glu Arg Asn Phe Lys Thr
                325                 330                 335

Ile Asn Ser Leu Leu Leu Asp Glu Leu Ser Gly Lys Thr Phe Gly Asn
            340                 345                 350

Ile Tyr Val Arg Gly Asp Tyr Asp Ser Val Lys Asn Ala Glu Ile Pro
        355                 360                 365

Phe Gly Val Phe Val Tyr Val Phe Glu Lys Trp Ala Ala Glu Val Tyr
    370                 375                 380
```

```
Asn Cys Ser Pro Asp Ser Gln Gly Met Lys Val Pro Ser Ile Ile Trp
385                 390                 395                 400


Gln Asp Gly Ile Glu Lys Phe Pro Pro Ala Arg Leu Ser Asp Ser Asp
                405                 410                 415


Ile Arg Glu Leu Pro Lys Ile Thr Gly Phe Lys Glu Ser Arg Lys Ile
            420                 425                 430


Gln Ser Ser Gly Ile Thr Tyr Lys His Leu Arg Tyr Asp Ser Thr Asp
        435                 440                 445


Leu Ala Glu Tyr Arg Lys Gln Tyr Trp Ser Asn Thr Arg Lys Glu Leu
    450                 455                 460


Val Thr Ile Lys Ile Asp Val Asp Asp Leu Ser Lys Ile Phe Val Tyr
465                 470                 475                 480


Leu Pro Glu Ile Glu Lys Tyr Leu Thr Val Pro Cys Val Asp Gln Glu
                485                 490                 495


Tyr Thr Lys Asn Leu Ser Leu Asp Gln His Leu Ile Thr Arg Thr Tyr
            500                 505                 510


Thr Lys Lys Arg Asn Lys Leu Leu Gly Lys Ser Glu Leu Glu Leu Ala
        515                 520                 525


Lys Ser Arg Asp Glu Ile Arg Asp Ala Val Glu Gln His Asp Lys Lys
    530                 535                 540


Ala Ala Thr Ser Lys Lys Val Thr Thr Ser Lys Lys Val Ala Gln Phe
545                 550                 555                 560


Lys Gly Tyr Ser Asn Glu Ser Val Lys Asn Lys Val Ala Lys Pro Ile
                565                 570                 575


Ser Ser Asp Asp Lys Ser Arg Ala Glu His Ala Glu Met Ser Glu Leu
            580                 585                 590


Glu Ala Leu Trp Asn Ser Phe Arg Lys
```

595 600

<210> 47
<211> 207
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG110 transposition protein sequence

<220>
<223> MG110-1-A transposition protein

<400> 47
Met Tyr Asn Arg Asn Leu Arg Lys Pro Ser Pro Asn Lys Asn Leu Tyr
1 5 10 15

Lys Phe Val Ser Arg Lys Asn His Ser Thr Val Met Cys Glu Ser Gly
20 25 30

Leu Glu Phe Asp Ala Cys Phe His Leu Glu Phe Ser Pro Ser Ile Ala
35 40 45

Ser Phe Glu Ser Gln Pro Thr Gly Ile Glu Tyr Gln Ala Asp Asn Lys
50 55 60

Ile Arg Arg Tyr Thr Pro Asp Phe Lys Ile Val Lys Asn Thr Gly Glu
65 70 75 80

Ile Glu Tyr Ile Glu Val Lys Pro Glu Arg Ile His Ser Thr Lys Lys
85 90 95

Phe Arg Asp Glu Phe Glu Tyr Lys Arg Ala Ala Tyr Arg Ala Leu Gly
100 105 110

Phe Lys Leu Ile Leu Val Ser Glu Lys Gln Ile Arg Ser Glu Lys Leu
115 120 125

Leu Ser Asn Leu Lys Ile Leu His Arg Tyr Ala Ser Thr Asn Phe Ser
130 135 140

Glu Leu His Lys Leu Ala Leu Thr His Ile Lys Lys Phe Lys Ser Leu
145 150 155 160

Ser Ile Arg Glu Leu Ala Ile Lys Leu Gly Ile Arg Ile Cys Asp Cys
165 170 175

Ile Ala Ala Cys Ala Leu Leu Ile Gly Ile Gly Ala Val Lys Ala Asp
180 185 190

Leu Glu Ser Asp Phe Leu Cys Glu His Ser Leu Leu Asn Glu Ala
195 200 205

<210> 48
<211> 215
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG110 effector sequence

<220>
<223> MG110-2-6 effector

<400> 48
Met Thr Asp Lys Arg Tyr Phe Phe Ala Ile His Tyr Val Pro Ala Lys
1 5 10 15

Ala Asp Cys Gly Leu Leu Ala Gly Arg Cys Ile Ser Thr Leu His Gly
20 25 30

Tyr Leu Leu Asn His Ala Val Thr Gln Ile Gly Val Ala Phe Pro Cys
35 40 45

Trp Ser Asn Lys Ser Ile Gly Arg Thr Ile Ala Phe Val Ser Glu His
50 55 60

Ser Ser His Leu Thr Gln Phe Arg Glu Arg Thr Tyr Phe Gln Thr Met
65 70 75 80

Gln Gly Asp Gly Leu Phe Glu Leu Ser Pro Val Leu Glu Val Pro Asp
85 90 95

Asp Cys Ala Glu Val Arg Phe Ile Arg Asn Gln Asn Leu Ala Lys Leu
100 105 110

```
Phe Val Gly Glu Arg Arg Arg Arg Leu Leu Arg Ser Lys Arg Arg Val
        115                 120                 125


Ala Glu Arg Gly Glu Leu Phe Val Pro Lys Ala Pro Ala Ile Asn Gln
    130                 135                 140


Glu Ile Ala Pro Phe His Cys Ala Leu Ile Gln Ser Ala Ser Asn Ala
145                 150                 155                 160


Gln Ser Tyr Val Leu His Ile Gln Lys Gln Ser Cys Asp Ser Glu Asp
                165                 170                 175


Ser Thr Asn Thr Tyr Ser Arg Tyr Gly Leu Ala Ser Pro Asn Ser Tyr
            180                 185                 190


Arg Gly Ser Val Pro Glu Leu Ser Ser Ser Ile Ser Thr Leu Phe Pro
        195                 200                 205


Asp Ser Leu Tyr Leu Lys Arg
    210                 215


<210> 49
<211> 345
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
      MG110 effector sequence

<220>
<223> MG110-2-7 effector

<400> 49
Met Glu Leu Cys Lys His Leu Ser Tyr Ser Arg Ser Leu Ser Pro Gly
1               5                   10                  15


Lys Ala Val Phe Phe Tyr Lys Thr Ala Asp Ser Asp Phe Val Pro Leu
            20                  25                  30


Arg Val Glu Thr Ala Arg Val Val Gly Gln Lys Ser Gly Tyr Ser Glu
        35                  40                  45
```

```
Gly Phe Asp Ala Asn Phe Gln Pro Lys Asn Val Glu Arg Tyr Glu Leu
    50                  55                  60

Ala Tyr Gly Asn Pro Gln Thr Ile Glu Ala Cys Tyr Val Pro Pro Asn
65                  70                  75                  80

Ile Val Glu Leu Tyr Cys Arg Phe Ser Leu Arg Ile Glu Ala Asn Ser
                85                  90                  95

Leu Ser Pro Asp Leu Cys Ser Asp Pro Gly Val Val Lys Thr Leu Lys
            100                 105                 110

Gln Leu Ala Gly Leu Tyr Gln Glu His Gly Gly Tyr Ile Glu Leu Ala
        115                 120                 125

Arg Arg Tyr Ser Arg Asn Leu Leu Met Gly Thr Trp Leu Trp Arg Asn
    130                 135                 140

Arg His Asn Gln Gly Thr Arg Ile Glu Val Lys Thr Ser Gln Gly Ser
145                 150                 155                 160

His Phe Ala Ile Asp Asp Val Arg His Leu Asp Trp Phe Ser Glu Trp
                165                 170                 175

Pro Ala Leu Ala Glu Ala Gln Leu Asn Gln Leu Ala Ala Glu Met Ala
            180                 185                 190

Ile Ala Leu Ser Arg Arg Asp Met Phe Trp Phe Ala Asp Ile Thr Ala
        195                 200                 205

Val Met Lys Thr Gly Phe Cys Gln Glu Ile Phe Pro Ser Gln Arg Phe
    210                 215                 220

Thr Glu Arg Pro Asp Asn His Asp Glu Pro Ser Lys Lys Leu Ala Thr
225                 230                 235                 240

Val Glu Cys Leu Asp Gly Gln Leu Ala Ala Cys Leu His Ala Gln Lys
                245                 250                 255

Val Gly Ala Ala Leu Gln Gln Ile Asp Asp Trp Trp Ala Glu Asp Ala
            260                 265                 270
```

Asp Gln Pro Leu Arg Val His Gln Tyr Gly Ala Asp Ser Lys Asn Leu
        275   280   285

Thr Ser Met Arg His Pro Val Ser Gly Leu Asp Phe Tyr His Leu Leu
    290   295   300

Gly Arg Ala Asp Glu Phe Leu Glu Leu Met Arg Ser Ser Pro Ala Ser
305   310   315   320

Ser Pro Leu Pro Gly Glu Ile His Tyr Leu Met Ala Val Leu Thr Lys
                325   330   335

Gly Gly Leu Phe Gln Lys Gly Lys Gly
            340   345

<210> 50
<211> 710
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
      MG110 effector sequence

<220>
<223> MG110-2-8 effector

<400> 50
Met Pro Leu His Glu Leu Met Asn Leu His Asp Gln Ala Glu Gln Asp
1   5   10   15

Lys Ala Leu Arg Arg Ala Phe Ala Pro Tyr Thr Pro Asp Ile Glu Leu
            20   25   30

Glu Gly Glu Glu Tyr Arg Ala Leu Val Val Leu Val Asn Leu Thr Leu
        35   40   45

Lys Arg Asp Gln Ala Val Asp Leu Cys Asp Glu Ser Gln Ala Gln Arg
    50   55   60

Leu Val Asn Asp Asp Lys His Leu Glu His Cys Val His Pro Val Ala
65   70   75   80

```
Trp Leu His Ser His Asn Leu Lys Tyr Pro Asp Thr Arg Val Ser Gly
                85                  90                  95

Gln Arg Leu Val Ile Asp Ala Pro Pro Leu Ile Pro Gly Val Val Thr
            100                 105                 110

Ser Ala Gly Leu Gln Asn Arg Leu Gly Trp Ala Asn Asn Ser Ala Asp
        115                 120                 125

Ile Asn Tyr Ala Lys Leu Phe Cys Ser Ser Phe Leu Tyr Gln Gly Leu
    130                 135                 140

Ser Ser Asn Leu Ala Leu Gln Leu Val Ala Asp Val Pro Ala Trp Thr
145                 150                 155                 160

Gly Ala Phe Arg Gln Leu Gly Val Ala Asn Thr Ala Ile Ala Ala Leu
                165                 170                 175

Gln Ala Gln Leu Ala His His Leu Ala Ala Thr Ala Ile Pro Ser Glu
            180                 185                 190

Val Ser Val Tyr Ser Lys Gln Val Arg Phe Trp Tyr Gln Gly Asp Tyr
        195                 200                 205

Cys Ala Ile Thr Pro Val Val Ser His Ala Leu Met Ala His Leu Gln
    210                 215                 220

His Met Ile Tyr Glu Lys Arg Cys Ser His Leu Thr Ile Ser His Asp
225                 230                 235                 240

His Pro Ser Ser Val Gly Asn Leu Val Gly Ala Val Gly Gly Lys Ile
                245                 250                 255

Ser Val Leu Asn Tyr Pro Pro Pro Ile Ser Ile Pro Lys Arg Arg His
            260                 265                 270

Phe Ser Gln Ser Arg Thr Gln Arg Leu Asn Glu Gly Arg Ser Leu Phe
        275                 280                 285

Asp Arg Gly Met Leu Tyr Asp Arg Ile Phe Pro Gln Ala Leu Glu His
```

```
    290                 295                 300

Leu Ile Thr Pro Ala Gly Leu Thr Arg Arg Gln Arg Lys Gln Ser Arg
305                 310                 315                 320

Leu Ser Ala Met Arg Tyr Leu Arg Arg Gln Leu Val Ala Trp Ile Gly
                325                 330                 335

Pro Val Ile Glu Trp Arg Asp Glu Ile Gln Leu Leu Pro Asn Ile Glu
            340                 345                 350

Leu Pro Ile Thr Pro Glu Arg Leu Glu Trp Lys Ile Val Thr Ser Pro
        355                 360                 365

Val Ala Glu Leu Pro Asp Leu Ala Thr Glu Leu Ala Gly Val Phe His
    370                 375                 380

Leu Glu Leu Gln Thr His His Ala Thr Arg Arg Phe Ala Tyr His Pro
385                 390                 395                 400

Glu Leu Leu Ile Pro Ile Lys Ser Gln Leu Arg Trp Leu Leu Asn Lys
                405                 410                 415

Leu Ala Asp Asp Glu Asn Arg Pro Ser Pro Gln His Ala Ala Thr Ser
            420                 425                 430

Cys Tyr Leu His Leu Ser Gly Leu Arg Val Tyr Asp Ala Leu Ala Leu
        435                 440                 445

Ala Asn Pro Tyr Leu Cys Gly Ile Pro Ser Leu Ser Ala Leu Ala Gly
    450                 455                 460

Phe Cys His Asp Tyr Glu Arg Arg Leu Thr Ala Leu Leu Lys Arg Pro
465                 470                 475                 480

Val Arg Phe Thr Gly Val Ser Trp Tyr Leu Ser His Tyr His Leu Leu
                485                 490                 495

Ser Gly Lys Ser Leu Gln Glu Pro Leu Ser Pro Met His Ala Arg Glu
            500                 505                 510
```

```
Val Ser Ala Ile Arg Arg Pro Gly Leu Ile Asp Ser Lys Tyr Cys Asp
        515                 520                 525

Leu Gly Met Asp Leu Val Ile Ala Leu Gln Val Asp Asp Glu His Pro
    530                 535                 540

Leu Pro Ser Ala Asn Glu Gln Asp Leu Leu Gln Ala Ala Phe Pro Ser
545                 550                 555                 560

Arg Phe Ala Gly Gly Cys Leu His Pro Pro Ser Leu Tyr Glu Gly Lys
                565                 570                 575

Pro Trp Cys Ser Leu Tyr Thr Asn Arg Asn Glu Leu Phe Asn Thr Leu
            580                 585                 590

Ser Arg Leu Pro Arg Thr Gly Cys Trp Val Tyr Pro Tyr Gln Ser Arg
        595                 600                 605

Ile Ser Asn Leu Asp Asp Leu Ile Asp Thr Leu Ser Met Asp Arg Arg
    610                 615                 620

Val Arg Pro Val Ser Thr Gly Tyr Val Phe Leu Asp Glu Pro Arg Val
625                 630                 635                 640

Arg Ala Gly Ser Val Glu Lys Cys His Val Tyr Ala Glu Ser Ala Leu
                645                 650                 655

Gly Leu Ala Leu Cys Val Asn Pro Val Glu Met Arg Leu Tyr Gly Asn
            660                 665                 670

Asn His Phe Phe Asn His Gly Val Trp Gln Leu Thr Ala Ile Asn Arg
        675                 680                 685

Ala Ile Leu Met Thr Gly Ala Gly Asn Met Glu Gln Arg His Gly Ala
    690                 695                 700

Met Gln Thr Ser Glu Leu
705                 710
```

<210> 51

<211> 409
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG110 transposition protein sequence

<220>
<223> MG110-2-Q transposition protein

<400> 51
Met Arg Leu Leu Ile Arg Pro Glu Pro Tyr Pro Asp Glu Ser Leu Glu
1 5 10 15

Ser Tyr Leu Leu Arg Leu Ser Gln Glu Asn Gly Phe Glu Arg Tyr Thr
20 25 30

Leu Leu Ser Gly Ala Ile Arg Asp Ser Leu Leu Val Gln Asp His Glu
35 40 45

Ala Ser Gly Ala Phe Pro Leu Glu Leu Ala Arg Val Asn Val Phe His
50 55 60

Ala Gln Arg Ser Ser Gly Leu Arg Val Arg Ala Leu Gln Leu Phe Ala
65 70 75 80

Gln Leu Thr Asp Leu Asp Pro Ala Cys Leu Leu Gln Leu Ser Leu Met
85 90 95

His Ser Gly Leu His Phe Gly Ser Gly His Ala Cys Val His Arg Ala
100 105 110

Gly Val Asp Thr Pro Leu Gly Phe Ile Arg Thr Gly Pro Ile Pro Val
115 120 125

Cys Pro Ala Cys Leu Arg Glu Ser Ala Tyr Ile Arg Gln His Trp His
130 135 140

Tyr Ser Pro Tyr Leu Ala Cys His Gln His Gly Ser Gln Leu Leu His
145 150 155 160

Leu Cys Pro Ser Cys Gly Glu Ala Leu Asp Tyr Gln His Ser Glu Ser
165 170 175

```
Phe Thr His Cys Arg Cys Gly Phe Asp Leu Arg Thr Ala Ile Thr Gln
            180                 185                 190

Pro Ala Ser Ala Glu Ala Gln Gln Leu Ser Thr Leu Ile Cys Gly Ala
        195                 200                 205

Arg Phe Glu Ser Thr Asn Pro Leu Leu Ser Ser Glu His Pro Ser Leu
    210                 215                 220

Met Phe Gly Ala Leu Tyr Trp Tyr Trp Leu Arg Asn Lys Arg Ala Asp
225                 230                 235                 240

Ala Gly Gln Pro Glu Phe Gly Thr Leu Thr Met Ala Ile Asp Tyr Phe
                245                 250                 255

Ser Ala Trp Pro Ala Asn Phe His His Glu Leu Gln Gln Met Ala Gln
            260                 265                 270

Lys Ala Leu Gln Thr Gln Thr Arg Leu Leu Asn His Thr Ala Phe Arg
        275                 280                 285

Glu Val Phe Gly Ser Val Leu Thr Asp Cys Arg Gln Leu Pro Thr Arg
    290                 295                 300

Asp Ser His Arg Asn Phe Ile Leu Arg Gly Leu Leu Asp Tyr Leu Thr
305                 310                 315                 320

Glu Leu Val Cys Ala Asn Pro Lys Thr Arg Val Ala Asn Leu Ala Asp
                325                 330                 335

Ile Gln Leu Ser Ala Ile Glu Ala Ala Ala Leu Leu Gly Thr Ser Val
            340                 345                 350

Glu Gln Val Phe Arg Leu Leu Gln Asp Gly Tyr Leu Thr Pro Ala His
        355                 360                 365

Arg Arg Ile His Gly Gly Leu Lys Pro His Glu Pro Leu Phe His Leu
    370                 375                 380
```

Arg His Val Ile Glu Cys Arg Gln Ala His Arg Ser Thr Phe Asn Asp
385 390 395 400

Ala Tyr His Pro Tyr Leu Pro Ala Trp
405

<210> 52
<211> 335
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG110 transposition protein sequence

<220>
<223> MG110-2-C transposition protein

<400> 52
Met Glu Leu Ser Ser Thr Asp Ala Asp Lys Leu Lys Ser Phe Ile Asp
1 5 10 15

Cys Tyr Val Glu Thr Pro Leu Leu Arg Ile Ile Gln Asp Asp Phe Asp
20 25 30

Arg Leu Arg Tyr Asp Lys Gln Phe Ala Gly Glu Pro Ile Cys Met Leu
35 40 45

Leu Thr Gly Asp Ala Gly Thr Gly Lys Ser Ser Leu Leu Arg His Tyr
50 55 60

Met Ala Gln Ser Pro Glu Gln Ser Gly Gln Gly Phe Val Arg Lys Pro
65 70 75 80

Leu Leu Val Ser Arg Ile Pro Ser Lys Pro Thr Leu Glu Ser Thr Met
85 90 95

Val Glu Leu Leu Lys Asp Leu Gly Gln Trp Gly Ser Glu Tyr Arg Leu
100 105 110

His Arg Ser Ser Ala Glu Ser Leu Thr Glu Ala Leu Ile Lys Cys Leu
115 120 125

Lys Arg Cys Glu Thr Glu Leu Ile Ile Ile Asp Glu Phe Gln Glu Leu

```
    130                 135                 140

Ile Glu Asn Lys Thr Arg Glu Lys Arg Asn Gln Ile Ala Asn Arg Leu
145                 150                 155                 160

Lys Tyr Ile Ser Glu Thr Ala Lys Ile Pro Ile Val Leu Val Gly Met
                165                 170                 175

Pro Trp Ala Ala Lys Ile Ala Glu Glu Pro Gln Trp Ala Ser Arg Leu
            180                 185                 190

Met Val Arg Arg Thr Ile Pro Phe Phe Lys Leu Ser Glu Asp Ala Glu
        195                 200                 205

Ser Phe Val Arg Phe Val Met Gly Leu Ala Arg Arg Met Pro Phe Ala
    210                 215                 220

Thr Pro Pro Lys Leu Glu Ala Lys His Thr Ile Phe Ala Leu Phe Ala
225                 230                 235                 240

Ser Cys His Gly Ser Ile Arg Arg Leu Lys His Leu Leu Asp Glu Ser
                245                 250                 255

Val Lys Gln Ala Leu Ala Ala His Ser Glu Thr Leu Gln His Glu His
            260                 265                 270

Ile Ala Val Ala Phe Ala Leu Phe Tyr Pro Asp Gln Ile Asn Pro Phe
        275                 280                 285

Leu Gln Pro Ile Asp Glu Ile Gln Thr Cys Glu Val Lys Gln Tyr Ser
    290                 295                 300

Arg Tyr Glu Ile Asp Ala Ala Gly Lys Asp Glu Val Leu Ile Pro Thr
305                 310                 315                 320

Gln Phe Thr Asp Lys Ile Pro Ile Ser Gln Leu Leu Lys Lys Arg
                325                 330                 335

<210> 53
<211> 630
<212> PRT
```

<213> Unknown

<220>
<223> Description of Unknown:
MG110 transposition protein sequence

<220>
<223> MG110-2-B transposition protein

<400> 53
Met Met Cys Ala Gln Gln Ser Ser Glu His Pro Ser Gly Leu Phe Glu
1 5 10 15

Asp Glu Phe Ala Leu Pro Glu Pro Ala Ala Ser Ser Thr Pro Ser Glu
20 25 30

Ser Thr Val Thr Gly Ile Ser Ala Thr Phe Asp Ser Phe Pro Asp Glu
35 40 45

Leu Lys Asn Glu Ala Leu His Arg Leu Lys Tyr Leu Gln Trp Ile Glu
50 55 60

Ala Arg Leu Ile Gly Gly Trp Thr Glu Lys Asn Ile Ser Pro Leu Leu
65 70 75 80

Val Glu Ala Ala Ser Ile Leu Pro Pro Pro Val Pro Asn Trp Arg Thr
85 90 95

Leu Ala Arg Trp Arg Lys Asn Tyr Ile Gln Gln Gly Lys Lys Ile Ile
100 105 110

Ala Leu Ile Pro Arg His Gln Ala Lys Gly Asn Ser Gln Ser Arg Leu
115 120 125

Pro Leu Ser Asp Glu Ile Phe Phe Glu Glu Ala Val His Lys Tyr Leu
130 135 140

Val Asp Glu Glu Pro Ser Ile Ala Ser Ser Tyr Gln Leu Tyr Lys Ser
145 150 155 160

Lys Val Glu Leu Glu Asn Lys Thr Ile Val Gln Asn Pro Ile Lys Ile
165 170 175

```
Leu Ser Tyr Lys Ala Phe Tyr Asp Arg Ile Lys Met Leu Pro Ala Tyr
            180                 185                 190

Gln Val Met Lys Cys Arg Lys Gly Leu His Leu Ala Asn Ala Gln Phe
        195                 200                 205

Lys Ala Ile Gly Ser His Lys Arg Pro Thr Arg Ile Met Glu Arg Val
    210                 215                 220

Glu Ile Asp His Thr Pro Leu Asp Leu Ile Leu Leu Asp Asp Glu Leu
225                 230                 235                 240

Leu Val Pro Leu Gly Arg Pro Cys Leu Thr Leu Leu Ile Asp Cys Tyr
                245                 250                 255

Ser His Cys Val Val Gly Phe Asn Leu Asn Phe Asn Gln Pro Gly Tyr
            260                 265                 270

Glu Ser Val Arg Asn Ala Leu Leu Asn Ser Ile Pro Gln Lys Asn Tyr
        275                 280                 285

Ile Lys Asp Lys Tyr Pro Val Ile Glu His Glu Trp Pro Cys Tyr Gly
    290                 295                 300

Lys Pro Glu Thr Leu Val Val Asp Asn Gly Val Glu Phe Trp Ser Asn
305                 310                 315                 320

Ser Leu Glu Gln Ala Cys Leu Glu Leu Gly Ile Asn Ile Gln Tyr Asn
                325                 330                 335

Pro Val Arg Lys Pro Trp Leu Lys Pro Met Ile Glu Arg Met Phe Arg
            340                 345                 350

Thr Ile Asn Cys Lys Leu Ile Asp Pro Ile Pro Gly Lys Thr Phe Ser
        355                 360                 365

Asn Tyr Leu Glu Lys Gly Glu Tyr Asn Pro Glu Lys Asp Ala Val Met
    370                 375                 380

Arg Phe Ser Val Phe Leu Glu Ile Phe His Gln Trp Ile Ile Asp Ile
385                 390                 395                 400
```

```
Tyr His Tyr Glu Pro Asp Ser Arg His Arg Tyr Ile Pro Ile Leu Ser
               405                 410                 415

Trp Gln Tyr Gly Phe Asp Arg Leu Pro Pro Ala Lys Val Thr Gly Glu
            420                 425                 430

Asp Met Ala Lys Leu Glu Val Ile Leu Ser Leu Cys Ile Arg Cys Lys
        435                 440                 445

His Thr Arg Gly Gly Val Glu His Leu Tyr Leu Arg Tyr Asp Ser Glu
    450                 455                 460

Glu Phe Ala Ser Tyr Arg Met Lys Tyr Pro Ser Lys Thr Asp Gly Lys
465                 470                 475                 480

Gln Tyr Val Leu Val Lys Leu Asn Pro Arg Asp Ile Ser Tyr Val Tyr
               485                 490                 495

Val Phe Ile Asp Lys Ile Gly Glu Tyr Ile Arg Val Pro Cys Val Asp
            500                 505                 510

Ser Glu Gly Tyr Thr Arg Gly Leu Ser Leu Gln Ala His Lys Ile Asn
        515                 520                 525

Val Lys Leu His Arg Asp Phe Ile Gly Lys Lys Ile Asp Val Val Ser
    530                 535                 540

Leu Ala Ile Ser Arg Ala Lys Ile Glu Ser Arg Ile Ile Lys Glu Leu
545                 550                 555                 560

Thr Glu Val Arg Gln Thr Leu Lys Lys His Asn Ile Lys Gly Ile Asn
               565                 570                 575

Lys Ile Ala Lys Tyr Arg Asp Ile Gly Ser Gln Thr Ala Ala Asn Leu
            580                 585                 590

Leu Ser Ser Thr Gln Thr Pro Glu Asn Thr Asn Asp Asn Pro Val Gln
        595                 600                 605
```

Pro Lys Thr Asp Gln Leu Pro Leu Glu Asp Asp Trp Asp Ser Phe Thr
610 615 620

Ser Glu Leu Glu Pro Tyr
625 630

<210> 54
<211> 207
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG110 transposition protein sequence

<220>
<223> MG110-2-A transposition protein

<400> 54
Met Tyr Arg Arg His Leu Lys His Ser Arg Val Lys Asn Leu Phe Lys
1 5 10 15

Phe Val Ser Ala Lys Met Asn Thr Val Phe Thr Val Glu Ser Ser Leu
20 25 30

Glu Phe Asp Thr Cys Phe His Leu Glu Tyr Ser Pro Ala Val Lys Thr
35 40 45

Phe Glu Ala Gln Pro Glu Gly Tyr Tyr Tyr Ile Phe Glu Asp Arg Glu
50 55 60

Cys Pro Tyr Thr Pro Asp Phe Arg Val Val Asp Glu Arg Gly Asn Pro
65 70 75 80

Tyr Phe Ile Glu Val Lys Pro Ser Ala Lys Val Ala Thr Pro Asp Phe
85 90 95

Leu Gln Arg Phe Pro Ile Lys Gln Gln Lys Ala Ile Glu Leu Arg Ser
100 105 110

Pro Leu Lys Leu Val Thr Glu Arg Gln Ile Arg Val Asp Pro Ile Leu
115 120 125

Gly Asn Leu Lys Leu Leu His Arg Tyr Ser Gly Phe Gln Ser Phe Thr

130 135 140

Pro Leu His Ile Gln Leu Leu Gly Leu Val Arg Gln Leu Gly Arg Val
145 150 155 160

Ser Leu Ala Lys Leu Ser Gly Ser Thr Gly Ala Pro Pro Glu Glu Ile
165 170 175

Leu Ala Thr Thr Leu Ser Leu Met Ala Arg Gly Leu Ile His Ser Asp
180 185 190

Leu Thr Glu Arg Glu Met Gly Leu Ser Ser Val Val Trp Ala Ile
195 200 205

<210> 55
<211> 400
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-8 effector

<400> 55
Gln Phe Asn Ile Glu Asn Glu Thr Gly Ser Ala Pro Asn Ala Lys Lys
1 5 10 15

Arg Lys Lys Ser Lys Ser Ala Asn Asp Pro Asn Leu Met Ser Tyr Leu
20 25 30

Phe Lys Ala His Lys Ala Thr Glu Asp Ile Leu Ser His Cys Ser Ile
35 40 45

Ser His Leu Leu Lys Asn Asp Cys Lys Val Ser Glu Thr Glu Glu Asp
50 55 60

Pro Asp Lys Phe Ala His Arg Ile His Arg Lys Gln Lys Gln Ile Glu
65 70 75 80

Gln Leu Asp Ala Lys Leu Asn Ala Arg Leu Pro Lys Gly Arg Asp Leu
85 90 95

```
Thr Gly Glu Glu Phe Phe Arg Thr Leu Glu Thr Ala Thr His Gln Ile
            100                 105                 110

Ser Asp Asn Val Ile Gln Ala Arg Glu Trp Asp Ala Lys Leu Leu Thr
        115                 120                 125

Arg Pro Ala Thr Leu Pro Tyr Pro Ile Ile Tyr Gly Ser Ser Thr Asp
    130                 135                 140

Leu Arg Trp Gly Lys Thr Lys Thr Gly Arg Ile Thr Val Asn Phe Asn
145                 150                 155                 160

Gly Ile Asp Lys Tyr Leu Asn Ala Ala Asp Pro Asn Leu Lys Glu Trp
                165                 170                 175

Phe Lys Thr His Glu Glu His Pro Phe Gln Leu Tyr Cys Asp Arg Arg
            180                 185                 190

Gln Leu His Phe Phe Gln Arg Phe Leu Glu Asp Trp Gln Ala Tyr Gln
        195                 200                 205

Ala Asn Lys Asp Thr Tyr Pro Ala Gly Leu Leu Thr Leu Ser Ser Ala
    210                 215                 220

Ile Leu Ile Trp Thr Glu Cys Glu Gly Lys Gly Asp Pro Trp Asn Val
225                 230                 235                 240

Asn His Leu Ser Leu His Cys Thr Tyr Asp Thr Arg Leu Met Thr Ala
                245                 250                 255

Glu Gly Thr Leu Val Val Gln Gln Glu Lys Ser Ala Lys Ala Ala Lys
            260                 265                 270

Asn Leu Glu Arg Glu Lys Ser Asp Pro Arg Asn Gln Ser Thr Leu Asn
        275                 280                 285

Arg Leu Asn Asn Leu Pro Ala Arg Pro Ser Lys Gln Pro Tyr Gln Gly
    290                 295                 300
```

Asn Pro Glu Ile Ser Val Gly Leu Ser Ile Gly Leu Thr Asn Pro Ile
305 310 315 320

Thr Ala Ala Val Val Asn Val Thr Thr Gly Glu Val Leu Thr Tyr Arg
325 330 335

Thr Pro Lys Thr Leu Leu Gly Glu Gln Tyr Arg Leu Leu Asn Arg His
340 345 350

Arg Asp Arg Gln Gln Gln Asn Thr Leu Gln Arg Gln Lys Asn Gln Lys
355 360 365

Arg Cys Val Arg Tyr Gln Pro Ser Glu Ser Glu Leu Gly Glu Tyr Val
370 375 380

Asp Arg Leu Leu Ala Asn Lys Ile Ile Gln Leu Ala Gln Gln Tyr Gln
385 390 395 400

<210> 56
<211> 748
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-9 effector

<400> 56
Met Gly Met Lys Thr Ile Gln Ser Asp Leu Ile Val Arg Leu Asp Pro
1 5 10 15

Asn Pro Arg Arg Gly Lys Lys Asn Leu Asp Asp Leu Ser Glu Phe Asp
20 25 30

Ala Val Ala Trp Cys Arg Leu Cys Glu Arg Ala Tyr Gln His Thr Leu
35 40 45

Leu Ile Asp Gln Ile Ser Glu Gln Ile Lys Leu Asn Pro Asp Val Ile
50 55 60

Ala Trp Ile Tyr Ala Val Gln Asn Ser Glu Asp Gln Asn Ser Ala Asp

```
65                  70                  75                  80

Met Asp Thr Asp Thr Asp Ser Glu Gln Gly Tyr Ser Trp Leu Pro Asp
                85                  90                  95

Gly Leu Ile Asn Gly Leu Cys Ala Leu Leu Lys Asn Thr Pro Gln Phe
            100                 105                 110

Glu Gly Met Ser Gly Arg Ala Tyr Thr Ser Ala Thr Asp Arg Val Glu
        115                 120                 125

Glu Asn Phe Lys Gly Trp Phe Ala Asn His Gln Lys Leu Ile Arg Gln
    130                 135                 140

Ile Lys Gly Lys Gln His Trp Leu Thr Val Leu Glu Ser Asp Ala Glu
145                 150                 155                 160

Leu Ala Glu Asn Ser Asn Phe Asn Gln Ser Glu Ile Glu Asn Gln Ala
                165                 170                 175

Lys Gln Ile Leu Asn Glu Ile Glu Ala Asp Asn Glu Ser Lys Asp Glu
            180                 185                 190

Ser Val Asn Asp Asn Arg Arg Val Phe Ser Ile Leu Phe Asp Arg Phe
        195                 200                 205

Glu Ala Thr Glu Asp Val Leu Asn Arg Arg Ala Ile Ile His Leu Leu
    210                 215                 220

Lys Asn Gly Gly Lys Val Arg Trp Glu Pro Lys Lys Ser Arg Asn Arg
225                 230                 235                 240

Lys Ser Gln Lys His Pro Ser Lys Pro Met Thr Phe Glu Glu Arg Leu
                245                 250                 255

Val Ala Lys Arg Ile Glu Ile Ala Arg Leu Glu Lys Gln Leu Leu Ser
            260                 265                 270

Gln Leu Pro Arg Ala Arg Asn Leu Phe Pro Asp Ile Ala Phe Glu Glu
        275                 280                 285
```

```
Ser Leu Ala Glu Val Val Ala Leu Pro Lys Leu Ser Asp Glu Glu Gln
    290                 295                 300

Val Glu Ala Glu Gln Val Glu Ser Glu Phe Tyr Asn Trp His Glu Ser
305                 310                 315                 320

Ile Thr Phe Lys Leu Leu Glu Phe Leu Arg Glu Pro Lys Ser Leu Pro
                325                 330                 335

Tyr Pro Ile Ser Phe Gly Tyr Glu Asp Val Arg Ser Trp Gln Ile Asn
            340                 345                 350

Gln Ala Gly Lys Ile Phe Phe Lys Leu Asn Gly Trp Gly Asp Leu Ile
        355                 360                 365

Phe Glu Val Arg Cys His Arg Arg Gln Leu Pro Leu Ile Lys Ser Phe
    370                 375                 380

Leu Lys Asp Trp Gln Thr Lys Glu Gln Cys Gln Glu Gly Asp Gln Tyr
385                 390                 395                 400

Ser Gly Ser Leu Met Leu Leu Arg Ser Ile Glu Leu Val Trp Lys Pro
                405                 410                 415

Lys Pro Val Asn Glu Gln Asn Asp Ile Gln Leu Cys Ser Gln Cys Glu
            420                 425                 430

Val Phe Gln Gln Tyr Pro Ser Lys Gly Phe Trp Ser Glu Cys Lys Leu
        435                 440                 445

Ser Ile His Trp Ser Tyr Asp Ser Asp Ala Leu Ser Lys Gln Gly Leu
    450                 455                 460

Glu Lys Val Arg Gln Arg Lys Leu Glu Pro Gln Leu Glu Lys Leu Arg
465                 470                 475                 480

Lys Lys Gln Glu Glu Leu Glu Glu Lys Gln Gln Leu Leu Gln Ser Ile
                485                 490                 495

Glu Glu Val Pro Glu Ala Leu Arg Ser Lys Ala Gln Ser Lys Lys Met
```

```
            500                 505                 510

Arg Ser Leu Thr Lys Ala Ile Gln Glu Leu Gln Asp Asp Leu Ala Lys
        515                 520                 525

Pro Arg Pro Lys Leu Asp Cys Leu Gln Asn Ser Leu Leu Phe Asp Arg
    530                 535                 540

Pro Asp Arg Pro Leu His Glu Gly Val Pro Asn Ile Phe Val Gly Val
545                 550                 555                 560

Leu Leu Asp Leu Asp Lys His Leu Val Val Thr Val Val Asp Ala Met
                565                 570                 575

Arg Arg Lys Ile Leu Ala Ile Arg Asn Ala Arg Ser Ile Ser Lys Glu
            580                 585                 590

Gly Tyr Asp Leu Leu Gln Ser Tyr Phe Arg Gln Arg Arg Glu His Ser
        595                 600                 605

Lys Glu Arg Gln Ile Asp Gln Lys Ala His Arg His Val His Gln Thr
    610                 615                 620

Glu Ser Asn Leu Gly Gln His Val Ala Arg Leu Phe Ala Lys Gly Ile
625                 630                 635                 640

Val Glu Leu Gly Gln Lys Tyr Lys Ala Ser Thr Ile Val Ile Pro Glu
                645                 650                 655

Thr Asp Gly Trp Arg Asp Arg Leu Tyr Ser Gln Leu Val Ala Ser Ala
            660                 665                 670

Ala Ile Lys Cys Lys Gly Val Lys Lys Ala Met Thr Leu Tyr Thr Lys
        675                 680                 685

Gln His Gly Glu Lys Leu His Gln Trp Asp Tyr Asn Arg Leu Ser Gln
    690                 695                 700

Ala Ile Ser Asp Arg Ala Ala Thr Asp Gly Leu Lys Val Met Arg Gln
705                 710                 715                 720
```

Asn Thr Val Tyr Glu Glu Asp Ala Phe Gln Gln Ala Ala Asn Leu Ala
725 730 735

Ile Ala Ala Tyr Asp Ser Leu Asn Ser Val Glu Lys
740 745

<210> 57
<211> 422
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-10 effector

<400> 57
Ser Trp Gln Leu Asn Gln Glu Gly Lys Ile Cys Phe Lys Leu Asn Gly
1 5 10 15

Leu Gly Asp Tyr Leu Phe Glu Val Arg Cys Asp Arg Arg Gln Leu Gly
20 25 30

Ile Val Lys Tyr Phe Leu Gln Asp Trp Gln Thr Gln Asn Lys Asn Lys
35 40 45

Asn Glu Tyr Ser Gly Gly Leu Thr Leu Leu Arg Ser Ala Glu Leu Leu
50 55 60

Val Lys Pro Lys Leu Gly Lys Gln Asn Ala Lys Leu Pro Pro Ile His
65 70 75 80

Asp Arg Gln Ala Val Val Thr Ala Tyr Lys Leu Ser Leu His Cys Thr
85 90 95

Tyr Asp Thr Asp Tyr Leu Thr His Gln Gly Leu Glu Cys Val Arg Gln
100 105 110

Arg Lys Ile Ala Asn Gln Leu Lys Gly Leu Thr Asp Lys Lys Ala Lys
115 120 125

```
Leu Thr Lys Gln Gln Glu Gln Leu Gln Gln Leu Glu Gln Glu Met Gln
    130                 135                 140

Gln Glu Gln Ile Gly Thr Ser Ala Lys Arg Ser Lys Arg His Ala Gln
145                 150                 155                 160

Arg Leu Lys Gln Ile Glu Gln Leu Lys Gln Ser Ile Ser Lys Leu Gln
                165                 170                 175

Ala Ala Ile Gln Ala Glu Leu Glu Arg Pro Arg Pro Lys Leu Glu Arg
            180                 185                 190

Leu Gln Gln Ser Gln Leu Phe Gln Arg Ala Asp Arg Pro Leu Tyr Ala
        195                 200                 205

Gly Val Ala His Leu Phe Val Gly Val Cys Leu Asp Leu Asp Gln His
    210                 215                 220

Leu Val Val Thr Ile Val Asp Ala Met Arg His Lys Val Leu Thr Lys
225                 230                 235                 240

Arg Thr Val Lys Gln Ile Met Gly Glu His Tyr Pro Leu Leu Gln Arg
                245                 250                 255

Tyr Arg Arg Leu Lys Gln Gln His Pro Lys Gln Arg Arg Gln Asp Gln
            260                 265                 270

Lys Val Gly Arg His Asn His Leu Ser Glu Thr Gly Leu Gly Glu Gln
        275                 280                 285

Val Ala Cys Ala Ile Ala Asn Gly Leu Leu Ser Leu Ala Gln Gln Tyr
    290                 295                 300

Lys Val Ser Thr Ile Val Leu Pro Glu Thr Lys Gly Trp Arg Glu Arg
305                 310                 315                 320

Leu Tyr Ser Gln Leu Val Ala Arg Ala Lys Ile Lys Cys Asn Gly Ser
                325                 330                 335

Lys Lys Ala Met Ala Arg Tyr Thr Lys Ala Tyr Gly Lys Arg Leu His
            340                 345                 350
```

Gln Trp Asp Tyr Asn Arg Leu Ser Arg Ala Ile Glu Thr Glu Ala Gln
355 360 365

Thr Val Gly Val Thr Val Ile Phe Gln Arg Leu Glu Phe Gln Ala Asn
370 375 380

Ala Glu Gln Asp Asn Gln Pro Ala Asp Glu Ala Asp Glu Gln Asp Asn
385 390 395 400

Gln Arg Val Asn Pro Phe Glu Leu Ala Leu Gln Ile Ala Ile Ala Ala
405 410 415

Tyr Asp Ser Leu Gln Ala
420

<210> 58
<211> 404
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-11 effector

<400> 58
Met Ala Glu Ser Thr Leu Trp Glu Lys Ser Glu Pro Glu Pro Leu Ala
1 5 10 15

Gln Leu Pro Thr Leu Pro His Thr Leu Pro Tyr Gln Ile Arg Phe Glu
20 25 30

Ser Ser Asp Asp Leu Tyr Trp Ser Leu Leu Pro Lys Gly Lys Gln Glu
35 40 45

Asp Ala Glu Lys Asn Glu Gly Lys Phe Glu Gln Ala Glu Pro Leu Pro
50 55 60

Arg Pro Ala Pro Lys Arg Lys Arg Arg Lys Arg Arg Lys Ile His Tyr
65 70 75 80

```
Gln Ala Ser Ile Cys Val Arg Phe Lys Gly Leu Ala Glu His Ser Phe
                85                  90                  95

Arg Val Gln Cys Ser His Arg Gln Leu Pro Ile Phe Gln Gln Cys Leu
            100                 105                 110

Asp Glu Trp Gln Leu Leu Asn Ala Leu Ser Asp Asp Val Lys Phe Ser
        115                 120                 125

Leu Ala Val Phe Pro Leu Arg Ser Ala Arg Leu Val Trp Cys Lys Asp
    130                 135                 140

Lys Gln Ala Ser Lys His Ser Lys Asn Pro Lys Ala Glu Trp Asn Gln
145                 150                 155                 160

Tyr Arg Leu Tyr Leu His Cys Thr Ile Asp Arg Arg Thr Leu Thr Ala
                165                 170                 175

Glu Gly Thr Glu Glu Leu Arg Gln Gln Lys Leu Val Glu Ala Gln Lys
            180                 185                 190

Gln Lys Glu Ala Ser Val Lys Lys Gln Lys Pro Pro Lys Ile Pro Ala
        195                 200                 205

Pro Leu Thr Glu Glu Gln Arg Gln Ala Lys Asn Gln Ala Glu Ala Leu
    210                 215                 220

Lys Arg Asn His Ser Thr Leu Leu Arg Leu Gln Asn Pro Ser Pro Pro
225                 230                 235                 240

Arg Leu Ser Lys Pro Cys Tyr Val Gly Gln Pro His Leu Val Val Gly
                245                 250                 255

Val Cys Phe Ser Arg Arg Asp Arg Ala Ser Ala Ile Val Tyr Asp Ile
            260                 265                 270

Gln Gln His Gln Glu Ile Glu Tyr Val Gly Val Arg Glu Leu Leu Thr
        275                 280                 285

Asp Glu Ser Ala Glu Glu His Tyr Arg His Phe Ser Ser Asn Pro Gln
```

```
    290                 295                 300

Arg Lys Gly Lys Arg Ser Leu Glu Gln Met Arg Leu Glu Gln Tyr Arg
305                 310                 315                 320

Leu Val Glu Arg Phe Tyr Arg Glu Gln Glu Lys Tyr Gln Gln Arg Arg
                325                 330                 335

Ala Glu Ala Gln Glu Cys Gly Leu Glu Leu Glu Asp Glu Ser Val Ser
            340                 345                 350

Asn Leu Gly Glu Tyr Leu Asp Arg Leu Leu Ala Ala Arg Ile Val Glu
        355                 360                 365

Phe Ala Val Lys Gln Gln Ala Ser Ser Ile Val Ile Pro Glu Leu Gly
    370                 375                 380

Asn Ile Arg Glu Ser Ile Glu Cys Glu Val Gln Met Gln Met Glu Asn
385                 390                 395                 400

Arg Gly Phe Glu
```

<210> 59
<211> 731
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-12 effector

<400> 59

```
Met Ala Ser Gly Ile Leu Arg Thr Ile Gln Ala Arg Leu Val Pro Ser
1               5                   10                  15

Glu Trp Met Pro Leu Asp Gln Lys Ala Ser Glu Asp Ser Gly Glu Ser
            20                  25                  30

Gln Pro Val Pro Ile Ala Val Val Ser Ser Phe Val Glu Asp Cys Thr
        35                  40                  45
```

```
Leu Gln Leu Gly Leu Thr Ser Thr Arg Leu Cys Arg Ala Leu Asn Glu
    50                  55                  60

Cys Ile Asn Gln His Ser Asp Phe Glu Val Trp Lys Glu Gly Gly Ser
65                  70                  75                  80

Ile Pro Pro Lys Ala Leu Lys Glu Ile Trp Asn Ser Ile Arg Thr Met
                85                  90                  95

Pro Pro Tyr Gln Asp Leu Pro Glu Arg Phe Leu Arg Ser Ala Trp Leu
            100                 105                 110

Arg Val Glu Gly Ile Tyr Ala Ser Trp Phe Ala Ile Arg Thr Arg Leu
        115                 120                 125

Ala Ala Lys Gln Asp Gly Leu Asn Arg Trp Leu Ser Ile Val Arg Ser
    130                 135                 140

Asp Glu Tyr Leu Thr Lys Thr Cys Asp Cys Ser Leu Glu Glu Met Gln
145                 150                 155                 160

Ala Arg Ala Thr Gln Leu Leu Thr Glu Pro Ser Gly Ser Ser Leu Ser
                165                 170                 175

Ala Leu Phe Asp Ala Tyr Phe Ala Leu Val Lys Ala Asp Gly Ser Leu
            180                 185                 190

Leu Asp Arg Cys Ala Ile Val His Leu Ile Lys Asn Gly Cys Ala Val
        195                 200                 205

Ala Ser Lys Pro Glu Asp Leu Lys Lys Phe Ala Ala Thr Cys Gln Lys
    210                 215                 220

Lys Arg Gln Gln Val Glu Arg Ile Thr Lys Gln Leu Ala Ala Pro Ala
225                 230                 235                 240

Pro Gln Val Arg Asp Leu Gly Glu Ala Ala Glu Ala Ala Leu Thr Asp
                245                 250                 255
```

```
Ser Val Glu Arg Ile Pro Ser Asn Asn Ala Asp Phe Val Ala Gln Leu
            260                 265                 270

Ala Ala Leu Gln Arg Lys Pro Ile Pro Leu Pro Tyr Ala Val Val Phe
        275                 280                 285

Tyr Ser Asn Asp Asp Leu Glu Trp His Gln Ile Lys Arg Arg Asn Ser
    290                 295                 300

Ala Thr Glu Lys Ile Glu Glu Arg Ile Phe Val Lys Phe Lys Gly Leu
305                 310                 315                 320

Asn Lys Tyr Leu Arg Asn Arg Val Lys Tyr Asn Leu Glu Ala Gln Leu
                325                 330                 335

Ala Gln Ile Leu Glu Ser Ser Gly Leu Lys Leu Asp Asp Ile Ser Trp
            340                 345                 350

Lys Ile Ile Arg Arg Lys Thr Ala Gln Ser Thr Val Lys Tyr Ile Ser
        355                 360                 365

Phe Lys Leu Lys Thr Pro Asp Pro Thr Leu Lys Gln His Leu Lys Glu
    370                 375                 380

Gln Phe Glu Gln Ile Asp Leu Glu Asp Lys Tyr Asn Leu Gln Lys Glu
385                 390                 395                 400

Tyr Gly Phe Glu Leu Cys Cys Gly Tyr Arg Gln Leu Ala Asp Phe Gln
                405                 410                 415

Thr Phe Leu Ser Asp Trp Gln Leu Tyr Cys Thr Asn Arg Asp Arg Tyr
            420                 425                 430

Ser Val Ser Pro Phe Ala Leu Arg Ser Ala Ala Leu Asp Trp Gln Lys
        435                 440                 445

Glu Val Lys Gln Gly Lys Ser Lys Leu Gln Val Tyr Leu Arg Cys Thr
    450                 455                 460

Leu Asp Gln Gln Glu Leu Thr Ala Glu Gly Ala Glu Arg Ala Arg Ala
465                 470                 475                 480
```

```
Lys Ala Ala Thr Glu Leu Asn Lys Lys Ile Ala Lys Leu Glu Ala Ala
                485                 490                 495

Gln Glu Gln Gly Glu Val Leu Ser Glu Glu Gln Gln Arg Asp Leu Glu
            500                 505                 510

Arg Ala Arg Ser Gln Leu Arg Gly Ala Glu His Pro Tyr Pro Arg Pro
        515                 520                 525

Ser Gln Pro Leu Tyr Glu Gly Asn Pro Asp Ile Leu Val Gly Val Cys
    530                 535                 540

Leu Ser Trp Glu Gln Val Ala Thr Val Ala Val Val Asn Trp Ser Thr
545                 550                 555                 560

Arg Gln Val Leu Val Tyr Arg Thr Ile His Gln Leu Met Gly Glu Asp
                565                 570                 575

Tyr Pro Leu Leu Ser Ser Tyr Arg His Glu Gln Pro Arg Asn Ala Asn
            580                 585                 590

Lys Arg His Lys Arg Gln Lys Arg Gly Gln Ser Ser Asn Ile Ser Glu
        595                 600                 605

Ala Ser Lys Gly Lys His Ile Asp Arg Val Ile Ala Lys Arg Leu Val
    610                 615                 620

Glu Val Val Val Glu Phe Asn Ala Gly Cys Leu Val Leu Pro Asn Leu
625                 630                 635                 640

Asn Gly Leu Arg Glu Asn Leu Gln Ala Glu Leu Met Ala Lys Ala Ala
                645                 650                 655

Trp Lys His Pro Gly Asp Val Ala Lys Gln Lys Ala Tyr Leu Lys Gln
            660                 665                 670

Tyr Lys Lys Asn Leu His Cys Trp Pro Tyr Arg Arg Leu Ala Gln Phe
        675                 680                 685
```

Val Lys Glu Arg Ala Asp Lys Ile Gly Val Pro Ile Glu Leu Gly Gln
690 695 700

Gln Pro Arg Glu Gly Asp Phe Lys Glu Gln Ala Val Gln Val Ala Trp
705 710 715 720

Ser Phe His Glu Ala Gln Gln Ala Asp Glu Gly
725 730

<210> 60
<211> 741
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-13 effector

<400> 60
Met Ala Ile Lys Thr Leu Arg Ser Gln Leu Cys Ala Thr Glu Glu Thr
1 5 10 15

Arg Glu Tyr Leu Trp His Leu Phe Ile Thr Tyr Thr Leu Leu Ile Asn
20 25 30

Gln Leu Leu Glu Arg Leu Pro Lys Asp Ala Lys Phe Ser Asp Trp Arg
35 40 45

Lys Glu Gly Arg Val Pro Arg Lys Ala Val Ala Ala Leu Cys Asn Lys
50 55 60

Val Leu Glu Glu Asp Gly Ser Leu Lys Gly Leu Pro Ala Arg Phe Tyr
65 70 75 80

Thr Ser Ala Val Leu Cys Thr Ser Tyr Thr Phe Ala Ser Ile Phe Ala
85 90 95

Ile Gln Asn Lys Leu Arg Ala Arg Cys Glu Gly Lys Gln Arg Trp Leu
100 105 110

Asn Val Ala Glu His Asp Leu Glu Leu Ala Glu Thr Thr Ala Phe Ser

```
        115                 120                 125

Leu Glu Glu Ile Arg Glu Arg Ala Ala Gln Val Leu Gly Gln Ala Glu
    130                 135                 140

Ser Glu Arg Gln Glu Cys Leu Asp Glu Glu Glu Glu Pro Ser Asn Ile
145                 150                 155                 160

Pro Ser Ile Met Ser Leu Leu Phe Lys Leu Leu Asp Lys Ser Asn Asp
                165                 170                 175

Leu Leu Ser Tyr Arg Ala Ile Ala His Leu Leu Arg Asn Asp Cys Lys
            180                 185                 190

Val Ser Tyr Glu Glu Glu Asp Pro Glu Lys Leu Ala Leu Arg Leu Ser
        195                 200                 205

Lys Lys Arg Ile Gln Ile Glu Arg Leu Gln Glu Gln Leu Asn Ser Gln
    210                 215                 220

Leu Pro Ile Gly Pro Asp Pro Thr Gly Glu Arg Thr Asn Gln Phe Ile
225                 230                 235                 240

Glu Glu Ala Ile Phe Leu Pro Arg Asp Ser Asp Ser Ser Glu Asn Arg
                245                 250                 255

Ile Lys Glu Phe Glu Thr Trp Glu Asp Leu Leu Pro Glu Arg Ile Ala
            260                 265                 270

Asn Leu Ser Thr Gln Leu Glu Thr Leu Pro Tyr Pro Leu Leu Leu Gly
        275                 280                 285

Ser Asn Asp Asp Val Ser Trp Ser Trp Glu Thr Ile Ser Thr Ala Ala
    290                 295                 300

Ser Ser Ser Ser Thr Ser Pro Asn Pro Ser Pro Lys Arg Thr Arg Thr
305                 310                 315                 320

Arg Arg Arg Arg Arg Lys Ala Lys Lys Arg Ile Cys Val Ser Phe Lys
                325                 330                 335
```

```
Gly Lys Gly Leu Ser Gln Ile Arg Leu Arg Leu Tyr Cys Asp Arg Arg
            340                 345                 350

Gln Leu Pro Ile Phe Arg Gln Phe Val Glu Asp Ser Lys Ala Asn Lys
        355                 360                 365

Ala Arg Asn Lys Lys Glu Lys Phe Ser Leu Ala Leu Tyr Pro Leu Arg
    370                 375                 380

Ser Ala Thr Leu Leu Trp Ala Glu Asp Pro Lys Leu Ser Thr Glu Lys
385                 390                 395                 400

Asn Arg Asp Lys Pro Trp Asn Ser His Arg Leu Tyr Leu His Val Thr
                405                 410                 415

Val Asp Thr Arg Leu Leu Thr Ala Glu Gly Thr Glu Gln Val Arg Gln
            420                 425                 430

Glu Lys Ile Glu Leu Met His Lys Phe Leu Lys Gly Leu Val Glu Ala
        435                 440                 445

Glu Glu Thr Gln Glu Ala Asn Glu Asp Glu Glu Ala Phe Leu Val Ala
    450                 455                 460

Gln Lys Asn Arg Ser Arg Ala Thr Lys Lys Asn Gln Thr Thr Leu Lys
465                 470                 475                 480

Leu Leu Glu Ser Asn Pro Pro Pro Pro Arg Pro Arg Gln Ser Ala Tyr
                485                 490                 495

Arg Gly Asn Ala Glu Ile Ala Val Lys Val Ala Phe Ser Arg Glu His
            500                 505                 510

Val Val Gly Val Ala Val Ser Asp Gly Gln Gln Pro Ile Leu Asp Tyr
        515                 520                 525

Arg Asp Val Glu Thr Leu Leu Ile Asp Ser His Leu Glu Leu Leu Glu
    530                 535                 540

Gln Arg Ser His Lys Leu Arg Asn Gln Pro Glu Arg Leu Arg Lys Val
```

```
545                 550                 555                 560

Lys Leu Gly Asp Ser Lys Lys Ala Lys Lys Ala Lys Arg Thr Arg Tyr
                565                 570                 575

Lys Pro Lys Ile Ser Thr Arg Gln Leu Gln Leu Gln Pro Tyr Arg Leu
            580                 585                 590

Leu Lys Arg Trp His Arg Leu Lys Arg Lys Asn Val Thr Glu Arg Arg
        595                 600                 605

Ala Glu Gln Lys His Gly Leu Tyr Arg Gln Ser Gln Ala Glu Thr Asn
    610                 615                 620

Leu Ala His Gln Ile Asn His Ile Leu Ala Arg Asn Ile Val Asp Leu
625                 630                 635                 640

Cys Gln Lys Trp Ser Ala Gly Thr Ile Ile Leu Pro Glu Phe Gly Asp
                645                 650                 655

Leu Arg Glu Ser Ile Glu Ser Glu Ile Gln Ala Lys Ala Arg Arg Lys
            660                 665                 670

Tyr Pro Asp Asp Asn Val Glu Arg Gln Lys Gln Tyr Ala Lys Glu Phe
        675                 680                 685

Arg Met Gln Phe His Arg Trp Asn Tyr Lys His Leu Ser Glu Cys Ile
    690                 695                 700

Arg Leu Arg Ala Ala Lys Val Gly Ile Ala Cys Ile Ala Gly Gln Gln
705                 710                 715                 720

Pro Arg Leu Gly Thr Leu Arg Glu Lys Ala Ile Ala Val Thr Ala Ile
                725                 730                 735

Pro Pro Lys Pro Lys
            740
```

<210> 61
<211> 642
<212> PRT

<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-14 effector

<400> 61
Met Ser Ile Ile Thr Ile His Cys Arg Leu Ile Ala Ser Glu Pro Ile
1 5 10 15

Arg Arg His Leu Trp Gln Leu Met Ser Arg Ser Asn Thr Pro Leu Ile
20 25 30

Asn Asp Leu Leu Lys Gln Val Ser His His Ala Asp Phe Glu Thr Trp
35 40 45

Gln Thr Arg Gly Thr Val Pro Ser Asn Ala Ile Arg Asp Leu Cys Glu
50 55 60

Pro Leu Lys Glu Val Tyr Pro Gly Gln Pro Ala Arg Phe Tyr Ala Ser
65 70 75 80

Ala Ile Leu Met Val Thr Tyr Thr Tyr Glu Ser Trp Leu Ala Leu Gln
85 90 95

Gln Thr Arg Arg Arg Arg Leu His Gly Lys Gln Arg Trp Leu Asn Val
100 105 110

Val Lys Ser Asp Ala Glu Leu Leu Gly Leu Ser Gly Ser Thr Leu Glu
115 120 125

Ser Ile Arg Gln Arg Ala Gln Asp Ile Leu Ser Gln Leu Ser Thr Glu
130 135 140

Met Glu Thr Gln Ser Ala Pro Asn Pro Lys Lys Arg Ser Gln Arg Gln
145 150 155 160

Gly Gln Thr His Ser Ser Asn Asp Ala Ser Leu Met Pro Arg Leu Phe
165 170 175

```
Thr Ala Tyr Asp Thr Ala Asp Asp Ile Leu Ser Gln Cys Ala Ile Val
            180                 185                 190

His Leu Ile Lys Asn Gly Cys Lys Ile Ser Glu Thr Glu Glu Asp Ser
        195                 200                 205

Glu Lys Phe Ala His Arg Ile His Arg Lys Gln Lys Glu Ile Asp Gln
    210                 215                 220

Ile Glu Ala Gln Leu Gln Ala Arg Leu Pro Lys Gly Arg Asp Leu Thr
225                 230                 235                 240

Gly Asp Val Phe Leu Glu Thr Leu Ala Ile Ala Thr Gln Gln Ile Pro
                245                 250                 255

Glu Thr Val Ile Gln Ala Arg Glu Trp Gln Ala Lys Leu Leu Ser Arg
            260                 265                 270

Pro Ala Ser Leu Pro Tyr Pro Met Ile Tyr Gly Ser Ser Thr Asp Val
        275                 280                 285

Arg Trp Gly Lys Thr Ala Asn Asp Arg Ile Ala Val Ser Phe Asn Gly
    290                 295                 300

Ile Asp Lys Tyr Leu Lys Glu Ala Asp Pro Glu Ile Gln Glu Trp Phe
305                 310                 315                 320

Lys Leu His Lys Glu Tyr Pro Phe Arg Val Tyr Cys Asp Gln Arg Gln
                325                 330                 335

Leu Pro Phe Phe Gln Arg Phe Leu Glu Asp Trp Gln Ala Tyr Gln Ala
            340                 345                 350

Asn Lys Asp Thr Tyr Pro Ala Gly Leu Leu Thr Leu Ser Ser Ala Thr
        355                 360                 365

Leu Ala Trp Arg Glu Gly Glu Gly Lys Gly Glu Pro Trp Glu Ala Asn
    370                 375                 380

His Leu Ala Leu Tyr Cys Ala Phe Asp Thr Arg Leu Met Ser Ala Glu
385                 390                 395                 400
```

```
Gly Thr Leu Glu Ile Gln Gln Glu Lys Ser Thr Lys Ala Leu Lys Asn
                405                 410                 415

Leu Thr His Asp Asn Pro Asp Pro Arg Asn Gln Ser Thr Leu Asn Arg
            420                 425                 430

Leu Lys Asn Val Pro Asp Arg Pro Ser Arg Lys Pro Tyr Arg Gly Asn
        435                 440                 445

Pro Glu Ile Leu Val Gly Leu Ser Ile Gly Leu Ala Asp Pro Val Thr
    450                 455                 460

Ala Ala Val Val Asn Gly Arg Thr Gly Glu Val Leu Thr Tyr Arg Ser
465                 470                 475                 480

Pro Arg Thr Leu Leu Ala Glu His Tyr His Leu Leu Asn Arg Tyr Arg
                485                 490                 495

Gln Gln Gln Gln Gln Asn Ala Leu Gln Arg His Arg Asn Gln Lys Arg
            500                 505                 510

Gly Val Thr Tyr Gln Pro Ser Glu Ser Glu Leu Gly Gln Tyr Val Asp
        515                 520                 525

Arg Leu Leu Ala Asn Ser Ile Ile Gln Leu Ala Gln Thr His Gln Ala
    530                 535                 540

Gly Ser Ile Val Ile Pro Ser Leu Thr His Val Arg Glu Leu Leu Ala
545                 550                 555                 560

Ser Glu Ile Thr Ala Lys Ala Glu Arg Lys Ser Arg Ile Val Glu Val
                565                 570                 575

Gln Asp Lys Tyr Ala Lys Glu Tyr Arg Ile Ala Ile His Arg Trp Ser
            580                 585                 590

Tyr Asn Arg Leu Ile Asn Ala Val Cys Ser Lys Ala Gln Gln Phe Gly
        595                 600                 605
```

Met Thr Val Glu Ser Gly Phe Gln Pro Leu Gln Gly Asn Pro Gln Glu
610 615 620

Gln Ala Lys Asp Ile Ala Ile Ala Ala Tyr Tyr Ala Arg Thr Ile Ala
625 630 635 640

Ile Lys

<210> 62
<211> 828
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-15 effector

<400> 62
Met Ser Phe Lys Thr Ile Arg Ser Leu Leu Cys Ala Thr Glu Glu Thr
1 5 10 15

Arg Glu Tyr Leu Trp His Leu Phe Leu Asn His Thr Leu Leu Ile Asn
20 25 30

Gln Leu Leu Glu Lys Leu Pro Gln Lys Thr Ala Phe Ala Lys Trp Arg
35 40 45

Lys Ser Gly Arg Ile Pro Arg Lys Ala Ile Val Ala Ala Cys Asn Thr
50 55 60

Ser Leu Gln Glu Asp Glu Ser Leu Glu Gly Leu Pro Ala Arg Phe Tyr
65 70 75 80

Thr Ser Ala Val Leu Ser Val Ser Tyr Thr Phe Ala Phe Ile Phe Ala
85 90 95

Ile Gln Asn Gly Leu Arg Ala Lys Val Glu Gly Lys Gln Arg Trp Leu
100 105 110

Asn Val Ala Glu Lys Asp Leu Glu Leu Ala Gln Thr Thr Ser Phe Ser

```
        115                 120                 125

Pro Glu Ala Ile Arg Ala Val Ala Ala Gln Val Val Asp Gln Ala Glu
    130                 135                 140

Ala Glu Arg Gln Arg Glu Phe Glu Ala Arg Glu Asn Arg Thr Gly Gln
145                 150                 155                 160

Pro Leu Pro Leu Met Thr Val Leu Phe Lys Phe Trp Asn Lys Thr Lys
                165                 170                 175

Lys Pro Leu Arg Arg Arg Ala Ile Ala His Leu Leu Arg Asn Asp Cys
            180                 185                 190

Gln Val Asn Ser Glu Ala Glu Asp Pro Asp Lys Leu Thr Leu Arg Leu
        195                 200                 205

Ser Lys Lys Arg Ile Gln Ile Glu Arg Leu Gln Lys Gln Leu Asp Ser
    210                 215                 220

Gln Leu Pro Thr Gly Arg Asp Pro Leu Gly Asp Arg Ile Glu Lys Phe
225                 230                 235                 240

Ile Glu Asp Ala Ile Ala Phe Ala Glu His Ser Thr Phe Val Pro Thr
                245                 250                 255

Gln Phe Trp Leu Lys Trp His Arg Leu Leu Leu Asp Val Gln Ser Ser
            260                 265                 270

Glu Pro Met Leu Asp Leu Trp Phe Phe Gly Trp Thr Tyr Tyr Arg Leu
        275                 280                 285

Asn Ala Ile Ala Glu Phe Glu Ala Trp Glu Gln Ser Leu Ser Glu Arg
    290                 295                 300

Thr Ala Asn Leu Ser Thr Gln Leu Lys Ser Leu Pro Tyr Pro Leu Leu
305                 310                 315                 320

Phe Gly Ser Thr Asp Asp Leu Tyr Trp Ser Trp Glu Lys Gln Val Gln
                325                 330                 335
```

```
Pro Lys Thr Pro Gln Pro Ala Pro Ala His Pro Gln Ala Ser Ala Asp
            340                 345                 350

Pro Thr Ile Lys Arg Lys Pro Lys Arg Cys Arg Thr Arg Lys Arg Lys
        355                 360                 365

Lys Lys Ser Thr Pro Arg Ile Cys Val Ser Phe Lys Ser Arg Gly Leu
    370                 375                 380

Ser His Leu Cys Leu Arg Leu Tyr Cys Asp Arg Arg Gln Leu Pro Val
385                 390                 395                 400

Phe Arg Gln Leu Val Glu Glu Ser Glu Ala Asn Lys Ala Arg Lys Lys
                405                 410                 415

Pro Asp Arg Phe Ser Leu Ala Leu Ser Pro Leu Arg Ser Ala Gly Leu
            420                 425                 430

Met Trp Val Glu Asp Pro Gln Gln Leu His Lys Lys Lys His Trp Lys
        435                 440                 445

Leu Lys Asn Leu Trp Leu Lys Trp Phe Cys Glu Met Ser Ser Gln Ala
    450                 455                 460

Phe Pro Asp Glu Ile Arg Glu Gln Trp Phe Gln Ser Leu Val Tyr Leu
465                 470                 475                 480

Cys Leu Ser Leu Pro Trp Gln Thr Ala Ala Ile Asp Pro Arg Leu Leu
                485                 490                 495

Thr Ala Glu Gly Thr Glu Val Val Arg Gln Glu Lys Ile Thr Leu Met
            500                 505                 510

Gly Lys Phe Leu Glu Gly Ser Lys Lys Ala Asp Glu Lys Gln Ala Ile
        515                 520                 525

Glu Glu Leu Ser Lys Lys Lys Gln Ala Lys Leu Leu Thr Ala Glu Lys
    530                 535                 540

Asn Arg Cys Lys Val Thr Lys Lys Asn Gln Thr Thr Leu Thr Leu Leu
```

```
545                 550                 555                 560

Gln Ser Asn Ser Pro Leu Pro Arg Pro Ser Arg Val Ala Tyr Gln Gly
                565                 570                 575

Asn Pro Asp Ile Ala Val Lys Val Ala Phe Ser Arg Glu His Ile Ala
            580                 585                 590

Gly Val Ala Val Ser Asp Gly Cys Gln Pro Val Leu Asp Tyr His Asp
        595                 600                 605

Ile Lys Ser Leu Leu Val Asp Pro Arg Val Asp Leu Leu Glu Glu Arg
    610                 615                 620

Ser His Lys Leu Arg Asn Gln Pro Glu Arg Leu Trp Lys Ala Thr Leu
625                 630                 635                 640

Ala Thr Gln Lys Ser Lys Pro Ser Lys Ser Lys Tyr Pro Arg Ser Lys
                645                 650                 655

Pro Lys Ile Ser Thr Arg Gln Leu Gln Leu Gln Pro Tyr Arg Leu Leu
            660                 665                 670

Lys Arg Trp Arg Arg Leu Lys Arg Glu Asn Gly Ala Lys Arg Gln Leu
        675                 680                 685

Glu Gln Arg His Gly Leu Tyr Arg Pro Ser Gln Ala Glu Ser Asn Leu
    690                 695                 700

Ala Gln Tyr Ile Asn His Leu Leu Ala Arg Asn Ile Val Asp Leu Cys
705                 710                 715                 720

Gln Arg Trp Ser Ala Gly Ser Ile Ile Leu Pro Lys Phe Gly Asp Leu
                725                 730                 735

Arg Glu Ser Ile Glu Ser Glu Ile Gln Ala Lys Ala Lys Arg Lys Tyr
            740                 745                 750

Pro Asp Asp Asn Val Glu Arg Gln Lys Gln Tyr Ala Lys Glu Phe Arg
        755                 760                 765
```

Met Glu Phe His Arg Trp Asn Tyr Gln His Leu Thr Gln Cys Ile Arg
    770   775   780

Ser Arg Ala Ala Thr Val Gly Val Lys Cys Val Ala Gly Gln Gln Pro
785   790   795   800

Arg Leu Gly Thr Leu Arg Glu Lys Ala Thr Leu Ala Val Ser Ala Glu
                805   810   815

Gln Arg Val Val Thr Ala Ile Pro Pro Lys Pro Lys
            820   825

<210> 63
<211> 601
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
      MG64 effector sequence

<220>
<223> MG64-16 effector

<400> 63
Met Thr Lys Leu Thr Ile His Cys Arg Leu Val Ala Glu Glu Glu Thr
1   5   10   15

Arg Arg Gln Val Trp Glu Met Met Ala Gly Arg Tyr Ala Pro Leu Ile
            20   25   30

Ala Thr Thr Leu Glu Gln Val Ser Gln His Lys Asp Phe Pro Gln Trp
        35   40   45

Val Ala Thr Gly Glu Leu Pro Thr Gln Val Val Lys Asn Leu Val Asn
    50   55   60

Gln Ala Gln Ser Gly Leu Pro Ala Arg Trp Cys Ala Ser Ala Gln Arg
65   70   75   80

Gln Val Gln Glu Thr Tyr Lys Ala Trp Leu Thr Lys Arg Arg Lys Leu
                85   90   95

```
Gln Gln Lys Leu Gln Gly Gln Gln Thr Trp Leu Ser Val Leu Arg Pro
            100                 105                 110

Asp Ala Glu Leu Ala Lys Asn Ala Gly Leu Ser Leu Glu Glu Met Lys
        115                 120                 125

Ile Arg Ala Gln Ala Leu Leu His Arg Glu Ile Asn Asn Trp Phe Gln
    130                 135                 140

Val Tyr Gln Gln Cys Gln Asp Val Val Glu Arg Ser Ile Phe Ala Tyr
145                 150                 155                 160

Leu Leu Lys His Arg Leu Thr Val Pro Ala Glu Pro Glu Asn Thr Asp
                165                 170                 175

Lys Leu Arg Gln Lys Arg Arg Gln Val Glu Ile Lys Ile Glu Arg Leu
            180                 185                 190

Glu Thr Gln Leu Ala Gly Arg Ser Pro Gln Gly Arg Asp Leu Thr Gly
        195                 200                 205

Ser Arg Tyr Ala Thr Ala Leu Asn Asp Gly Glu Gln Cys Asp Trp Glu
    210                 215                 220

Asp Asp Ala Asp Phe Leu Ala Trp Gln Ala Glu Ile Leu Ser Arg Pro
225                 230                 235                 240

Asp Ser Leu Pro Pro Pro Val Glu Tyr Ala Thr Asn Thr Asp Met Thr
                245                 250                 255

Trp His Lys Asp Glu Gln Gly Arg Leu Ala Val Thr Phe Asn Gly Leu
            260                 265                 270

Gly Lys Leu Lys Phe Lys Ile Ala Cys Asp Gln Arg Gln Leu His Trp
        275                 280                 285

Phe Gln Arg Phe Tyr Gln Asp Gln Glu Gln Phe Lys Ala Gln Lys Gly
    290                 295                 300

Gln Arg Ser Gln Ala Leu Phe Thr Leu Arg Ser Ala Glu Leu Leu Trp
305                 310                 315                 320
```

```
Lys Pro Gly Asn Arg Ser Gly Asp Pro Trp Gln Ala Asn Phe Leu Tyr
                325                 330                 335

Leu His Cys Thr Val Asp Ser Arg Leu Trp Thr Gln Glu Gly Thr Ala
            340                 345                 350

Met Val Gln Gln Glu Lys Ala Lys Lys Ser Gln Ala Ile Val Lys Lys
        355                 360                 365

Leu Ser Glu Arg Ser Asp Leu Thr Ala Gln Gln Lys Asp Cys Leu Gln
    370                 375                 380

Arg His Gln Ser Thr Leu Ala Arg Leu His Met Gly Tyr Asp Arg Pro
385                 390                 395                 400

Gln Arg Arg Met Tyr Gln Gly Lys Ser His Leu Val Val Gly Ile Ser
                405                 410                 415

Leu Asp Met Glu Asn Leu Val Thr Val Ala Leu Val Asp Val Val Lys
            420                 425                 430

Gln Lys Val Ile Thr Gly Cys Thr Met Lys Ser Leu Leu Gly Gln Asp
        435                 440                 445

Tyr Ala Leu Val Gln Arg Leu Arg Tyr Glu Lys Arg Gln Asn Ser His
    450                 455                 460

Leu Arg Lys Val Ala Gln Glu Arg Gly Ser Lys Ile Val Asn Tyr Glu
465                 470                 475                 480

Ala Asn Leu Ala Ile His Val Glu Arg Leu Leu Val Lys Ala Ile Ile
                485                 490                 495

His Phe Ala Gln Gln His Leu Ala Gly Ser Leu Cys Val Pro Thr Leu
            500                 505                 510

Lys Asp Ile Arg Glu Thr Ile Gln Ala His Leu Gln Cys Arg Ala Glu
        515                 520                 525
```

Glu Arg Phe Pro Asp Ser Lys Glu Leu Gln Arg Arg Tyr Ala Lys Glu
530 535 540

Tyr Arg Ile Asn Ala His Arg Trp Ser Tyr Asn Arg Leu Leu Lys Leu
545 550 555 560

Leu Asn Gln Gln Ala Lys Phe Ala Gly Leu Val Val Glu Gln Gly Val
565 570 575

Gln Ser Ala Gly Glu Thr Ala Leu Glu Arg Ala Leu Gly Val Ala Leu
580 585 590

Ser Ala Tyr Tyr Gln Arg Ser Ala Ala
595 600

<210> 64
<211> 605
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-17 effector

<400> 64
Met Thr Lys Leu Thr Ile Gln Cys Arg Leu Val Ala Glu Glu Glu Thr
1 5 10 15

Arg Arg Gln Val Trp Glu Met Met Ala Gly Arg Tyr Ala Pro Leu Ile
20 25 30

Ala Thr Thr Leu Glu Gln Val Ser Arg His Lys Asp Phe Pro Gln Trp
35 40 45

Val Ser Ala Gly Glu Ile Pro Ala Gln Val Val Lys Ser Leu Leu His
50 55 60

Ser Ala Lys Thr Ser Tyr Gly Gly Leu Pro Ala Arg Trp Tyr Ala Ser
65 70 75 80

Ala Gln Arg Gln Val Gln Glu Ile Tyr Lys Gly Trp Leu Lys Arg Arg

```
                85                  90                  95

Arg Gln Leu Gln His Lys Leu Gln Gly Gln Gln Thr Trp Leu Ser Val
            100                 105                 110

Leu Gln Pro Asp Ala Glu Leu Ala Gln Ser Ala Gly Leu Ser Leu Pro
        115                 120                 125

Glu Leu Gln Gln Arg Ala Gln Ala Leu Leu Ala Gly Asp Thr Val Thr
    130                 135                 140

Asn Trp Tyr His Ala Tyr Gln Gln Cys Gln Asp Ala Val Glu Arg Ser
145                 150                 155                 160

Ile Phe Ala Tyr Leu Leu Lys His Arg Leu Thr Val Pro Thr Glu Pro
                165                 170                 175

Glu Asp Thr Asp Lys Leu Lys Arg Lys Arg Arg Arg Val Glu Ile Arg
            180                 185                 190

Ile Glu Arg Leu Glu Thr Gln Leu Ala Gly Arg Ser Pro Gln Gly Arg
        195                 200                 205

Asp Leu Thr Gly Ile Arg Tyr Ala Thr Ala Leu His Glu Gly Glu Gln
    210                 215                 220

Cys Tyr Trp Glu Asp Asp Ala Asp Phe Leu Ala Trp Gln Ala Glu Ile
225                 230                 235                 240

Leu Ser Cys Pro Asp Ser Val Pro Pro Pro Val Glu Tyr Ala Thr Asn
                245                 250                 255

Thr Asp Met Thr Trp His Lys Asp Glu Gln Ser Arg Leu Trp Val Thr
            260                 265                 270

Phe Asn Gly Leu Gly Lys His Lys Phe Lys Ile Ala Cys Asp Gln Arg
        275                 280                 285

Gln Leu His Trp Phe Glu Arg Phe Tyr Lys Asp Gln Glu Leu Phe Lys
    290                 295                 300
```

```
Ala Gln Lys Gly Gln Arg Ser Gln Ala Leu Phe Thr Leu Arg Ser Cys
305                 310                 315                 320

Glu Leu Leu Trp Lys Pro Gly Asn Arg Ser Gly Glu Leu Trp Gln Val
                325                 330                 335

Asn Cys Leu Tyr Leu His Cys Thr Val Asp Pro Arg Leu Trp Thr Gln
            340                 345                 350

Glu Gly Thr Ala Ile Val Gln Gln Glu Lys Ala Gln Lys Ser Gln Ala
        355                 360                 365

Ile Val Lys Glu Leu Ser Glu Arg Ser Asp Leu Thr Pro Gly Gln Glu
    370                 375                 380

Asp Tyr Leu Arg Arg His Gln Ser Thr Leu Ala Arg Leu His Thr Gly
385                 390                 395                 400

Tyr Ser Arg Pro Gln Arg Arg Met Tyr Gln Gly Lys Ser His Leu Ser
                405                 410                 415

Val Gly Ile Ser Leu Asp Met Asp Asn Leu Val Thr Val Ala Leu Val
            420                 425                 430

Asp Val Val Glu Gln Arg Ile Ile Thr Gly Thr Thr Met Lys Asn Leu
        435                 440                 445

Leu Gly Ser Gly Tyr Ala Leu Val Gln Arg Leu Arg Tyr Glu Lys Arg
    450                 455                 460

Arg Asn Ala His Leu Arg Lys Val Ala Gln Gln Arg Gly Ser Lys Ile
465                 470                 475                 480

Val Asn His Glu Ala Asn Leu Ala Thr His Ile Glu Arg Leu Leu Ala
                485                 490                 495

Lys Ala Ile Ile Asn Phe Ala Gln Gln His Leu Ala Gly Ser Leu Cys
            500                 505                 510

Val Pro Thr Leu Lys Asp Ile Arg Glu Thr Ile Gln Ala His Ile Gln
```

515 520 525

Ser Arg Ala Glu Glu Arg His Pro Asn Ser Lys Glu Leu Gln Arg Arg
530 535 540

Tyr Ala Lys Glu Tyr Arg Ile Asn Ala His Arg Trp Ser Tyr Asn Arg
545 550 555 560

Leu Leu Glu Leu Ile Ser Gln Gln Ala Ile Val Ala Gly Leu Leu Val
565 570 575

Glu Gln Gly Leu Gln Gly Ser Gly Gly Thr Ala Ile Lys Arg Ala Leu
580 585 590

Gly Val Ala Leu Ser Ala Tyr Tyr Gln Arg Ser Ala Ala
595 600 605

<210> 65
<211> 613
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-18 effector

<400> 65
Met Ser Ile Lys Thr Ile Arg Cys Arg Leu Val Ala Ser Glu Ser Thr
1 5 10 15

Arg Lys Ala Leu Trp Glu Leu Met Ala Gly Lys Asn Thr Pro Leu Met
20 25 30

Asn His Val Leu Ala Leu Val Ser Ser Asp Thr Asn Phe Asp Ser Trp
35 40 45

Arg Gln Lys Gly Glu Ile Pro Ala Ser Ile Ile Lys Asn Tyr Ile Asn
50 55 60

Thr Leu Lys His Glu Pro Gln Phe Ala Gly Gln Pro Trp Trp Cys Tyr
65 70 75 80

```
Ala Ser Ser Glu Lys Gln Ala Thr Arg Ile Phe Asp Asn Trp Leu Ala
                85                  90                  95

Thr Gln Lys Gln Leu Gln Ser Lys Leu Asn Gly Lys Glu Gln Trp Leu
            100                 105                 110

Ser Val Leu Gln Pro Asp Ser Glu Leu Ala Ser Gln Ala Ser Cys Thr
        115                 120                 125

Leu Ala Ala Leu Arg Arg Leu Ala Ser Gln Ile Asn Lys Gln Gly Leu
    130                 135                 140

Thr Phe Asn Gln Leu Phe Asn Leu Tyr Thr Thr Val Gln Val Leu Lys
145                 150                 155                 160

Gln Arg Ala Ala Ile Ala Tyr Leu Leu Lys Arg Gln Gly Lys Leu Glu
                165                 170                 175

Pro Asn Glu Glu Asp Pro Glu Gln Ile Ala Arg Lys Arg Arg Lys Thr
            180                 185                 190

Glu Ile Ser Ile Gln Arg Leu Lys Lys Gln Ile Gln Ala Arg Val Pro
        195                 200                 205

Lys Gly Arg Asp Leu Thr Gly Ser Glu Tyr Thr Ala Ala Leu Asp Glu
    210                 215                 220

Cys Ile Gln Thr Ala Ile Thr Thr Asp Glu Glu Tyr Glu Ala Trp Gln
225                 230                 235                 240

Gly Lys Leu Thr Ala Lys Thr Val Ser Val Pro Phe Pro Leu Ile Cys
                245                 250                 255

Glu Ser Ser Glu Val Leu Lys Trp Ser Gln Thr Ser Ser Gly Arg Ile
            260                 265                 270

Val Val Lys Phe Ser Gly Leu His Gly Leu Ser Phe Asp Ile Tyr Cys
        275                 280                 285
```

```
Asp Lys Thr His Leu Arg Trp Phe Ser Arg Phe Leu Ala Asp Gln Glu
    290                 295                 300

Val Lys Lys Ala Ser Gly Gly Lys His Ser Ala Ala Leu Phe Thr Leu
305                 310                 315                 320

Arg Ser Ala Thr Leu Leu Trp Arg Pro Asn Lys His Pro Ser Lys Gly
                325                 330                 335

Asp Pro Trp Ser Thr Asn Tyr Leu Glu Leu His Cys Thr Val Asp Thr
            340                 345                 350

Arg Leu Trp Thr Ala Glu Gly Thr Glu Gln Val Arg Arg Glu Lys Ser
        355                 360                 365

Thr Glu Val Glu Lys Ile Leu Thr Glu Ile Gly Ser Lys Asp Ser Leu
    370                 375                 380

Ser Asn Asn Gln Leu Val Tyr Ile Lys Arg Lys Gln Ala Thr Leu Thr
385                 390                 395                 400

Arg Leu Gln Gly Asn Phe Asp Arg Pro His Gln Ser Pro Tyr Cys Gly
                405                 410                 415

Asp Pro Asp Ile Ser Val Gly Ile Ser Met Gly Leu Asp Lys Pro Val
            420                 425                 430

Thr Leu Ala Ile Val Asn Leu Lys Thr Gly Gln Val Ser Ala Tyr Arg
        435                 440                 445

Thr Thr Arg Gln Leu Leu Gly Ser Asn Tyr Cys Leu Leu Gly Arg Arg
    450                 455                 460

Arg Lys Glu Gln Glu Arg Ala Ser His Arg Ala His Ile Asn Arg Ala
465                 470                 475                 480

Arg His Arg Ser Asp Tyr His Thr Tyr Lys Glu Ser Gln Val Gly Thr
                485                 490                 495

Gln Ile Asp His Leu Leu Ala Cys Ser Ile Ile Glu Val Ala Met Gln
            500                 505                 510
```

Tyr Lys Ala Ser Ser Ile Val Leu Pro Asp Leu Glu Tyr Ile Arg Glu
515 520 525

Val Val Glu Ala Glu Ile Lys Glu Arg Ala Glu Arg Arg Ile Pro Asp
530 535 540

Tyr Ser Glu Gly Gln Lys Arg Tyr Ala Lys Glu Tyr Arg Lys Lys Val
545 550 555 560

His Ser Trp Ser Tyr Arg Arg Leu His Asn Phe Ile Lys Asn Lys Cys
565 570 575

Glu Ala Val Gly Ile Ala Ile Glu Val Gln Arg Gln Glu Thr Ile Gly
580 585 590

Thr Pro Gln Gln Lys Ala Arg Lys Leu Val Glu Lys Ala Tyr Ala Asn
595 600 605

Arg Cys Phe Val Ala
610

<210> 66
<211> 543
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-19 effector

<400> 66
Lys Asn Leu Val Asn Gln Ala Gln Ser Gly Leu Pro Ala Arg Trp Cys
1 5 10 15

Ala Ser Ala Gln Arg Gln Val Gln Glu Thr Tyr Lys Ala Trp Leu Thr
20 25 30

Lys Arg Arg Lys Leu Gln Gln Lys Leu Gln Ala Gln Gln Thr Trp Leu
35 40 45

```
Ser Val Leu Arg Pro Asp Ala Glu Leu Ala Gln Glu Ala Gly Leu Ser
    50                  55                  60

Leu Pro Glu Leu Gln Gln Arg Ala Gln Ala Leu Leu Ala Gly Asp Thr
65                  70                  75                  80

Ile Thr Asn Trp Tyr Gln Ala Tyr Gln Glu Cys Gln Asp Val Val Gln
                85                  90                  95

Arg Ser Val Phe Ala Tyr Leu Leu Lys His Arg Leu Thr Val Pro Asn
            100                 105                 110

Glu Pro Glu Asp Thr Asp Lys Leu Lys Arg Lys Arg Arg Gln Val Glu
        115                 120                 125

Ile Lys Ile Glu Arg Leu Glu Thr Gln Leu Ala Gly Arg Ser Pro Gln
    130                 135                 140

Gly Arg Asp Leu Thr Gly Ser Arg Tyr Ala Thr Ala Leu Asn Asp Gly
145                 150                 155                 160

Glu Gln Cys Asp Trp Glu Asp Asp Ala Asp Phe Leu Ala Trp Gln Ala
                165                 170                 175

Glu Ile Leu Ser Arg Pro Asp Ser Leu Pro Pro Pro Val Glu Tyr Ala
            180                 185                 190

Thr Asn Thr Asp Met Thr Trp His Lys Asp Glu Gln Gly Arg Leu Ala
        195                 200                 205

Val Thr Phe Asn Gly Leu Gly Lys Leu Lys Phe Lys Ile Ala Cys Asp
    210                 215                 220

Gln Arg Gln Leu His Trp Phe Gln Arg Phe Tyr Gln Asp Gln Glu Gln
225                 230                 235                 240

Phe Lys Ser Gln Lys Gly Gln Arg Ser Gln Ala Leu Phe Thr Leu Arg
                245                 250                 255

Ser Ala Glu Leu Leu Trp Lys Pro Gly Asn Arg Ser Gly Asp Pro Trp
```

```
            260                 265                 270

Gln Ala Asn Phe Leu Tyr Leu His Cys Thr Val Asp Ser Arg Leu Trp
        275                 280                 285

Thr Gln Glu Gly Thr Ala Met Val Gln Gln Glu Lys Ala Lys Lys Ser
    290                 295                 300

Gln Ala Ile Val Lys Lys Leu Ser Glu Arg Ser Asp Leu Thr Ala Gln
305                 310                 315                 320

Gln Lys Asp Cys Leu Gln Arg His Gln Ser Thr Leu Ala Arg Leu His
                325                 330                 335

Met Gly Tyr Asp Arg Pro Gln Arg Arg Met Tyr Gln Gly Lys Ser His
            340                 345                 350

Leu Val Val Gly Ile Ser Leu Asp Met Glu Asn Leu Val Thr Val Ala
        355                 360                 365

Leu Val Asp Val Val Lys Gln Lys Val Ile Thr Gly Cys Thr Met Lys
    370                 375                 380

Ser Leu Leu Gly Gln Asp Tyr Ala Leu Val Gln Arg Leu Arg Tyr Glu
385                 390                 395                 400

Lys Arg Gln Asn Ser His Leu Arg Lys Val Ala Gln Glu Arg Gly Ser
                405                 410                 415

Lys Ile Val Asn Tyr Glu Ala Asn Leu Ala Ile His Val Glu Arg Leu
            420                 425                 430

Leu Val Lys Ala Ile Ile His Phe Ala Gln Gln His Leu Ala Gly Ser
        435                 440                 445

Leu Cys Val Pro Thr Leu Lys Asp Ile Arg Glu Thr Ile Gln Ala His
    450                 455                 460

Leu Gln Cys Arg Ala Glu Glu Arg Phe Pro Asp Ser Lys Glu Leu Gln
465                 470                 475                 480
```

Arg Arg Tyr Ala Lys Glu Tyr Arg Ile Asn Ala His Arg Trp Ser Tyr
485 490 495

Asn Arg Leu Leu Lys Leu Leu Asn Gln Gln Ala Lys Phe Ala Gly Leu
500 505 510

Val Val Glu Gln Gly Val Gln Ser Ala Gly Glu Thr Ala Leu Glu Arg
515 520 525

Ala Leu Gly Val Ala Leu Ser Ala Tyr Tyr Gln Arg Ser Ala Ala
530 535 540

<210> 67
<211> 605
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-20 effector

<400> 67
Met Ser Lys Leu Thr Ile Gln Cys Arg Leu Val Ala Cys Glu Asp Thr
1 5 10 15

Arg Arg Gln Val Trp Glu Met Met Ala Gly Arg Tyr Ala Pro Leu Ile
20 25 30

Ala Thr Thr Leu Glu Gln Val Ser Gln His Lys Asp Phe Pro Gln Trp
35 40 45

Val Ser Ala Gly Glu Ile Pro Ala Gln Val Val Lys Ser Leu Leu His
50 55 60

Ser Ala Lys Thr Ser Tyr Gly Gly Leu Pro Ala Arg Trp Tyr Ala Ser
65 70 75 80

Ala Gln Arg Gln Val Gln Glu Ile Tyr Lys Gly Trp Leu Lys Arg Arg
85 90 95

```
Arg Gln Leu Gln His Lys Leu Gln Gly Gln Gln Thr Trp Leu Ser Val
            100                 105                 110

Leu Gln Pro Asp Ala Glu Leu Ala Gln Ser Ala Gly Leu Ser Leu Pro
        115                 120                 125

Glu Leu Gln Gln Arg Ala Gln Ala Leu Leu Ala Gly Asp Thr Val Thr
    130                 135                 140

Asn Trp Tyr His Ala Tyr Gln Gln Cys Gln Asp Ala Val Glu Arg Ser
145                 150                 155                 160

Ile Phe Ala Tyr Leu Leu Lys His Arg Leu Thr Val Pro Thr Glu Pro
                165                 170                 175

Glu Asp Thr Asp Lys Leu Lys Arg Lys Arg Arg Arg Val Glu Ile Arg
            180                 185                 190

Ile Glu Arg Leu Glu Thr Gln Leu Ala Gly Arg Ser Pro Gln Gly Arg
        195                 200                 205

Asp Leu Thr Gly Ile Arg Tyr Ala Thr Ala Leu His Glu Gly Glu Gln
    210                 215                 220

Cys Tyr Trp Glu Asp Asp Ala Asp Phe Leu Ala Trp Gln Ala Glu Ile
225                 230                 235                 240

Leu Ser Cys Pro Asp Ser Val Pro Pro Pro Val Glu Tyr Ala Thr Asn
                245                 250                 255

Thr Asp Met Thr Trp His Lys Asp Glu Gln Ser Arg Leu Trp Val Thr
            260                 265                 270

Phe Asn Gly Leu Gly Lys His Lys Phe Lys Ile Ala Cys Asp Gln Arg
        275                 280                 285

Gln Leu His Trp Phe Glu Arg Phe Tyr Lys Asp Gln Glu Leu Phe Lys
    290                 295                 300

Ala Gln Lys Gly Gln Arg Ser Gln Ala Leu Phe Thr Leu Arg Ser Gly
305                 310                 315                 320
```

```
Glu Leu Leu Trp Lys Pro Gly Asn Arg Ser Gly Glu Leu Trp Gln Val
               325                 330                 335

Asn Cys Leu Tyr Leu His Cys Thr Val Asp Pro Arg Leu Trp Thr Gln
            340                 345                 350

Glu Gly Thr Ala Ile Val Gln Gln Glu Lys Ala Gln Lys Ser Gln Ala
        355                 360                 365

Ile Val Lys Glu Leu Ser Glu Arg Ser Asp Leu Thr Pro Gly Gln Glu
    370                 375                 380

Asp Tyr Leu Arg Arg His Gln Ser Thr Leu Ala Arg Leu His Thr Gly
385                 390                 395                 400

Tyr Ser Arg Pro Gln Arg Arg Met Tyr Gln Gly Lys Ser His Leu Ser
               405                 410                 415

Val Gly Ile Ser Leu Asp Met Asp Asn Leu Val Thr Val Ala Leu Val
            420                 425                 430

Asp Val Val Glu Gln Arg Ile Ile Thr Gly Thr Thr Met Lys Asn Leu
        435                 440                 445

Leu Gly Ser Gly Tyr Ala Leu Val Gln Arg Leu Arg Tyr Glu Lys Arg
    450                 455                 460

Arg Asn Ala His Leu Arg Lys Val Ala Gln Gln Arg Gly Ser Lys Ile
465                 470                 475                 480

Val Asn His Glu Ala Asn Leu Ala Thr His Ile Glu Arg Leu Leu Ala
               485                 490                 495

Lys Ala Ile Ile Asn Phe Ala Gln Gln His Leu Ala Gly Ser Leu Cys
            500                 505                 510

Val Pro Thr Leu Lys Asp Ile Arg Glu Thr Ile Gln Ala His Ile Gln
        515                 520                 525
```

Ser Arg Ala Glu Glu Arg His Pro Asn Ser Lys Glu Leu Gln Arg Arg
530 535 540

Tyr Ala Lys Glu Tyr Arg Ile Asn Ala His Arg Trp Ser Tyr Asn Arg
545 550 555 560

Leu Leu Glu Leu Ile Ser Gln Gln Ala Ile Val Ala Gly Leu Leu Val
565 570 575

Glu Gln Gly Leu Gln Gly Ser Gly Gly Thr Ala Ile Lys Arg Ala Leu
580 585 590

Gly Val Ala Leu Ser Ala Tyr Tyr Gln Arg Ser Ala Ala
595 600 605

<210> 68
<211> 638
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-21 effector

<400> 68
Met Gly Gln Ile Thr Ile Gln Ser Arg Leu Val Ala Ser Glu Ala Thr
1 5 10 15

Arg Gln His Leu Trp Met Leu Met Ala Asp Arg Asn Thr Pro Leu Ile
20 25 30

Asn Gln Leu Leu Glu Gln Ile Asn Gln His Pro Asp Phe Leu Thr Trp
35 40 45

Arg Gln Lys Gly Lys Ile Pro Val Gly Thr Val Lys Gln Leu Cys Gln
50 55 60

Pro Leu Arg Ser Asp Pro Arg Phe Ser Gly Gln Pro Gly Arg Phe Tyr
65 70 75 80

Thr Ser Ala Ile Ala Leu Val Asp Tyr Ile Tyr Lys Ser Trp Leu Thr

```
                85                  90                  95

Leu Gln Gln Arg Leu Gln His Lys Leu Asn Gly Gln Gln Arg Trp Leu
            100                 105                 110

Glu Met Leu Lys Ser Asp Glu Glu Leu Ala Glu Ile Ser Gly Cys Ser
        115                 120                 125

Leu Asp Thr Leu Arg Gln Arg Ala Ala Asn Leu Leu Gln Ala Asp Asn
    130                 135                 140

Pro Ile His Pro Glu Pro Ser Thr Leu Pro Lys Ser Gly Lys Arg Arg
145                 150                 155                 160

Lys Pro Lys Gln Thr Glu Thr Ser Ala Leu Ser Pro Asn Ile Ser Gln
                165                 170                 175

Ala Leu Phe Phe Ala Tyr Arg Glu Thr Glu Asp Val Lys Ile Arg Ala
            180                 185                 190

Ala Ile Ala Tyr Leu Leu Lys Asn Gly Cys Thr Leu Pro Glu Lys Pro
        195                 200                 205

Glu Asp Pro Gln Lys Phe Ala Lys Arg Arg Arg Lys Val Glu Ile Gln
    210                 215                 220

Ile Gln Arg Leu Thr Glu Gln Leu Glu Gly Arg Val Pro Lys Gly Arg
225                 230                 235                 240

Asp Leu Thr Asn Thr Gln Trp Leu Ala Ala Leu Glu Thr Ala Thr Thr
                245                 250                 255

Thr Ala Pro Gln Ser Glu Ala Glu Ala Lys Ser Trp Gln Asp Ser Leu
            260                 265                 270

Leu Arg Glu Ser Ser Asn Leu Pro Phe Pro Val Ala Tyr Glu Thr Asn
        275                 280                 285

Thr Asp Leu Thr Trp Phe Arg Asn Gln Gln Gly Arg Leu Cys Val Arg
    290                 295                 300
```

```
Phe Asn Gly Leu Gly Asp His Ile Phe Gln Ile Tyr Cys Asp Arg Arg
305                 310                 315                 320

Gln Leu His Trp Phe Glu Arg Phe Leu Asp Asp Gln Gln Val Gln Lys
                325                 330                 335

Asp Ser Lys Asp Gln His Ser Ser Ala Leu Phe Thr Leu Arg Ser Val
            340                 345                 350

Arg Ile Ser Trp Gln Glu Gly Lys Gly Lys Gly Thr Pro Trp Gln Ile
        355                 360                 365

His Arg Leu Ala Leu Gln Cys Ser Leu Asp Thr Arg Leu Trp Thr Gln
    370                 375                 380

Glu Gly Thr Glu Gln Val Arg Asn Glu Lys Ala Ala Asp Ile Ala Lys
385                 390                 395                 400

Ile Leu Thr Lys Met Glu Asp Lys Gly Asn Leu Asn Asp Lys Gln Glu
                405                 410                 415

Ala Phe Ile Lys Arg Lys Gln Ser Thr Leu Asp Arg Ile Asn His Thr
            420                 425                 430

Phe Pro Arg Pro Ser Lys Pro Leu Tyr Gln Gly Gln Ser Gln Ile Ile
        435                 440                 445

Val Gly Val Ser Ile Gly Leu Glu Lys Leu Ala Thr Ala Ala Val Val
    450                 455                 460

Asp Ala Ser Thr Gly Lys Ile Leu Thr Tyr Arg Ser Ile Arg Gln Leu
465                 470                 475                 480

Leu Gly Asn Asn Tyr Arg Leu Leu Asn Arg Gln Arg Gln Gln Gln His
                485                 490                 495

His Asn Thr His Gln Arg Gln Ile Ala Gln Arg Gln Gly Lys Arg Gly
            500                 505                 510

Leu Leu Ser Glu Ser Glu Leu Gly Gln Tyr Val Asp Arg Leu Leu Ala
```

515 520 525

Asp Ala Ile Val Asp Leu Ala Lys Thr Tyr Tyr Ala Gly Ser Ile Val
530 535 540

Val Pro Lys Leu Gly Asp Val Arg Glu Arg Val Gln Ser Glu Ile Gln
545 550 555 560

Ala Arg Ala Glu Gln Val Cys Pro Asp Leu Leu Ser Gly Gln Gln Asn
565 570 575

Tyr Ala Lys Gln His Arg Ser Ala Ile His Arg Trp Ser Tyr Gly Arg
580 585 590

Leu Ile Asp Gln Ile Gln Ser Asn Ala Lys Arg Thr Gly Ile Val Val
595 600 605

Glu Asp Ala Gln Gln Pro Ile Gln Gly Thr Pro Gln Glu Ile Ala Arg
610 615 620

Ala Leu Ala Ile Thr Ala Tyr Gln Asp Arg Leu Lys Met Ile
625 630 635

<210> 69
<211> 502
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-22 effector

<400> 69
Ala Asn Leu Leu Gln Ala Asp Asn Pro Ile His Pro Glu Pro Ser Thr
1 5 10 15

Leu Pro Lys Ser Ser Lys Arg Arg Lys Pro Lys Gln Thr Glu Thr Ser
20 25 30

Val Leu Ser Pro Asn Ile Ser Gln Ser Leu Phe Val Ala Tyr Arg Glu
35 40 45

```
Thr Glu Asp Val Lys Thr Arg Ala Ala Ile Ala Tyr Leu Leu Lys Asn
    50                  55                  60

Gly Cys Thr Leu Pro Glu Lys Pro Glu Asp Pro Gln Lys Phe Ala Lys
65                  70                  75                  80

Arg Arg Arg Lys Val Glu Ile Gln Ile Gln Arg Leu Thr Glu Gln Leu
                85                  90                  95

Glu Gly Arg Val Pro Lys Gly Arg Asp Leu Thr Asn Thr Gln Trp Leu
            100                 105                 110

Ala Ala Leu Glu Thr Ala Thr Thr Thr Ala Pro Gln Ser Glu Ala Glu
        115                 120                 125

Ala Lys Ser Trp Gln Asp Ser Leu Leu Arg Glu Ser Ser Asn Leu Pro
    130                 135                 140

Phe Pro Val Ala Tyr Glu Thr Asn Thr Asp Leu Thr Trp Phe Arg Asn
145                 150                 155                 160

Gln Gln Gly Arg Leu Cys Val Arg Phe Asn Gly Leu Gly Asp His Ile
                165                 170                 175

Phe Gln Ile Tyr Cys Asp Arg Arg Gln Leu His Trp Phe Glu Arg Phe
            180                 185                 190

Leu Asp Asp Gln Gln Val Gln Lys Asp Ser Lys Asp Gln His Ser Ser
        195                 200                 205

Ala Leu Phe Thr Leu Arg Ser Val Arg Ile Ser Trp Gln Glu Gly Lys
    210                 215                 220

Gly Lys Gly Thr Pro Trp Gln Ile His Arg Leu Ala Leu Gln Cys Ser
225                 230                 235                 240

Leu Asp Thr Arg Leu Trp Thr Gln Glu Gly Thr Glu Gln Val Arg Asn
                245                 250                 255
```

```
Glu Lys Ala Ala Asp Ile Ala Lys Ile Leu Thr Lys Met Glu Asp Lys
            260                 265                 270

Gly Asn Leu Asn Asp Lys Gln Glu Ala Phe Ile Lys Arg Lys Gln Ser
        275                 280                 285

Thr Leu Asp Arg Ile Asn His Thr Phe Pro Arg Pro Ser Lys Pro Leu
    290                 295                 300

Tyr Gln Gly Gln Ser Gln Ile Ile Val Gly Val Ser Ile Gly Leu Glu
305                 310                 315                 320

Lys Leu Ala Thr Ala Ala Val Val Asp Ala Ser Thr Gly Lys Ile Leu
                325                 330                 335

Thr Tyr Arg Ser Ile Arg Gln Leu Leu Gly Asn Asn Tyr Arg Leu Leu
            340                 345                 350

Asn Arg Gln Arg Gln Gln Gln His His Asn Thr His Gln Arg Gln Ile
        355                 360                 365

Ala Gln Arg Gln Gly Lys Arg Gly Leu Leu Ser Glu Ser Glu Leu Gly
    370                 375                 380

Gln Tyr Val Asp Arg Leu Leu Ala Asp Ala Ile Val Asp Leu Ala Lys
385                 390                 395                 400

Thr Tyr Tyr Ala Gly Ser Ile Val Val Pro Lys Leu Gly Asp Val Arg
                405                 410                 415

Glu Arg Val Gln Ser Glu Ile Gln Ala Arg Ala Glu Gln Val Cys Pro
            420                 425                 430

Asp Leu Leu Ser Gly Gln Gln Asn Tyr Ala Lys Gln His Arg Ser Ala
        435                 440                 445

Ile His Arg Trp Ser Tyr Gly Arg Leu Ile Asp Gln Thr Gln Ser Asn
    450                 455                 460

Ala Lys Arg Thr Gly Ile Val Val Glu Asp Ala Gln Gln Pro Ile Gln
465                 470                 475                 480
```

Gly Thr Pro Gln Glu Ile Ala Arg Ala Leu Ala Ile Thr Ala Tyr Gln
485 490 495

Asp Arg Leu Lys Met Ile
500

<210> 70
<211> 541
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-23 effector

<400> 70
Ala Asp Ser Leu Arg Ser Asp Pro Arg Phe Lys Ser His Pro Gln Arg
1 5 10 15

Leu Val Ala Ser Ala Gly Arg Val Val Asn Ser Met Phe Glu Ser Trp
20 25 30

Leu Ala Ile Gln Lys Gly Asn Phe Ala Arg Leu Lys Gly Lys Gln Arg
35 40 45

Trp Leu Ala Met Leu Lys Ser Asp Glu Glu Leu Val Glu Tyr Ser Asp
50 55 60

Cys Asp Leu Tyr Gln Ile Arg Asn His Ala Ala Gln Ile Leu Glu Ser
65 70 75 80

Val Ala Ala Ser Val Pro Ser Gln Thr Ser Asn Pro Ala Asn Lys Lys
85 90 95

Ser Lys Gln Gln Lys Ser Asp Arg Lys Lys Ser Lys Gln Lys Asp Asn
100 105 110

Asn Ile Pro Lys Ala Leu His Gly Leu Tyr Asp Ser Thr Asp Asn Val
115 120 125

```
Leu Thr Gln Ala Ala Ile Ile Tyr Leu Leu Lys Asn Gly Ser Lys Val
    130                 135                 140

Ser Asp Cys Glu Glu Asp Pro Glu Lys Phe Thr Gln Lys Tyr Arg Lys
145                 150                 155                 160

Thr Glu Leu Gln Val Glu Arg Leu Gln Lys Gln Ile Asp Ala Asn Met
                165                 170                 175

Pro Lys Gly Arg Asp Leu Thr Gly Gln Val Trp Glu Glu Thr Leu Glu
            180                 185                 190

Ile Ala Thr Leu Thr Ala Pro Arg Asn Gly Asp Glu Ala Lys Glu Trp
        195                 200                 205

Gln Asn Thr Leu Leu Arg Lys Pro Lys Asn Thr Pro Tyr Pro Ile Phe
    210                 215                 220

Tyr Arg Asn Lys Glu Asp Leu Arg Trp Tyr Glu Ile Glu Ile Asn Gly
225                 230                 235                 240

Lys Ala Arg Arg Arg Gly Lys Asn His Asn Ser Pro Lys Lys Gln Leu
                245                 250                 255

Cys Val Thr Leu Ser Gly Phe Gly Gly His Ala Leu Lys Val Ser Cys
            260                 265                 270

Asp Arg Arg Gln Leu His Trp Phe Lys Arg Phe Leu Glu Asp Gln Gln
        275                 280                 285

Thr Lys Lys Asn Ser Asp Gly Gln Cys Ser Ser Ser Leu Phe Thr Leu
    290                 295                 300

Arg Ser Ala Gln Leu Leu Trp His Glu Gln Lys Gly Asn Gly Asn Ser
305                 310                 315                 320

Trp Asp Val His His Leu His Leu His Cys Thr Val Glu Thr Pro Arg
                325                 330                 335

Trp Thr Ala Glu Gly Thr Glu Lys Ile Arg Gln Ala Glu Arg Gln Glu
```

```
            340                 345                 350

Val Ala Lys Arg Ile Glu Ser Leu Asn Gln Lys Glu Ala Leu Ser Pro
        355                 360                 365

Ser Gln Gln Lys Tyr Leu Gln Arg Leu Asp Ser Gln Leu Lys Gly Leu
    370                 375                 380

Asp Tyr Pro Phe Asp Arg Pro Ser Arg Pro Leu Tyr Thr Pro Gln Pro
385                 390                 395                 400

Asp Ile Ile Ala Ala Val Ser Met Gly Leu Gln Tyr Pro Ala Thr Leu
                405                 410                 415

Ile Ile Met Asn Ser Ser Thr Gln Glu Val Leu Ala Thr Arg Thr Ile
            420                 425                 430

Arg Glu Leu Leu Gly Ser Asn Tyr Gly Leu Leu Leu Arg Arg Arg Arg
        435                 440                 445

Gln Gln Gln Lys Asn Ala His Gln Arg His Lys Ala Gln Lys Arg Ser
    450                 455                 460

Ala Ser Asp Gln Phe Gly Glu Ser Asn Leu Gly Glu Tyr Leu Asp Arg
465                 470                 475                 480

Leu Ile Ala Lys Glu Ile Val Ala Ile Ala Lys Val His Lys Ala Gln
                485                 490                 495

Ser Val Val Ile Pro Lys Leu Ala Arg Ile Arg Glu Ile Leu Asn Cys
            500                 505                 510

Glu Val Met Ser Arg Ala Glu Gln Arg Cys Pro Gly Ser Val Glu Ile
        515                 520                 525

Gln Lys Arg Tyr Ala Lys Glu Tyr Arg Thr Lys Ile His
    530                 535                 540

<210> 71
<211> 614
<212> PRT
```

<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-24 effector

<400> 71
Met Ser Gln Lys Thr Val Arg Cys Arg Leu Ile Ser Thr Glu Glu Asn
1 5 10 15

Arg Lys Ala Leu Trp Leu Leu Met Ala Glu Arg Asn Thr Pro Leu Ile
20 25 30

Asn Glu Ala Leu Arg Gln Leu Pro Ser His Ser Asp Phe Pro Lys Trp
35 40 45

Arg Gln Lys Gly Lys Leu Pro Asp Ile Ala Ala Lys Cys Leu Ile Asp
50 55 60

Arg Leu Lys Thr Asp Ala Arg Phe Ala Asn Gln Pro Val Trp Tyr Cys
65 70 75 80

Ile Ser Ala Gln Lys His Val Thr Tyr Ile Phe Arg Ser Trp Leu Ala
85 90 95

Ile Gln Arg Arg Lys Gln Trp Lys Leu Glu Gly Lys Arg Arg Trp Leu
100 105 110

Glu Ile Leu Gln Pro Asp Thr Ala Leu Ala Glu Lys Ala Gly Cys Ser
115 120 125

Ile Glu Ala Leu Arg Glu Glu Ala Gly Glu Ile Ile Ala Glu Ala Glu
130 135 140

Ile Val Asp Pro Phe Arg His Leu Leu Ala Lys Tyr Arg Glu Thr Glu
145 150 155 160

Asp Val Arg Glu Gln Cys Ala Ile Ala Tyr Leu Leu Lys Arg Lys Ala
165 170 175

```
Glu Leu Glu Pro Glu Glu Glu Asn Leu Ala Lys Leu Val Glu Arg His
            180                 185                 190

Arg Lys Thr Glu Ile Phe Ile Gln Arg Leu Glu Ala Gln Leu Asp Ala
        195                 200                 205

Ser Leu Pro Lys Gly Arg Asp Leu Thr Gly His Leu Gln Ala Glu Ala
    210                 215                 220

Leu Ile Gln Ser Ile His Ser Pro Leu Leu Asp Asp Ser Asn Tyr Asn
225                 230                 235                 240

Thr Trp Lys Asp Ala Leu Thr Thr Glu Pro Ala Lys Phe Pro Phe Pro
                245                 250                 255

Ile Met Tyr Glu Thr Thr Glu Ser Leu Val Leu Ser Arg Asp Asp Arg
            260                 265                 270

Gly Arg Ile Leu Leu Arg Phe Ser Gly Leu Ser Gln Gln Thr Tyr Lys
        275                 280                 285

Ile Tyr Cys Asp Lys Pro His Gln His Trp Phe Glu Arg Phe Phe Glu
    290                 295                 300

Asp Gln Glu Thr Lys Arg Val Gly Gly Asp Gln His Ser Ala Ala Ala
305                 310                 315                 320

Phe Thr Leu Arg Ser Ala Gln Leu Met Trp Val Pro Ser Glu Lys His
                325                 330                 335

Arg Asp Glu Pro Asp Pro Trp Asn Arg Tyr Tyr Leu Asn Leu Ser Cys
            340                 345                 350

Thr Val Asp Thr Arg Leu Trp Thr Gln Glu Gly Thr Lys Val Val Val
        355                 360                 365

Gln Glu Lys Ala Val Lys Thr Ala Gln Lys Leu Thr Ser Met Gln Glu
    370                 375                 380

Lys Lys Ser Leu Thr Gln Thr Gln Arg Gly Tyr Ile Arg Arg Leu Glu
385                 390                 395                 400
```

```
Ser Thr Leu Gln Arg Leu Gln Asn Pro Tyr Pro Arg Pro Ser Arg Thr
                405                 410                 415

Val Tyr Trp Gly Gln Pro Glu Ile Leu Val Gly Val Ser Met Ser Leu
            420                 425                 430

Asp Lys Thr Val Thr Ile Ala Val Val Asn Ala Leu Thr Glu Gln Val
        435                 440                 445

Leu Thr Tyr Arg Ser Ala Lys Gln Leu Leu Gly Glu Arg Tyr Arg Leu
    450                 455                 460

Leu Gln Arg Ala Arg Lys Glu Ile Val Lys Ile Ser His Gln Gly His
465                 470                 475                 480

Arg Gln Arg Arg Lys Gly Gly Arg Arg Ile Ser Gln Glu Ser Asp Val
                485                 490                 495

Gly Lys Tyr Val Asp Arg Leu Ile Ala Lys Ala Ile Asp Thr Leu Ala
            500                 505                 510

Leu Lys Tyr Arg Ala Gly Ser Ile Val Leu Pro Asn Leu Ala Tyr Ile
        515                 520                 525

Arg Glu Ser Ile Glu Ala Glu Val Gln Gln Arg Ala Ile Glu Lys Val
    530                 535                 540

Pro Asp Phe Lys Asp Gly Gln Lys Gln Tyr Ala Lys Ala Tyr Arg Thr
545                 550                 555                 560

Gln Ile His Arg Trp Pro Phe Ser Arg Leu Gln Ser Ala Ile Ile Ser
                565                 570                 575

Lys Ala Glu Gln Ser Gly Ile Thr Ile Glu Ile Ala Thr Gln Gln Ile
            580                 585                 590

Ser Gly Ser Phe Gln Asp Lys Ala Arg Glu Leu Gly Leu Gln Ala Tyr
        595                 600                 605
```

Ala His Arg Ser Ala Ser
610

<210> 72
<211> 440
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-25 effector

<400> 72
Leu Lys Asn Arg Gln Arg Leu Pro Lys Lys Ala Glu Asp Pro Glu Lys
1 5 10 15

Leu Ala Gln Arg Arg Arg Lys Val Glu Ile Lys Ile Glu Arg Leu Arg
20 25 30

Asp Gln Ile Glu Ser Cys Ile Pro Lys Gly Arg Asp Ile Thr Gly Glu
35 40 45

Ala Trp Leu Asn Thr Leu Leu Thr Ala Ala Asn Thr Val Pro Gln Ser
50 55 60

Ala Gln Glu Ala Lys Thr Trp Gln Asp Ile Leu Leu Thr Arg Ser Lys
65 70 75 80

Ser Val Pro Tyr Pro Val Ala Tyr Glu Thr Asn Glu Asp Leu Thr Trp
85 90 95

Ser Lys Asn Glu Lys Asn Arg Leu Cys Val Arg Phe Asn Gly Met Ser
100 105 110

Asp Tyr Ser Phe Gln Ile Tyr Cys Asp Gln Arg Gln Leu Lys Trp Phe
115 120 125

Glu Arg Phe Phe Gln Asp Gln Glu Thr Lys Lys Gln Ser Gln Asn Gln
130 135 140

His Pro Ser Ala Leu Phe Thr Leu Arg Ser Ala Arg Ile Phe Trp Gln

```
145                 150                 155                 160

Glu His Gln Gly Lys Gly Gln Pro Trp Gln Lys His Arg Leu Ile Leu
                165                 170                 175

Tyr Cys Ser Leu Asp Thr Arg Phe Trp Thr Ala Glu Gly Thr Glu Gln
            180                 185                 190

Ile Arg Gln Glu Lys Ala Ala Lys Ile Thr Lys Thr Leu Ser Asn Leu
        195                 200                 205

Gln Ala Lys Asp Glu Leu Thr Asp Ser Gln Gln Ala Phe Leu Arg Arg
    210                 215                 220

Gln Thr Ser Thr Leu Glu Arg Ile Asn Thr Pro Phe Pro Arg Pro Asn
225                 230                 235                 240

Gln Ala Ala Tyr Gln Gly Gln Pro His Ile Leu Val Gly Val Ala Leu
                245                 250                 255

Gly Leu Glu Lys Pro Ala Thr Val Ala Val Val Asp Ile Asn Thr Gly
            260                 265                 270

Arg Ala Ile Ala Tyr Arg Ser Ile Lys Gln Leu Leu Gly Ser Asn Tyr
        275                 280                 285

Lys Leu Leu Asn Arg Gln Arg Arg Gln Lys Gln Arg Asn Ala His Gln
    290                 295                 300

Arg His Lys Ala Gln Lys Gln Asn Lys Pro Asn Gln Phe Gly Asp Ser
305                 310                 315                 320

Asn Leu Gly Glu Tyr Ile Asp Arg Leu Leu Ala Lys Ala Ile Leu Ser
                325                 330                 335

Val Ala Gln Gln Tyr Gln Ala Ser Ser Ile Val Leu Pro Lys Leu Gly
            340                 345                 350

Asp Leu Arg Glu Leu Val Gln Ser Glu Val Lys Val Lys Ala Glu Ala
        355                 360                 365
```

Lys Ile Pro Gly Tyr Leu Glu Gly Gln Glu Gln Tyr Ala Lys Gln Tyr
370 375 380

Arg Val Ser Val His Gln Trp Ser Tyr Gly Arg Leu Ile Asp Asn Ile
385 390 395 400

His Gln Gln Ala Ser Lys Ile Gly Ile Ala Ile Glu Gln Gly Gln Gln
405 410 415

Pro Thr Arg Ala Ser Pro Gln Asn Gln Ala Arg Glu Leu Ala Ile Thr
420 425 430

Ala Tyr Gln Ser Arg Thr Gln Lys
435 440

<210> 73
<211> 561
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-26 effector

<400> 73
Gln His Lys Asp Phe Pro Gln Trp Val Ser Ala Gly Glu Ile Pro Ala
1 5 10 15

Gln Val Val Lys Asn Leu Val Asn Gln Ala Gln Ser Gly Leu Pro Ala
20 25 30

Arg Trp Cys Ala Ser Ala Gln Arg Gln Val Gln Glu Thr Tyr Lys Ala
35 40 45

Trp Ile Thr Lys Arg Arg Lys Leu Gln Gln Lys Leu Gln Gly Gln Gln
50 55 60

Thr Trp Leu Ser Val Leu Arg Pro Asp Ala Glu Leu Ala Lys Asn Ala
65 70 75 80

```
Gly Leu Ser Leu Glu Glu Met Lys Ile Arg Ala Gln Ala Leu Leu His
                85                  90                  95

Arg Glu Ile His Asn Trp Phe Gln Val Tyr Gln Gln Cys Gln Asp Val
            100                 105                 110

Val Glu Arg Ser Ile Phe Ala Tyr Leu Leu Lys His Arg Leu Thr Val
        115                 120                 125

Pro Ala Glu Pro Glu Asn Thr Asp Lys Leu Arg Gln Lys Arg Arg Gln
    130                 135                 140

Val Glu Ile Lys Ile Glu Trp Leu Glu Thr Gln Leu Ala Gly Arg Ser
145                 150                 155                 160

Pro Gln Gly Arg Asp Leu Thr Gly Ser Arg Tyr Val Ala Ala Leu Asn
                165                 170                 175

Glu Gly Glu Gln Cys Tyr Trp Glu Asn Asp Ala Asp Phe Leu Ala Trp
            180                 185                 190

Gln Ala Glu Ile Leu Ser Arg Pro Asp Ser Leu Pro Pro Pro Val Glu
        195                 200                 205

Tyr Ala Thr Asn Thr Asp Met Thr Trp His Lys Asp Glu Gln Gly Arg
    210                 215                 220

Leu Ala Val Thr Phe Asn Gly Leu Gly Lys Leu Lys Phe Lys Ile Ala
225                 230                 235                 240

Cys Asp Gln Arg Gln Leu His Trp Phe Gln Arg Phe Tyr Gln Asp Gln
                245                 250                 255

Glu Gln Phe Lys Ser Gln Lys Gly Gln Arg Ser Gln Ala Leu Phe Thr
            260                 265                 270

Leu Arg Ser Ala Glu Leu Leu Trp Lys Pro Gly Asn Arg Ser Gly Asp
        275                 280                 285

Pro Trp Gln Ala Asn Phe Leu Tyr Leu His Cys Thr Val Asp Pro Arg
    290                 295                 300
```

```
Leu Trp Thr Gln Glu Gly Thr Ala Met Val Gln Gln Glu Lys Ala Lys
305                 310                 315                 320

Lys Ser Gln Ala Ile Val Lys Lys Leu Ser Glu Arg Ser Asp Leu Thr
                325                 330                 335

Ala Gln Gln Lys Asp Cys Leu Gln Arg His Gln Ser Thr Leu Ala Arg
            340                 345                 350

Leu His Met Gly Tyr Asp Arg Pro Gln Arg Arg Met Tyr Gln Gly Lys
        355                 360                 365

Ser His Leu Val Val Gly Ile Ser Leu Asp Met Glu Asn Leu Val Thr
    370                 375                 380

Val Ala Leu Val Asp Val Val Lys Gln Glu Val Ile Thr Gly Cys Thr
385                 390                 395                 400

Met Lys Ser Leu Leu Gly Gln Asp Tyr Ala Leu Val Arg Arg Leu Arg
                405                 410                 415

Tyr Glu Lys Arg Gln Asn Ser His Leu Arg Lys Val Ala Gln Glu Arg
            420                 425                 430

Gly Ser Lys Ile Val Asn His Glu Ala Asn Leu Ala Ile His Val Glu
        435                 440                 445

Arg Leu Leu Val Lys Ala Ile Ile His Phe Ala Gln Gln His Leu Ala
    450                 455                 460

Gly Ser Leu Cys Val Pro Thr Leu Lys Asp Ile Arg Glu Thr Ile Gln
465                 470                 475                 480

Ala His Leu Gln Cys Arg Ala Glu Glu Arg Phe Pro Asp Ser Lys Glu
                485                 490                 495

Leu Gln Arg Arg Tyr Ala Lys Glu Tyr Arg Ile Asn Ala His Arg Trp
            500                 505                 510
```

Ser Tyr Asn Arg Leu Leu Lys Leu Leu Asn Gln Gln Ala Lys Phe Ala
515 520 525

Gly Leu Val Val Glu Gln Gly Val Gln Ser Ala Gly Glu Thr Ala Leu
530 535 540

Glu Arg Ala Leu Gly Val Ala Leu Ser Ala Tyr Tyr Gln Arg Ser Ala
545 550 555 560

Ala

<210> 74
<211> 583
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-27 effector

<400> 74
Gln Val Trp Glu Met Met Ala Gly Arg Tyr Ala Ser Leu Ile Ala Thr
1 5 10 15

Thr Leu Glu Gln Val Ser Gln His Lys Asp Phe Pro Gln Trp Val Ser
20 25 30

Ala Gly Glu Ile Pro Ala Gln Val Val Lys Asn Leu Val Asn Gln Ala
35 40 45

Gln Ser Gly Leu Pro Ala Arg Trp Cys Ala Ser Ala Gln Arg Gln Val
50 55 60

Gln Glu Thr Tyr Lys Ala Trp Leu Thr Lys Arg Arg Lys Leu Gln Gln
65 70 75 80

Lys Leu Gln Gly Gln Gln Thr Trp Leu Ser Val Leu Arg Pro Asp Ala
85 90 95

Glu Leu Ala Lys Asn Ala Gly Leu Ser Leu Glu Glu Met Lys Ile Arg

```
            100                 105                 110

Ala Gln Ala Leu Leu His Arg Glu Ile Asn Asn Trp Phe Gln Val Tyr
        115                 120                 125

Gln Gln Cys Gln Asp Val Val Glu Arg Ser Ile Phe Ala Tyr Leu Leu
    130                 135                 140

Lys His Arg Leu Thr Val Pro Thr Glu Pro Glu Asp Thr Asp Lys Leu
145                 150                 155                 160

Arg Arg Lys Arg Arg Gln Val Glu Ile Lys Ile Glu Arg Leu Glu Thr
                165                 170                 175

Gln Leu Ala Gly Arg Ser Pro Gln Gly Arg Asp Leu Thr Gly Ser Arg
            180                 185                 190

Tyr Val Ala Ala Leu Asn Glu Gly Glu Gln Cys Tyr Trp Glu Asn Asp
        195                 200                 205

Ala Asp Phe Leu Ala Trp Gln Ala Glu Ile Leu Ser Cys Pro Asp Ser
    210                 215                 220

Leu Pro Pro Pro Val Glu Tyr Ala Thr Asn Thr Asp Met Thr Trp His
225                 230                 235                 240

Lys Asp Glu Gln Gly Arg Leu Ala Val Thr Phe Asn Gly Leu Gly Lys
                245                 250                 255

Leu Lys Phe Lys Ile Ala Cys Asp Gln Arg Gln Leu His Trp Phe Gln
            260                 265                 270

Arg Phe Tyr Gln Asp Gln Glu Gln Phe Lys Ser Gln Lys Gly Gln Arg
        275                 280                 285

Ser Gln Ala Leu Phe Thr Leu Arg Ser Ala Glu Leu Leu Trp Lys Pro
    290                 295                 300

Gly Asn Arg Ser Gly Asp Pro Trp Gln Ala Asn Phe Leu Tyr Leu His
305                 310                 315                 320
```

```
Cys Thr Val Asp Pro Arg Leu Trp Thr Gln Glu Gly Thr Ala Met Val
                325                 330                 335

Gln Gln Glu Lys Ala Lys Lys Ser Gln Ala Ile Val Lys Lys Leu Ser
            340                 345                 350

Glu Arg Ser Asp Leu Thr Ala Gln Gln Arg Asp Cys Leu Gln Arg His
        355                 360                 365

Gln Ser Thr Leu Ala Arg Leu His Met Gly Tyr Asp Arg Pro Gln Arg
    370                 375                 380

Arg Met Tyr Gln Gly Lys Ser His Leu Val Val Gly Ile Ser Leu Asp
385                 390                 395                 400

Met Glu Asn Leu Val Thr Val Ala Leu Val Asp Val Val Lys Gln Lys
                405                 410                 415

Val Ile Thr Gly Cys Thr Met Lys Ser Leu Leu Gly Gln Asp Tyr Ala
            420                 425                 430

Leu Val Arg Arg Leu Arg Tyr Glu Lys Arg Gln Asn Ser His Leu Arg
        435                 440                 445

Lys Val Ala Gln Glu Arg Gly Ser Lys Ile Val Asn Tyr Glu Ala Asn
    450                 455                 460

Leu Ala Ile His Val Glu Arg Leu Leu Val Lys Ala Ile Ile His Phe
465                 470                 475                 480

Ala Gln Gln His Leu Ala Gly Ser Leu Cys Val Pro Thr Leu Lys Asp
                485                 490                 495

Ile Arg Glu Thr Ile Gln Ala His Leu Gln Cys Arg Ala Glu Glu Arg
            500                 505                 510

Phe Pro Asp Ser Lys Glu Leu Gln Arg Arg Tyr Ala Lys Glu Tyr Arg
        515                 520                 525

Ile Asn Ala His Arg Trp Ser Tyr Asn Arg Leu Leu Lys Leu Leu Asn
```

530 535 540

Gln Gln Ala Lys Phe Ala Gly Leu Val Val Glu Gln Gly Val Gln Ser
545 550 555 560

Ala Gly Glu Thr Ala Leu Glu Arg Ala Leu Gly Val Ala Leu Ser Ala
565 570 575

Tyr Tyr Gln Arg Ser Ala Ala
580

<210> 75
<211> 417
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-28 effector

<400> 75
Val Glu Ile Lys Ile Glu Arg Leu Glu Thr Gln Leu Ala Gly Arg Ser
1 5 10 15

Pro Gln Gly Arg Asp Leu Thr Gly Ser Arg Tyr Ala Thr Ala Leu Asn
20 25 30

Asp Gly Glu Gln Cys Asp Trp Glu Asp Asp Ala Asp Phe Leu Ala Trp
35 40 45

Gln Ala Glu Ile Leu Ser Arg Pro Asp Ser Leu Pro Pro Pro Val Glu
50 55 60

Tyr Ala Thr Asn Thr Asp Met Thr Trp His Lys Asp Glu Gln Gly Arg
65 70 75 80

Leu Ala Val Thr Phe Asn Gly Leu Gly Lys Leu Lys Phe Lys Ile Ala
85 90 95

Cys Asp Gln Arg Gln Leu His Trp Phe Gln Arg Phe Tyr Gln Asp Gln
100 105 110

```
Glu Gln Phe Lys Ser Gln Lys Gly Gln Arg Ser Gln Ala Leu Phe Thr
        115                 120                 125

Leu Arg Ser Ala Glu Leu Leu Trp Lys Pro Gly Asn Arg Ser Gly Asp
    130                 135                 140

Pro Trp Gln Ala Asn Phe Leu Tyr Leu His Cys Thr Val Asp Ser Arg
145                 150                 155                 160

Leu Trp Thr Gln Glu Gly Thr Ala Met Val Gln Gln Glu Ile Ala Lys
                165                 170                 175

Lys Ser Gln Ala Ile Val Lys Lys Leu Ser Glu Arg Ser Asp Leu Thr
            180                 185                 190

Ala Gln Gln Lys Asp Cys Leu Gln Arg His Gln Ser Thr Leu Ala Arg
        195                 200                 205

Leu His Met Gly Tyr Asp Arg Pro Gln Arg Arg Met Tyr Gln Gly Lys
    210                 215                 220

Ser His Leu Val Val Gly Ile Ser Leu Asp Met Glu Asn Leu Val Thr
225                 230                 235                 240

Val Ala Leu Val Asp Val Val Lys Gln Lys Val Ile Thr Gly Cys Thr
                245                 250                 255

Met Lys Ser Leu Leu Gly Gln Asp Tyr Ala Leu Val Gln Arg Leu Arg
            260                 265                 270

Tyr Glu Lys Arg Gln Asn Ser His Leu Arg Lys Val Ala Gln Glu Arg
        275                 280                 285

Gly Ser Lys Ile Val Asn Tyr Glu Ala Asn Leu Ala Ile His Val Glu
    290                 295                 300

Arg Leu Leu Val Lys Ala Ile Ile His Phe Ala Gln Gln His Leu Ala
305                 310                 315                 320
```

Gly Ser Leu Cys Val Pro Thr Leu Lys Asp Ile Arg Glu Thr Ile Gln
325 330 335

Ala His Leu Gln Cys Arg Ala Glu Glu Arg Phe Pro Asp Ser Lys Glu
340 345 350

Leu Gln Arg Arg Tyr Ala Lys Glu Tyr Arg Ile Asn Ala His Arg Trp
355 360 365

Ser Tyr Asn Arg Leu Leu Lys Leu Leu Asn Gln Gln Ala Lys Phe Ala
370 375 380

Gly Leu Val Val Glu Gln Gly Val Gln Ser Ala Gly Glu Thr Ala Leu
385 390 395 400

Glu Arg Ala Leu Gly Val Ala Leu Ser Ala Tyr Tyr Gln Arg Ser Ala
405 410 415

Ala

<210> 76
<211> 488
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-29 effector

<400> 76
Met Gly Lys Asp Ile Leu Asp Glu Ile Ala Ser Ile Gly Val Ile Glu
1 5 10 15

Asp Glu Thr Glu Trp Asn Ser Ile Glu Ser Ala Leu Leu Lys Lys Gln
20 25 30

Thr Leu Met Pro His Pro Ile Leu Phe Glu Ser Ser Asp Asp Phe Ile
35 40 45

Trp Ala Glu Pro Ala Arg Phe Gln Ile Glu Asp Leu Gln Ala Glu Glu

```
    50                  55                  60

Glu Ile Ser Arg Asp Thr Glu Ala Lys Glu Leu Lys Ala Lys Asn Leu
65                  70                  75                  80

Gln Ala Glu Lys Glu Ser Ser Asn Thr Glu Pro Arg Lys Arg Val Cys
                85                  90                  95

Val Cys Phe Lys Ser Phe Asp Glu Lys Tyr Thr Phe Glu Val Ala Gly
            100                 105                 110

Asp Tyr Arg His Ile His Ala Val Trp Gln Ala Leu Lys Glu Arg Lys
        115                 120                 125

Glu Tyr Asp Lys Asn Ala Asp Lys Asn Thr Ser Ala Leu Phe Leu Val
    130                 135                 140

Arg Ser Ala Thr Leu Ile Trp Arg Glu Tyr Lys Lys Asn Asp Asn Arg
145                 150                 155                 160

Ile Val Arg Arg Arg Lys Ala Ala Lys Lys Arg Ala Lys Arg Glu Gly
                165                 170                 175

Leu Val Ala Ser Thr Glu Ser Asp Ser Ser Leu Ala Pro Glu Phe Tyr
            180                 185                 190

Asp Pro Glu Phe Pro Trp Asn Arg Tyr Gln Leu Phe Leu His Cys Thr
        195                 200                 205

Ile Glu Thr Arg Tyr Leu Ser Lys Glu Gly Thr Glu Ile Asp Ile Glu
    210                 215                 220

Lys Gln Lys Lys Ser Thr Ala Lys Ser Ile Gln Thr Leu Glu Lys Asn
225                 230                 235                 240

Ile Ala Glu Leu Glu Lys Lys Gly Glu Ser Ala Lys Thr Arg Lys Asp
                245                 250                 255

Arg His Asn Arg Arg Ser Gly Thr Leu Arg Arg Leu Glu Ser Tyr Asp
            260                 265                 270
```

```
Asn Asn Tyr Glu Arg Ser Ser Lys Pro Leu Tyr Ala Gly Gln Pro His
        275                 280                 285

Ile Val Thr Gly Val Ala Leu Gly Ser Ser Gly Leu Val Thr Thr Thr
    290                 295                 300

Ile Ser Asp Ala Ile Ser Gly Lys Ile Leu Glu Cys Arg Gly Leu Glu
305                 310                 315                 320

Ala Leu Leu Gly Lys Asp Tyr Glu Leu Val Lys Arg Arg Gln Phe Glu
                325                 330                 335

Arg Gln Leu Asn Ser Arg Arg Arg Thr Gln Asn Gln Lys Arg Gly Ala
            340                 345                 350

Asn Asp Gln Phe Gly Glu Ser Asn Leu Gly Asp Thr Ile Asp Lys Arg
        355                 360                 365

Ile Ala Asn Ala Val Ile Asp Phe Ala Lys Lys His Gln Ser Gly Phe
    370                 375                 380

Ile Val Leu Pro Asp Met Asn Asp Tyr Arg Arg Arg Lys Gln Ser Glu
385                 390                 395                 400

Ile Ala Ala Phe Ala Glu Arg Glu Cys Gly Gly Trp Lys Gly Ile Glu
                405                 410                 415

Lys Lys Phe Ala Lys Ala Gln Lys Glu Lys Ile His Ser Trp Ser Tyr
            420                 425                 430

Gly Arg Leu Thr Thr Tyr Ile Thr Asn Gln Ala Glu Lys Glu Gly Ile
        435                 440                 445

Leu Val Arg Thr Arg Arg Gln Pro Ile Gln Gly Ser Ser Gln Glu Gln
    450                 455                 460

Gly Lys Leu Met Ala Ile Glu Ala Tyr Lys Asp Lys Pro Lys Pro Lys
465                 470                 475                 480

Lys Ser Arg Pro Lys Lys Ser Ala
```

485

<210> 77
<211> 467
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-30 effector

<400> 77
Lys Thr Ala Glu Ile Leu Ala Gln Thr Thr Pro Gln Pro Ala Glu Ala
1 5 10 15

Arg Thr Ala Lys Ser Ser Lys Lys Lys Gln Lys Ser Lys Leu Asn Gly
20 25 30

Gln Lys Gln Ser Phe Asp Leu Leu Phe Glu Gln Tyr Gln Glu Thr Asp
35 40 45

Asp Glu Leu Thr Arg Ser Ala Ile Val Tyr Leu Leu Lys Asn Gly Arg
50 55 60

Lys Leu Pro Asn Glu Pro Glu Asp Pro Lys Gln Phe Ala Lys Arg Arg
65 70 75 80

Arg Lys Ala Glu Ile Arg Val Glu Arg Thr Thr Asn Thr Leu His Arg
85 90 95

Met Gln Leu Pro Thr Gly Arg Asp Met Thr Asp Gln Glu Trp Leu Lys
100 105 110

Thr Leu Ala Thr Ala Val Ser Asn Val Pro Glu Asp Asp Gly Glu Ala
115 120 125

Ala Ser Trp Gln Ala Ile Leu Met Ser Glu Ala His Lys Leu Pro Phe
130 135 140

Pro Ile Leu Tyr Glu Thr Asn Glu Asp Leu Ser Trp Phe Leu Asn Asp
145 150 155 160

```
Gly Glu Arg Leu Cys Val Thr Phe Asn Gly Leu Ser Glu His Thr Phe
                165                 170                 175

Glu Ile His Cys Asp Arg Arg Gln Leu His Trp Phe Lys Arg Phe Leu
            180                 185                 190

Glu Asp Gln Glu Ile Lys Lys Ala Ser Lys Asn Gln His Ser Cys Ala
        195                 200                 205

Leu Phe Thr Leu Arg Ser Ala Arg Ile Ala Trp Gln Glu Gly Lys Glu
    210                 215                 220

Lys Gly Glu Pro Trp Asn Thr His Arg Leu Thr Leu Ala Cys Thr Val
225                 230                 235                 240

Glu Thr Arg Ala Trp Thr Ala Glu Gly Thr Glu Gln Ile Arg Gln Glu
                245                 250                 255

Lys Ala Ala Glu Cys Lys Lys Val Ile Leu Gly Thr Lys Ala Lys Gly
            260                 265                 270

Asn Leu Asn Gln Arg Gln Glu Gln Phe Ile Gln Arg Arg Glu Lys Thr
        275                 280                 285

Ile Val Leu Met Asn Asn Ser Phe Pro Arg Pro Ser Arg Pro Leu Tyr
    290                 295                 300

Gln Gly Gln Pro Glu Ile Ile Ala Ser Val Ser Tyr Gly Leu Asp Arg
305                 310                 315                 320

Pro Ala Thr Leu Ala Ile Val Asp Ile Thr Thr Gly Lys Ala Ile Ala
                325                 330                 335

Tyr Arg Ser Ile Arg Gln Leu Leu Gly Asp His Tyr Lys Leu Leu Thr
            340                 345                 350

Arg Tyr Arg Leu Arg Gln Gln Gln Asn Ala His Arg Arg His Asn Arg
        355                 360                 365
```

Gln Arg Lys Gly Ala Ser Asn Arg Ile Gln Glu Ser Gln Leu Gly Glu
370 375 380

Tyr Leu Asp Cys Leu Ile Ala Gln Ala Ile Val Ser Val Ala Gln Thr
385 390 395 400

Tyr Gln Ala Ser Ser Ile Val Leu Pro Asp Leu Gly Asn Ile Arg Glu
405 410 415

Val Val Glu Ala Glu Val Arg Ala Arg Ala Glu Gln Arg Ile Val Gly
420 425 430

Tyr Leu Glu Gly Gln Gln Gln Tyr Ala Lys Gln Tyr Arg Ala Ser Ile
435 440 445

His Arg Trp Ser Tyr Gly Arg Leu Ser Glu Lys Ile Gln Ser Gln Ala
450 455 460

Ala Gln Ala
465

<210> 78
<211> 550
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-31 effector

<400> 78
Pro Leu Ile Asn Ala Leu Ile Arg Glu Leu Val Ser His Pro Glu Phe
1 5 10 15

Glu Ser Trp Arg Glu Lys Gly Arg His Pro Ser Asp Val Val Ala Lys
20 25 30

Leu Asn Arg Thr Leu Lys Thr Asp Ala Gln Phe Ala Gly Gln Pro Ser
35 40 45

Arg Phe Phe Met Ser Ala Glu Lys Val Val Asn Tyr Ile Phe Lys Ser

```
    50                  55                  60

Trp Phe Lys Ile Gln Ser Arg Leu Gln Gln Lys Leu Ser Gly Lys Gln
65                  70                  75                  80

Lys Trp Leu Glu Ile Leu Lys Ser Asp Glu Glu Leu Val Glu Leu Cys
                85                  90                  95

Gly Gln Ser Ile Asp Val Leu Gln Lys Lys Ala Val Gln Val Leu Gln
            100                 105                 110

Glu Val Gly Lys Ala Ile Asp Lys Glu Thr Ser Asp Thr Ile Pro Glu
        115                 120                 125

Lys Lys Arg Ala Asn Leu Ile Arg Gln Arg Leu Phe Lys Lys Leu Ser
    130                 135                 140

Ser Ala Lys Gln Pro Leu Pro Gln Tyr Ala Ala Ile Tyr Leu Leu Lys
145                 150                 155                 160

Asn His Cys Lys Ile Pro Asn Glu Pro Glu Asp Val Glu Asn Phe Ala
                165                 170                 175

Gln Arg Arg Arg Lys Thr Glu Ile Gln Ile Gln Ile Leu Gln Asp Gln
            180                 185                 190

Ile Glu Ser Arg Ile Pro Lys Gly Arg Asp Leu Thr Gly Lys Ala Trp
        195                 200                 205

Thr Glu Thr Leu Leu Lys Ala Val Asn Ser Ile Pro Gln Asp Asn Gln
    210                 215                 220

Glu His Lys Gln Trp Gln Asp Arg Leu Leu Ser Met Pro Ser Thr Ala
225                 230                 235                 240

Pro Phe Pro Ile Leu Phe Glu Thr Asn Glu Asp Leu Val Trp Ser Ala
                245                 250                 255

Asn Glu Arg Asp Arg Ile Cys Val Arg Phe Asn Gly Leu Ser Glu His
            260                 265                 270
```

```
Thr Phe Gln Ile Tyr Cys Asp Gln Arg Gln Leu His Trp Phe Lys Arg
        275                 280                 285

Phe Leu Glu Asp Gln Lys Thr Lys Arg Ala Ser Lys Asn Gln His Ser
    290                 295                 300

Ser Ala Leu Phe Thr Leu Arg Ser Ala Arg Ile Ser Trp Gln Glu Asn
305                 310                 315                 320

Asp Arg Lys Gly Asn Pro Trp Glu Thr His Tyr Ile Thr Leu Ser Cys
                325                 330                 335

Thr Val Asp Val Arg Leu Trp Ser Ala Glu Gly Thr Asp Glu Val Arg
            340                 345                 350

His Glu Lys Ala Ala Asp Val Ala Lys Val Leu Thr Arg Leu Asn Glu
        355                 360                 365

Lys Asp Ser Leu Thr Glu Thr Gln Ala Gly Tyr Ala Arg Arg Leu Thr
    370                 375                 380

Ser Thr Leu Glu Lys Leu Ser Ser Pro Phe Glu Arg Pro Ser Leu Pro
385                 390                 395                 400

Arg Tyr Gln Gly Lys Asp Asn Ile Ile Ala Gly Ile Ser Leu Gly Trp
                405                 410                 415

Asp Lys Pro Val Thr Leu Ala Ile Trp Asn Thr Asp Thr Gln Glu Val
            420                 425                 430

Leu Thr Tyr Arg Ser Leu Lys Gln Leu Leu Gly Lys Asp Tyr Pro Leu
        435                 440                 445

Tyr Leu Arg His Arg Arg Glu Gln Gln Lys Gln Ser His Glu Arg His
    450                 455                 460

Lys Ala Gln Lys Gln Gly Lys Gly Asn Arg Phe Gly Thr Ser Asn Leu
465                 470                 475                 480

Gly Glu His Leu Asp Arg Leu Leu Ala Lys Ala Ile Val Ala Ile Ala
```

485 490 495

Gln Gln Tyr Ser Ala Gly Ser Ile Ala Val Pro Lys Leu Asp Ser Ile
500 505 510

Arg Asp Thr Leu Gln Ala Glu Ile Asp Ala Lys Ala Glu Gln Arg Ile
515 520 525

Pro Gly Tyr Leu Glu Gly Gln Lys Arg Tyr Ser Arg Asp Tyr Lys Arg
530 535 540

Ser Ile His Arg Trp Ser
545 550

<210> 79
<211> 446
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-32 effector

<400> 79
Ser Ile Phe Ala Tyr Leu Leu Lys His Arg Leu Thr Val Pro Thr Glu
1 5 10 15

Ser Glu Asp Thr Asp Glu Leu Arg Arg Lys Arg Arg Gln Val Glu Ile
20 25 30

Lys Ile Glu Arg Leu Glu Thr Gln Leu Ala Gly Arg Ser Pro Gln Gly
35 40 45

Arg Asp Leu Thr Gly Ser Arg Tyr Ala Thr Ala Leu Asn Asp Gly Glu
50 55 60

Gln Cys Asp Trp Glu Asp Asp Ala Asp Phe Leu Ala Trp Gln Ala Glu
65 70 75 80

Ile Leu Ser Arg Pro Asp Ser Leu Pro Pro Pro Val Glu Tyr Ala Thr
85 90 95

```
Asn Thr Asp Met Thr Trp His Lys Asp Glu Gln Gly Arg Leu Ala Val
            100                 105                 110

Thr Phe Asn Gly Leu Gly Lys Leu Lys Phe Lys Ile Ala Cys Asp Gln
        115                 120                 125

Arg Gln Leu His Trp Phe Gln Arg Phe Tyr Gln Asp Gln Glu Gln Phe
    130                 135                 140

Lys Ser Gln Lys Gly Gln Arg Ser Gln Ala Leu Phe Thr Leu Arg Ser
145                 150                 155                 160

Ala Glu Leu Leu Trp Lys Pro Gly Asn Arg Ser Gly Asp Pro Trp Gln
                165                 170                 175

Ala Asn Phe Leu Tyr Leu His Cys Thr Val Asp Ser Arg Leu Trp Thr
            180                 185                 190

Gln Glu Gly Thr Ala Met Val Gln Gln Glu Lys Ala Lys Lys Ser Gln
        195                 200                 205

Ala Ile Val Lys Lys Leu Ser Glu Arg Ser Asp Leu Thr Ala Gln Gln
    210                 215                 220

Lys Asp Cys Leu Gln Arg His Gln Ser Thr Leu Ala Arg Leu His Met
225                 230                 235                 240

Gly Tyr Asp Arg Pro Gln Arg Arg Met Tyr Gln Gly Lys Ser His Leu
                245                 250                 255

Val Val Gly Ile Ser Leu Asp Met Glu Asn Leu Val Thr Val Ala Leu
            260                 265                 270

Val Asp Val Val Lys Gln Lys Val Ile Thr Gly Cys Thr Met Lys Ser
        275                 280                 285

Leu Leu Gly Gln Asp Tyr Ala Leu Val Gln Arg Leu Arg Tyr Glu Lys
    290                 295                 300
```

```
Arg Gln Asn Ser His Leu Arg Lys Val Ala Gln Glu Arg Gly Ser Lys
305                 310                 315                 320

Ile Val Asn Tyr Glu Ala Asn Leu Ala Ile His Val Glu Arg Leu Leu
                325                 330                 335

Val Lys Ala Ile Ile His Phe Ala Gln Gln His Leu Ala Gly Ser Leu
            340                 345                 350

Cys Val Pro Thr Leu Lys Asp Ile Arg Glu Thr Ile Gln Ala His Leu
        355                 360                 365

Gln Cys Arg Ala Glu Glu Arg Phe Pro Asp Ser Lys Glu Leu Gln Arg
    370                 375                 380

Arg Tyr Ala Lys Glu Tyr Arg Ile Asn Ala His Arg Trp Ser Tyr Asn
385                 390                 395                 400

Arg Leu Leu Glu Leu Ile Ser Gln Gln Ala Ile Val Ala Gly Leu Leu
                405                 410                 415

Val Glu Gln Gly Leu Gln Gly Ser Gly Gly Thr Ala Ile Lys Arg Ala
            420                 425                 430

Leu Gly Val Ala Leu Ser Ala Tyr Tyr Gln Arg Ser Ala Ala
        435                 440                 445


<210> 80
<211> 649
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
      MG64 effector sequence

<220>
<223> MG64-44 effector

<400> 80
Met Ser Gln Ile Thr Ile Gln Cys Asn Leu Val Ala Ser Glu Ala Thr
1               5                   10                  15

Arg Gln Tyr Leu Trp His Leu Met Ala Asp Ile Tyr Thr Pro Phe Ile
```

```
            20                  25                  30

Asn Glu Met Leu Ala Thr Ile Ala Gln His Pro Asn Phe Glu Glu Trp
        35                  40                  45

Ser Gln Asn Gly Arg Ile Pro Ala Asp Val Phe Glu Asp Ile Arg Lys
    50                  55                  60

Thr Leu Lys Ala His Pro Asp Phe Gln Gly Met Pro Gly Arg Trp Tyr
65                  70                  75                  80

Tyr Ala Gly Arg Asp Leu Val Lys Arg Ile Phe Lys Ser Trp Leu Ala
                85                  90                  95

Leu Arg Arg Arg Leu Arg His Gln Leu Ser Gly Gln Thr His Trp Leu
            100                 105                 110

Glu Ile Phe Gln Ser Asp Asp Asp Leu Val Ala Ala Cys Gly Gln Asp
        115                 120                 125

Leu Pro Ala Ile Arg Ala Glu Ala Ala Ser Ile Leu Thr Lys Ile Gln
    130                 135                 140

Ile Glu Ala Pro Asn Thr Ser Lys Gln Pro Lys Lys Thr Lys Gln Pro
145                 150                 155                 160

Lys Lys Ala Gly Ser Lys Thr Gln Lys Pro Glu Glu Glu Gln Arg Asn
                165                 170                 175

Arg Asn Leu Phe Pro Ala Leu Phe Lys Glu Tyr Asp Gly Ala Glu Thr
            180                 185                 190

Glu Leu Val Lys Cys Ala Ile Ala Cys Leu Leu Lys Asn Asn Cys Gln
        195                 200                 205

Ile Pro Thr Lys Ala Glu His Pro Glu Lys Phe Gln Lys Arg Arg Arg
    210                 215                 220

Lys Thr Glu Ile Arg Val Glu Arg Ile Ile Glu Gln Leu Ala Arg Thr
225                 230                 235                 240
```

```
Arg Leu Pro Lys Gly Arg Asp Leu Thr Asn Glu Lys Trp Leu Asp Thr
                245                 250                 255

Leu Lys Met Ala Val Gln Gln Val Pro Lys Asp Glu Thr Glu Ala Ala
            260                 265                 270

Ala Trp Glu Ala Asp Leu Gln Thr Asp Ser Ser Pro Leu Pro Phe Pro
        275                 280                 285

Ile Ala Tyr Glu Ser Asn Glu Asp Leu Lys Trp Ser Gln Asn Ala Lys
    290                 295                 300

Gly Arg Leu Cys Val Arg Phe Asn Gly Leu Gly Lys His Thr Phe Glu
305                 310                 315                 320

Ile Tyr Cys Asp Thr Arg Gln Leu His Trp Phe Lys Arg Phe Leu Asp
                325                 330                 335

Asp Gln Thr Ile Lys Lys Gln Gly Gly Asn Ser His Ser Ala Gly Ala
            340                 345                 350

Leu Thr Leu Arg Ser Gly Arg Ile Ser Trp Arg Leu Asp Ser Ser Lys
        355                 360                 365

Gly Asn Pro Trp Asp Arg Asn Arg Leu Val Leu Phe Cys Ser Val Asp
    370                 375                 380

Thr Leu Leu Trp Thr Lys Glu Gly Thr Glu Lys Ala Ser Gln Glu Lys
385                 390                 395                 400

Ala Ser Lys Ile Ala Gln Val Ile Ser Gly Thr Lys Ala Lys Gly Asn
                405                 410                 415

Leu Thr Ser Lys Gln Glu Asp Phe Val Arg Lys Arg Glu Lys Thr Leu
            420                 425                 430

Ala Leu Leu Gln Asn Pro Phe Pro Arg Pro Ser Arg Pro Leu Tyr Gln
        435                 440                 445

Gly Ser Pro Ala Ile Leu Ala Gly Val Ser Phe Gly Leu Asp Lys Pro
```

```
    450                 455                 460

Ala Thr Leu Ala Ile Val Asp Val Thr Thr Gly Lys Ala Ile Ala Tyr
465                 470                 475                 480

Arg Ser Ile Arg Gln Leu Leu Gly Asp Asp His Lys Leu Leu Asn Arg
                485                 490                 495

Gln Arg Gln Arg Gln Arg Gln Lys Ala Gln Arg Arg Arg Ser Asn Gln
            500                 505                 510

Leu Lys Phe Ala Ser Asn Arg Ile Ser Glu Gly Gly Leu Gly Gly Gln
        515                 520                 525

Ile Asp Ser Leu Ile Ala Lys Ala Ile Val Gln Ile Ala Gln Gln Tyr
    530                 535                 540

Asn Ala Ser Ser Ile Val Leu Gly Asp Leu Ala Asn Ile Arg Glu Ile
545                 550                 555                 560

Ile Glu Ser Glu Ile Gln Ala Lys Ala Glu Gln Lys Thr Thr Leu Lys
                565                 570                 575

Glu Ile Gln Ala Lys Tyr Ala Arg Asp Tyr Arg Ala Ser Ile His Arg
            580                 585                 590

Trp Ser Tyr Lys Arg Leu Ala Gln Lys Ile Glu Ser Asn Ala Leu Gln
        595                 600                 605

Ala Gly Leu Ile Val Ala Thr Ile Lys Gln Pro Leu Ala Gly Ser Pro
    610                 615                 620

Gln Asp Lys Ala Arg Asp Val Ala Ile Ala Gly Phe Gln Ser Arg Ser
625                 630                 635                 640

Val Ser Lys Ile Leu Asp Thr Gly Ser
                645

<210> 81
<211> 429
<212> PRT
```

<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-45 effector

<400> 81
Met Pro Pro Cys Arg Pro Pro Ala Thr Arg Gln Tyr Leu Trp Gly Leu
1 5 10 15

Ala Ala Glu Lys Asn Thr His Leu Ile Asn Ala Leu Ile Gln Glu Val
20 25 30

Val Ser His Glu Asp Phe Glu Asp Trp Arg Leu Lys Gly Arg His Pro
35 40 45

Ala Asp Val Val Ala Lys Ala Cys Arg Arg Leu Lys Thr Glu Val Pro
50 55 60

Phe Ser Gly Gln Pro Ser Arg Phe Tyr Ala Ser Ala Glu Lys Ala Val
65 70 75 80

Asn Tyr Ile Phe Lys Ser Trp Phe Thr Leu Gln Ser Arg Leu Gln Arg
85 90 95

Gln Ile Ser Gly Lys Gln Ala Trp Leu Asp Met Leu Lys Ser Asp Asp
100 105 110

Glu Leu Val Glu Leu Cys Gly Arg Pro Leu Glu Gly Ile Gln Ala Thr
115 120 125

Ala Ala Gln Leu Leu Lys Gln Ile Glu Lys Thr Leu Glu Ser Glu Ala
130 135 140

Thr Glu Glu Ser Gln Gly Lys Ser Lys Glu Asp Gln Val Arg Thr His
145 150 155 160

Leu Phe Lys Lys Tyr His Ser Ala Lys Gln Ala Leu Asn Leu Cys Ala
165 170 175

```
Ala Ala Tyr Leu Leu Lys Asn Asn Ser Lys Ile Pro Asp Lys Pro Glu
            180                 185                 190

Asp Pro Asn Lys Phe Ala Gln Arg Gln Arg Lys Ala Glu Ile Gln Val
        195                 200                 205

Gln Arg Leu Gln Asp Gln Ile Glu Gly Arg Ile Pro Lys Gly Arg Asp
    210                 215                 220

Leu Thr Gly Gln Ala Trp Leu Ser Thr Leu Leu Thr Val Thr Thr Thr
225                 230                 235                 240

Val Pro Arg Asp Asn Arg Glu His Lys Gln Trp Gln Asp Arg Leu Leu
                245                 250                 255

Ala Gln Pro Arg Thr Ile Pro Phe Pro Ile Leu Phe Glu Thr Asn Glu
            260                 265                 270

Asp Leu Val Trp Ser Arg Asn Gln Ala Asp Arg Leu Cys Val Arg Phe
        275                 280                 285

Asn Gly Leu Ser Glu His Thr Phe Gln Ile Tyr Cys Asp Gln Arg Gln
    290                 295                 300

Leu Pro Trp Phe Gln Arg Phe Leu Glu Asp Gln Gln Thr Lys Arg Ala
305                 310                 315                 320

Ser Lys Asn Gln His Ser Ser Ala Leu Phe Ala Leu Arg Ser Ala Arg
                325                 330                 335

Ile Ser Trp Gln Glu Asn Asp Arg Lys Gly Gln Pro Trp Asp Ala His
            340                 345                 350

Ser Leu Thr Leu Ser Cys Thr Ile Asp Thr Arg Leu Trp Ser Ala Glu
        355                 360                 365

Gly Thr Asp Glu Val Arg Gln Glu Lys Ala Ala Asp Thr Ala Lys Val
    370                 375                 380

Leu Thr Arg Leu Asn Glu Lys Gly Ser Leu Ser Asp Thr Gln Ala Gly
385                 390                 395                 400
```

Tyr Ala Lys Arg Leu Thr Ser Thr Leu Asn Arg Leu Asp Ser Pro Phe
405 410 415

Lys Arg Pro Ser Gln Pro Arg Tyr Gln Gly Lys Pro His
420 425

<210> 82
<211> 477
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-46 effector

<400> 82
Glu Gln Gln Asn Arg Asn Leu Phe Pro Ala Leu Phe Lys Glu Tyr Asp
1 5 10 15

Gly Ala Glu Thr Glu Leu Val Lys Cys Ala Ile Ala Cys Leu Leu Lys
20 25 30

Asn Asn Cys Gln Ile Pro Thr Lys Ala Glu His Pro Glu Lys Phe Gln
35 40 45

Lys Arg Arg Arg Lys Thr Glu Ile Arg Val Glu Arg Ile Ile Glu Gln
50 55 60

Leu Ala Arg Thr Arg Leu Pro Lys Gly Arg Asp Leu Thr Asn Glu Lys
65 70 75 80

Trp Leu Asp Thr Leu Lys Met Ala Val Gln Gln Val Pro Lys Asp Glu
85 90 95

Thr Glu Ala Ala Ala Trp Glu Ala Asp Leu Gln Thr Asp Ser Ser Pro
100 105 110

Leu Pro Phe Pro Ile Ala Tyr Glu Ser Asn Glu Asp Leu Lys Trp Ser
115 120 125

```
Gln Asn Ala Lys Gly Arg Leu Cys Val Arg Phe Asn Gly Leu Gly Lys
    130                 135                 140

His Thr Phe Glu Ile Tyr Cys Asp Thr Arg Gln Leu His Trp Phe Lys
145                 150                 155                 160

Arg Phe Leu Asp Asp Gln Thr Ile Lys Lys Gln Gly Gly Asn Ser His
                165                 170                 175

Ser Ala Gly Ala Leu Thr Leu Arg Ser Gly Arg Ile Ser Trp Arg Leu
            180                 185                 190

Asp Ser Ser Lys Gly Asn Pro Trp Asp Arg Asn Arg Leu Val Leu Phe
        195                 200                 205

Cys Ser Val Asp Thr Leu Leu Trp Thr Lys Glu Gly Thr Glu Lys Ala
    210                 215                 220

Ser Gln Glu Lys Ala Ser Lys Ile Ala Gln Val Ile Ser Gly Thr Lys
225                 230                 235                 240

Ala Lys Gly Asn Leu Thr Ser Lys Gln Glu Asp Phe Val Arg Lys Arg
                245                 250                 255

Glu Lys Thr Leu Ala Leu Leu Gln Asn Pro Phe Pro Arg Pro Ser Arg
            260                 265                 270

Pro Leu Tyr Gln Gly Ser Pro Ala Ile Leu Ala Gly Val Ser Phe Gly
        275                 280                 285

Leu Asp Lys Pro Ala Thr Leu Ala Ile Val Asp Val Thr Thr Gly Lys
    290                 295                 300

Ala Ile Ala Tyr Arg Ser Ile Arg Gln Leu Leu Gly Asp Asp His Lys
305                 310                 315                 320

Leu Leu Asn Arg Gln Arg Gln Arg Gln Arg Gln Lys Ala Gln Arg Arg
                325                 330                 335

Arg Ser Asn Gln Leu Lys Phe Ala Ser Asn Arg Ile Ser Glu Gly Gly
```

340 345 350

Leu Gly Gly Gln Ile Asp Ser Leu Ile Ala Lys Ala Ile Val Gln Ile
355 360 365

Ala Gln Gln Tyr Asn Ala Ser Ser Ile Val Leu Gly Asp Leu Ala Asn
370 375 380

Ile Arg Glu Ile Ile Glu Ser Glu Ile Gln Ala Lys Ala Glu Gln Lys
385 390 395 400

Thr Thr Leu Lys Glu Ile Gln Ala Lys Tyr Ala Arg Asp Tyr Arg Ala
405 410 415

Ser Ile His Arg Trp Ser Tyr Lys Arg Leu Ala Gln Lys Ile Glu Ser
420 425 430

Asn Ala Leu Gln Ala Gly Leu Ile Val Ala Thr Ile Lys Gln Pro Leu
435 440 445

Ala Gly Ser Pro Gln Asp Lys Ala Arg Asp Val Ala Ile Ala Gly Phe
450 455 460

Gln Ser Arg Ser Val Ser Lys Ile Leu Asp Thr Gly Ser
465 470 475

<210> 83
<211> 439
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-47 effector

<400> 83
Leu Lys Asn Gly Gly Lys Ile Ser Asn Lys Glu Glu Asp Pro Glu Lys
1 5 10 15

Phe Ala Lys Arg Arg Arg Lys Leu Glu Ile Gln Ile Glu Arg Leu Arg
20 25 30

```
Glu Gln Leu Glu Ala Arg Ile Pro Lys Gly Arg Asp Leu Thr Asp Ala
        35                  40                  45

Lys Trp Leu Glu Thr Leu Leu Leu Ala Thr Asn Lys Ala Pro Glu Asn
    50                  55                  60

Glu Glu Glu Ala Lys Phe Trp Gln Asp Ser Leu Leu Lys Lys Ser Ser
65                  70                  75                  80

Thr Val Pro Phe Pro Val Ala Tyr Glu Thr Lys Glu Asp Met Thr Trp
                85                  90                  95

Phe Lys Asn Glu Arg Gly Arg Ile Cys Val Lys Phe Asn Gly Leu Ser
            100                 105                 110

Glu His Ser Phe Gln Val Tyr Cys Asp Thr Arg His Leu His Trp Phe
        115                 120                 125

Gln Arg Phe Leu Lys Asp Gln Gln Ile Lys His Asn Ser Lys Asn Gln
    130                 135                 140

His Ser Ser Ser Leu Phe Thr Leu Arg Ser Gly Arg Ile Val Trp Gln
145                 150                 155                 160

Glu Gly Glu Ser Lys Gly Asp Pro Trp Lys Val Asn His Leu Ile Leu
                165                 170                 175

Tyr Cys Ser Val Asp Thr Arg Leu Trp Thr Ala Glu Gly Thr Asn Leu
            180                 185                 190

Val Arg Glu Glu Lys Ala Glu Glu Ile Ala Lys Thr Ile Thr Gln Thr
        195                 200                 205

Lys Val Lys Gly Glu Leu Asn Asp Lys Gln Leu Ala His Ile Lys Arg
    210                 215                 220

Lys Asn Ser Ser Leu Ala Arg Ile Asn Asn Pro Phe Pro Arg Pro Ser
225                 230                 235                 240
```

```
Lys Pro Leu Tyr Lys Gly Gln Pro His Ile Leu Val Gly Val Ser Leu
                245                 250                 255

Gly Leu Glu Lys Pro Thr Thr Leu Ala Ile Val Asp Gly Thr Thr Ser
            260                 265                 270

Lys Val Ile Thr Tyr Arg Ser Ile Lys Gln Leu Leu Gly Asp Asn Tyr
        275                 280                 285

Lys Leu Leu Asn Arg Gln Arg Gln Gln Lys His Leu Leu Ser His Gln
    290                 295                 300

Arg Gln Ile Ala Gln Thr Leu Ala Ser Pro Asn Gln Leu Gly Glu Ser
305                 310                 315                 320

Glu Leu Gly Gln Tyr Val Asp Arg Leu Leu Ala Lys Glu Ile Val Ala
                325                 330                 335

Ile Ala Gln Thr Tyr Lys Ala Gly Ser Ile Val Leu Pro Lys Leu Gly
            340                 345                 350

Asp Met Arg Glu Gln Leu Gln Ser Glu Ile Gln Ala Lys Ala Glu Gln
        355                 360                 365

Lys Ser Asp Leu Leu Glu Val Gln Gln Lys Tyr Ala Lys Gln Tyr Arg
    370                 375                 380

Val Ser Val Tyr Gln Trp Ser Tyr Gly Arg Leu Ile Ala Asn Ile Arg
385                 390                 395                 400

Ser Ser Ala Ala Lys Ala Gly Ile Val Ile Glu Glu Ser Lys Gln Pro
                405                 410                 415

Ile Arg Gly Ser Pro Lys Glu Lys Ala Lys Glu Leu Ala Ile Val Ala
            420                 425                 430

Tyr Asn Ser Arg Gln Thr Thr
        435
```

<210> 84
<211> 641

<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-48 effector

<400> 84
Met Ser Val Ile Thr Ile Gln Cys Arg Leu Val Ala Lys Glu Phe Ser
1 5 10 15

Leu Arg Tyr Leu Trp Glu Leu Met Ala Gln Lys Asn Thr Pro Leu Val
20 25 30

Asn Glu Leu Leu Gln Ala Ile Val Gln His Ser Asp Phe Glu Ser Trp
35 40 45

Leu Gln Lys Gly Ser Ile Pro Gln Ala Thr Ile Lys Ala Leu Cys Asp
50 55 60

Ser Leu Lys Asn Asp Pro Arg Phe Ser Gly Gln Pro Gly Cys Phe Tyr
65 70 75 80

Thr Ser Ala Thr Ser Leu Val Ser Tyr Asn Tyr Lys Ser Trp Phe Ala
85 90 95

Leu Gln Lys Arg Arg Gln Gln Lys Ile Glu Gly Lys Glu Arg Trp Leu
100 105 110

Asn Leu Leu Lys Ser Glu Arg Glu Leu Glu Glu Glu Cys Asp Arg Thr
115 120 125

Leu Glu Glu Ile Arg Ala Lys Ala Thr Gln Ile Leu Thr Gln Phe Gln
130 135 140

Pro Gln Ala Glu Asn Ala Glu Thr Pro Pro Gly Glu Ala Thr Gln Lys
145 150 155 160

Lys Lys Gly Lys Lys Ser Lys Ser Thr Pro Ser Lys Gly Lys Ala Ala
165 170 175

```
Ser Leu Phe Asn Thr Leu Phe Asn Ala Tyr Ala Gln Thr Lys Asp Pro
            180                 185                 190

Leu Glu Arg Cys Ala Leu Ala Tyr Leu Leu Lys Asn Asn Gly Glu Val
        195                 200                 205

Ser Leu Val Glu Glu Asp Leu Glu Ala Tyr Gln Gln Arg Arg Arg Lys
    210                 215                 220

Lys Glu Ile Glu Ile Glu Arg Leu Lys Glu Gln Leu Lys Ser Arg His
225                 230                 235                 240

Pro Lys Gly Arg Asp Leu Thr Ala Glu Lys Trp Leu Glu Thr Leu Glu
                245                 250                 255

Ala Ala Arg Thr Gln Gln Pro Glu Ser Glu Ala Glu Ala Ala Thr Trp
            260                 265                 270

Gln Ala Asn Leu Leu Arg Asn Ser Ser Pro Val Pro Phe Pro Val Asn
        275                 280                 285

Tyr Glu Thr Asn Glu Asp Leu Thr Trp Phe Lys Asn Glu Ala Gly Arg
    290                 295                 300

Ile Cys Val Phe Phe Ser Gly Leu Arg Glu His His Phe Glu Val Trp
305                 310                 315                 320

Cys Asp Asn Arg Gln Leu His Trp Phe Glu Arg Phe Phe Gln Asp Gln
                325                 330                 335

Gln Ile Lys Arg Gln Ser Lys Asn Gln His Ser Ser Ser Leu Phe Thr
            340                 345                 350

Leu Arg Ser Gly Arg Ile Gly Trp Gln Glu Cys Thr Gly Glu Gly Glu
        355                 360                 365

Ala Trp Asn Ile His His Leu Lys Leu Tyr Cys Thr Leu Asp Thr Arg
    370                 375                 380

Leu Trp Thr Ala Glu Gly Thr Glu Gln Val Ala Gln Glu Lys Ala Ala
```

```
385                 390                 395                 400

Glu Ile Ser Asn Ile Leu Thr Lys Thr Lys Gln Lys Asp Asp Leu Asn
                405                 410                 415

Glu Lys Gln Leu Ala Phe Ile Lys Arg Gly Glu Thr Met Leu Ala Arg
            420                 425                 430

Ile Asn Thr Pro Phe Pro Arg Pro Ser Lys Pro Leu Tyr Gln Gly Gln
        435                 440                 445

Ser Glu Ile Val Val Gly Val Ser Leu Gly Leu Glu Lys Pro Ala Thr
    450                 455                 460

Val Ala Val Ile Asp Val Pro Gly Asp Arg Val Leu Ala Tyr Arg Asn
465                 470                 475                 480

Val Lys Gln Leu Leu Gly Asp His Tyr Pro Leu Leu Asn Arg Arg Arg
                485                 490                 495

Gln Gln Gln Arg Ser Leu Ser Gln Lys Arg Gln Lys Ala Gln Lys Gln
            500                 505                 510

Gly Ala Ala Asn Glu Phe Gly Glu Ser Glu Leu Gly Gln Tyr Ile Asp
        515                 520                 525

Arg Leu Leu Ala Lys Glu Ile Val Ala Ile Ala Arg Ser Tyr Gln Ala
    530                 535                 540

Gly Ser Ile Ala Leu Pro Lys Leu Thr Asp Ile Arg Glu Ile Leu His
545                 550                 555                 560

Ser Glu Met Gln Thr Arg Val Glu Lys Lys Ile Pro Gly Tyr Lys Glu
                565                 570                 575

Ala Gln Asp Lys Tyr Ala Lys Gln Tyr Arg Val Asn Val His Arg Trp
            580                 585                 590

Ser Tyr Gly Arg Leu Met Asp Ala Ile Ala Ser Gln Ala Ala Lys Val
        595                 600                 605
```

Gly Ile Val Ile Glu Val Gly Lys His Pro Ser Glu Gly Ser Phe Gln
610 615 620

Asp Lys Ala Arg Glu Val Ala Ile Ser Ala Tyr Cys Asp Arg Ser Val
625 630 635 640

Asn

<210> 85
<211> 709
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-49 effector

<400> 85
Met Ser Ile Ile Thr Ile Arg Cys Arg Leu Val Ala Gly Ile Lys Gln
1 5 10 15

Gln Ile Asp Lys Lys Ser Ile Lys Asn Phe Ser Ala Glu Asp Arg Ala
20 25 30

Leu Leu Gln Lys Leu Leu Asp Asp Lys Ser Gln Thr Asp Arg Asp Asn
35 40 45

Pro Lys Asp Lys Gln Lys Ile Phe Leu Ala Gln Ser Ser Glu Asp Val
50 55 60

Arg Gln Asn Leu Trp Gln Leu Phe Phe Thr Ser Ser Ala Leu Ile Asp
65 70 75 80

Gln Leu Leu Asp Arg Leu Ser Gln His Pro Asn Val Gln Thr Trp Gln
85 90 95

Gln Gln Gly Lys Leu Pro Asp Asp Glu Leu Lys Ala Cys Trp Leu Glu
100 105 110

```
Leu Lys Thr Ser Pro Leu Tyr Asp Glu Lys Leu Pro Gly Arg Phe Phe
        115                 120                 125

Ser Ser Val Gln Ser Met Val Lys Thr Ile Tyr Ala Ser Trp Leu Ala
    130                 135                 140

Leu His Gln Gln Lys Gln Arg Arg Leu Asn Gly Leu Asn Arg Leu Thr
145                 150                 155                 160

Glu Ile Val Tyr Ser Asp Glu Thr Leu Leu Glu Met Cys Asp Ser Thr
                165                 170                 175

Phe Ala Gln Leu Gln Leu Asn Ala Ala Ser Met Leu Ser Glu Ile Asp
            180                 185                 190

Lys Glu Val Val Asn Ser Glu Lys Ser Arg Ser Arg Ile Asp Leu Leu
        195                 200                 205

Phe Lys Lys Tyr Thr Glu Leu Pro Val Glu Asp Thr Ile Gly Arg Ser
    210                 215                 220

Ala Ile Ser Tyr Leu Ile Lys His Gly Cys Lys Ile Glu Ser Lys Ile
225                 230                 235                 240

Glu Leu Thr Pro Lys Phe Ala Arg Trp Phe Lys Thr Lys Leu Lys Gln
                245                 250                 255

Ala Gln Arg Leu Glu His Gln Leu Ala Gly His Phe Pro Arg Gly Arg
            260                 265                 270

Asp Leu Asn Gly Ile Ala Phe Leu Asn Ala Leu Glu Thr Ala Thr Lys
        275                 280                 285

Asp Glu Pro Gln Asp Asn Pro Glu Leu Met Leu Trp Gln Ser Gln Ile
    290                 295                 300

Leu Arg Asp Pro Ser Ser Leu Pro His Pro Ile Glu Phe Asn Ser Asn
305                 310                 315                 320

Thr Asp Leu Arg Trp Leu Lys Leu Tyr Arg Lys Gln Tyr Gln Cys Lys
                325                 330                 335
```

```
Arg Thr Val Ser Gly Asp Ser Ile Asp Ser Ile Glu Leu Thr Gln Arg
            340                 345                 350

Leu Phe Val Glu Phe Lys Gly Leu Thr Gln Gly Ser Asn Tyr Val Phe
        355                 360                 365

Glu Val Tyr Cys Asp Arg Arg Gln Leu Ala Ile Phe Gln Gln Phe Phe
    370                 375                 380

Asp Asp Asp Arg Leu Leu Arg Asn Ser Ser Ser Glu Glu Lys Tyr Ser
385                 390                 395                 400

Ser Ser Leu Phe Thr Leu Arg Ser Ala Tyr Leu Leu Trp Asp Arg Gln
                405                 410                 415

Glu Ser Gln Asp Arg Tyr Arg Arg His Thr Leu Ser Ile Gln Thr Ala
            420                 425                 430

Thr Glu Pro Trp Asn Thr Asn Gln Leu Tyr Leu His Cys Ser Ile Glu
        435                 440                 445

Thr Lys Ser Leu Thr Ala Glu Gly Met Arg Glu Ile Gln Gln Gln Lys
    450                 455                 460

Thr Gln Lys Val Asn Asn Ile Leu Ala Lys Gln Ser Lys Asn Ala Tyr
465                 470                 475                 480

Pro Ser Ile Asp Gln Gln Gln Ser Gln Arg Lys Asn Gln Thr Ser Leu
                485                 490                 495

Asn Leu Leu Asp Arg Ser Leu Pro Arg Pro Ser Arg Pro Ile Tyr Gln
            500                 505                 510

Gly Asn Pro Gln Ile Ile Val Gly Leu Ile Phe Asp Pro Val Arg Pro
        515                 520                 525

Ile Tyr Leu Ala Val Val Asp Val Thr Thr Gly Lys Pro Ile Thr Tyr
    530                 535                 540
```

```
Arg Ser Thr Arg Gln Leu Leu Gly Asp Lys Tyr Gln Lys Leu Ser Glu
545                 550                 555                 560

Tyr Arg Leu Lys Gln Gln Gln Asn Ser His His Arg Arg Lys His Asn
                565                 570                 575

Gln Gln Gly Gln Phe Gln Gln Pro Thr Glu Ser Thr Gln Gly Glu Tyr
            580                 585                 590

Leu Asn Arg Leu Leu Ala Lys Ala Val Ile Gln Ile Ala Gln Glu Phe
        595                 600                 605

Lys Ala Ala Ser Ile Ala Leu Pro Pro Ile Asn His Asn Ile Glu Lys
    610                 615                 620

Val Gln Ser Lys Ile Glu Ala Asp Ala Glu Glu Asp Leu Pro Glu Asp
625                 630                 635                 640

Val Gly Thr Gln Lys Lys Ile Thr Arg Lys Thr Ser Val Val Ile His
                645                 650                 655

Lys Trp Ser Tyr His Ser Leu Leu Glu Tyr Val Lys Ser Asn Ala Ala
            660                 665                 670

Lys Leu Gly Ile Thr Val Glu Thr Val Ser Leu Pro Ser Leu Gly Thr
        675                 680                 685

Pro Ser Gln Gln Ala Ala Glu Val Ala Ile Ser Ala Tyr Asn Ser Arg
    690                 695                 700

Lys His Ile Lys Lys
705
```

<210> 86
<211> 642
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>

<223> MG64-50 effector

<400> 86

```
Met Arg Tyr Ala Ile Ser Met Ser Gln Ile Thr Ile Gln Cys Gln Leu
1               5                   10                  15


Phe Ala Asp Glu Ala Ala Arg Lys Tyr Leu Trp Gln Leu Met Thr Ala
            20                  25                  30


Arg Asn Thr Pro Leu Val Asn Glu Leu Leu Arg Leu Ile Ala Leu His
        35                  40                  45


Pro Asp Phe Leu Lys Trp Arg Ser Ser Gly Lys Leu Pro Thr Thr Glu
    50                  55                  60


Val Thr Lys Leu Ala Lys Thr Leu Lys Thr Asp Glu Arg Phe Asn Asn
65                  70                  75                  80


Gln Pro Ala Lys Phe His Ile Ser Ala Glu Lys Thr Ala Ile Tyr Thr
                85                  90                  95


Phe Lys Ser Trp Leu Ala Ile Gln Lys Arg Ala Gln Gln Gln Leu Glu
            100                 105                 110


Gly Lys Ile Ser Trp Leu Arg Met Leu Arg Thr Asn Glu Glu Leu Val
        115                 120                 125


Ala Asp Cys Gly Gln Asp Leu Ala Gln Ile Arg Thr Lys Ala Gln Ala
    130                 135                 140


Leu Ile Leu Gln Tyr Gln Ser Ser Glu Asp Ser Ala Glu Ser Gln Lys
145                 150                 155                 160


Asn Phe Arg Lys Ser Leu Tyr Gln Ala Tyr Asp Ile Ala Glu Pro Arg
                165                 170                 175


Leu Gly Ala Gln Asn Glu Arg Ser Leu Thr Arg Ser Ala Ile Ala Tyr
            180                 185                 190


Leu Leu Lys Asn Arg Cys His Ile Pro Ser Asp Thr Asp Glu Asp Leu
        195                 200                 205
```

```
Lys Lys Phe Leu Lys Tyr Arg Arg Lys Ile Glu Asn Gln Val Lys Arg
    210                 215                 220

Leu Thr Gln Gln Leu Glu Asn Arg Leu Pro Gln Gly Arg Asp Leu Thr
225                 230                 235                 240

Gly Glu Arg Phe Leu Lys Thr Leu His Ser Ala Thr Lys Ser Val Ser
                245                 250                 255

Val Asp Asn Thr Glu Ala Ser Arg Trp Gln Ser Gln Leu Leu Glu Arg
            260                 265                 270

Pro Asp Leu Val Pro Phe Pro Ile Thr Leu Glu Ser Asn Met Asp Leu
        275                 280                 285

Met Trp Phe Ile Thr Gln Gln Asp Lys Ile Gly Leu His Ile Gly Gly
    290                 295                 300

Ile Ser Glu His Glu Phe Thr Ile Gly Cys Gly Gln Arg Gln Leu His
305                 310                 315                 320

Tyr Phe Gln Arg Phe Leu Ser Asp Tyr Gln Thr Met Leu Ala Ser Lys
                325                 330                 335

Arg Gln His Thr Ser Ser Leu Phe Leu Leu Arg Ser Ala Lys Ser Ile
            340                 345                 350

Trp Ile Pro Ser Lys Asp Lys Gly Glu Pro Trp Asn Val His Gln Leu
        355                 360                 365

Tyr Leu Ser Cys Thr Leu Asp Thr Arg Leu Leu Thr Ser Glu Gly Thr
    370                 375                 380

Glu Leu Val Lys Gln Glu Val Ala Ala Gly Thr Thr Lys Lys Leu Val
385                 390                 395                 400

Thr Met Gln Glu Lys Leu Asp Arg Thr Asp Asn Gln Asp Asn Tyr Val
                405                 410                 415

Lys Arg Leu Gln Ser Thr Leu Asp Arg Leu Asp Arg Pro Phe Asn Arg
```

```
            420                 425                 430

Pro Ser Lys Pro Leu Tyr Gln Gly Gln Ser Asn Ile Ile Val Ala Val
        435                 440                 445

Ser Met Gly Leu Gln Ser Pro Val Thr Ala Ile Ala Ile Asp Ile Thr
    450                 455                 460

Thr Gln Lys Ile Leu Ala Tyr Arg His Thr Lys Gln Leu Leu Gly Asp
465                 470                 475                 480

Asp Tyr Arg Leu Val Asn Arg Gln Arg Asn Leu Gln Thr Gln Gln Arg
                485                 490                 495

His Ser Ser His Lys Ala Gln Lys Gln Gly Leu Ser Arg Gln Cys Ser
            500                 505                 510

Asn Ser Glu Leu Gly Glu His Leu Asp Arg Leu Phe Ala Lys Ala Ile
        515                 520                 525

Val Glu Leu Ala Gln Thr Tyr Gln Ala Gly Ser Ile Ala Leu Pro Lys
    530                 535                 540

Leu Asp Gln Ile Arg Leu Ser Ile Gln Ser Glu Ile Asp Thr Lys Ala
545                 550                 555                 560

Gln Gln Lys Ile Pro Gly Tyr Val Glu Gly Gln Lys Lys Tyr Ala Lys
                565                 570                 575

Gln Ile Arg Ile Asn Leu His Asn Trp Ser Tyr Asn Arg Val Ser Gln
            580                 585                 590

Leu Ile Ile Asn Lys Ala Gly Gln Ser Gly Ile Ala Ile Glu Tyr Gly
        595                 600                 605

Asp Gln Pro Ala Arg Ala Ser Pro Asn Val Arg Glu Ala Ser Arg Asn
    610                 615                 620

Glu Asn Arg Ala Gly Glu Ile Ala Leu Ser Ala Tyr Ala Lys Arg Thr
625                 630                 635                 640
```

Ile Pro

<210> 87
<211> 467
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-51 effector

<400> 87
Met Phe Pro Ser Phe Cys Ser Pro Pro Phe Leu Ser Tyr Phe Ser Thr
1 5 10 15

Pro Leu Asn Gly Tyr Asn Arg Cys Ala Ile Ser Tyr Leu Leu Lys Asn
20 25 30

Gly Cys Lys Ile Asn Asp Lys Glu Glu Asn Pro Glu Lys Phe Ala Gln
35 40 45

Arg Arg Arg Lys Leu Glu Ile Gln Ile Gln Arg Leu Thr Glu Lys Leu
50 55 60

Ala Ala Arg Ile Pro Lys Gly Arg Asp Leu Thr Asp Thr Gln Trp Leu
65 70 75 80

Glu Ala Leu Phe Thr Ala Thr Gln Thr Val Pro Glu Asn Glu Thr Glu
85 90 95

Ala Lys Phe Trp Gln Asn Ser Leu Leu Lys Lys Ser Ser Ser Val Pro
100 105 110

Phe Pro Val Ala Tyr Glu Thr Asn Glu Asp Leu Val Trp Val Lys Asn
115 120 125

Gln Phe Gly Arg Ile Cys Ile Lys Phe Ser Gly Leu Ser Glu His Thr
130 135 140

```
Phe Gln Ile Tyr Cys Asp Ser Arg Gln Leu His Trp Phe Gln Arg Leu
145                 150                 155                 160

Leu Glu Asp Gln Lys Ile Lys Lys Asn Ser Lys Asn Gln His Ser Ser
                165                 170                 175

Ala Leu Phe Thr Leu Arg Ser Gly Arg Ile Ser Trp Gln Glu Gly Glu
            180                 185                 190

Asp Lys Gly Asp Pro Trp Asn Val His His Leu Thr Leu Tyr Cys Ser
        195                 200                 205

Val Asp Thr Arg Leu Trp Thr Glu Glu Gly Thr Asn Leu Val Lys Glu
    210                 215                 220

Glu Lys Ala Glu Glu Ile Ala Lys Thr Ile Thr Gln Thr Gln Ala Lys
225                 230                 235                 240

Gly Asp Leu Asn Asp Asn Gln Gln Ala His Ile Lys Arg Lys Asn Ser
                245                 250                 255

Ser Leu Ala Arg Ile Asn Asn Pro Phe Pro Arg Pro Ser Gln Pro Leu
            260                 265                 270

Tyr Gln Gly Gln Ser His Ile Leu Val Gly Val Ser Leu Gly Leu Glu
        275                 280                 285

Asn Pro Ala Thr Ile Ala Ile Val Asp Gly Thr Thr Ser Lys Val Leu
    290                 295                 300

Thr Tyr Arg Asn Ile Lys Gln Ile Leu Gly Asp Asn Tyr Lys Leu Leu
305                 310                 315                 320

Asn Lys Gln Arg Gln Gln Lys His Leu Leu Ser His Gln Arg His Ile
                325                 330                 335

Ala Gln Arg Met Ala Ala Ser Asn Glu Phe Gly Asp Ser Glu Leu Gly
            340                 345                 350

Glu Tyr Ile Asp Arg Leu Leu Ala Lys Glu Ile Ile Ala Ile Ala Gln
        355                 360                 365
```

```
Thr Tyr Lys Ala Gly Ser Ile Val Leu Pro Lys Leu Gly Asp Met Arg
    370                 375                 380

Glu Gln Val Gln Ser Glu Ile Lys Ala Lys Ala Glu Gln Lys Ser Asp
385                 390                 395                 400

Leu Val Glu Val Gln Gln Lys Tyr Ala Lys Gln Tyr Arg Val Ser Val
                405                 410                 415

His Gln Trp Ser Tyr Gly Arg Leu Ile Ala Asn Ile Gln Cys Gln Ala
            420                 425                 430

Asn Lys Ala Gly Ile Val Ile Glu Glu Ala Lys Gln Pro Ile Arg Gly
        435                 440                 445

Ser Pro Gln Glu Lys Ala Lys Glu Leu Ala Ile Asn Ala Tyr Gln Ser
    450                 455                 460

Arg Lys Ala
465
```

<210> 88
<211> 814
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-52 effector

<400> 88

```
Met Ala Ile Arg Thr Met Glu Thr Asp Val Ile Val Arg Thr Ser Lys
1               5                   10                  15

Pro Arg Glu Ser Lys Arg Ser Ile Asp Asp Leu Val Leu Phe Asp Glu
            20                  25                  30

Ala Ala Trp Cys Arg Val Cys Glu Phe Gly Tyr Gln His Thr Leu Phe
        35                  40                  45
```

```
Val Asp Thr Ile Met Lys Gln Val Lys Gln Leu Pro Ile Asn Trp Leu
    50                  55                  60

Tyr Thr Asn Gln Glu Leu Gln Lys Gln Asp Asp Thr Thr His Glu Thr
65                  70                  75                  80

Glu Thr Lys Gln Gly Gln Asn Gly Leu Pro Glu Gly Leu Val Lys Ala
                85                  90                  95

Leu His Asp Ala Leu Ala Ser Thr Pro Gln Leu Ser Lys Met Ser Gly
            100                 105                 110

Arg Phe Tyr Ala Ser Ala Ile Asp Arg Val Glu Glu Ile Ile Glu Ser
        115                 120                 125

Trp Phe Ala Ala His Gln Lys Leu Ile His Lys Leu Arg Gly Gln Arg
    130                 135                 140

Arg Trp Leu Ala Ala Val Glu Ser Asp Thr Ala Leu Ala Glu Ser Ser
145                 150                 155                 160

His Phe Ser Gln Tyr Glu Ile Glu Gln Arg Ala Ala Gln Ile Leu Ala
                165                 170                 175

Glu Leu Glu Ala Lys Gly Glu Leu Thr Ala Ala Ser Gly Asn Gln Thr
            180                 185                 190

Pro Phe Asn Val Leu Phe Lys Gln Phe Asp Glu Thr Glu Ala Val Leu
        195                 200                 205

Ser Arg Arg Ala Ile Ile His Leu Leu Lys Asn Ile Asn Glu Val Gln
    210                 215                 220

Val Lys Leu Lys Gln Pro Lys Lys Gly Lys Gly Arg Asn Lys Ala Ala
225                 230                 235                 240

Gln Arg Ser Leu Thr Glu Arg Leu Glu Ala Lys Arg Val Glu Ile Glu
                245                 250                 255

Arg Leu Glu Lys Gln Leu Leu Gly Gln Leu Pro Arg Ala Arg Asn Leu
```

```
            260                 265                 270

Phe Pro Asp Gln Ala Phe Asp Gln Arg Leu Glu Ala Leu Ile Asp Met
        275                 280                 285

Pro Leu Ser Glu Ala Thr Glu Leu Glu Arg Tyr Tyr Phe Leu Tyr Phe
    290                 295                 300

Ser Leu Leu Leu Tyr Leu Ser Gly Ser Asn Thr Tyr Val Gln Leu Glu
305                 310                 315                 320

Gln His Leu Leu Ile Ala Leu Val Leu Gln Trp Ser Lys Leu Asp Asp
                325                 330                 335

Leu His Tyr Tyr Arg Ala Leu Asn Tyr Ser Phe Ile Leu Tyr Ala Ala
            340                 345                 350

Ser Glu Gln Gln Tyr Leu Gln Leu Gly Ser Tyr Leu Leu Gln Ser Ile
        355                 360                 365

Lys Ile Glu Ala Glu Arg Val Glu Ala Ala Phe Phe Ala Trp His Glu
    370                 375                 380

Ser Ile Thr Pro Lys Leu Gln Asp Phe Leu Arg Glu Pro Lys Ala Leu
385                 390                 395                 400

Pro Tyr Pro Ile Ser Phe Gly Tyr Glu Asp Val Arg Ser Trp His Val
                405                 410                 415

Asn Gln Lys Gly Lys Ile Phe Phe Lys Leu Asn Gly Trp Gly Asp Leu
            420                 425                 430

Leu Phe Glu Val Arg Cys His Arg Arg Gln Leu Ser Leu Ile Lys Thr
        435                 440                 445

Phe Cys Lys Asp Trp Gln Thr Lys Asn Asn Ala Glu His Lys Gly Gln
    450                 455                 460

Tyr Ser Gly Ser Leu Met Leu Leu Arg Ser Ile Glu Leu Ile Trp Lys
465                 470                 475                 480
```

```
Pro Lys Lys Ala Ser Glu Gln Lys Asp Ala Gln Leu Cys Ser Gln Cys
               485                 490                 495

Glu Val Phe Gln Gln Tyr Gly Gly Lys Gly Phe Trp Asn Glu Cys Lys
            500                 505                 510

Leu Thr Ile His Trp Thr Phe Asp Thr Asp Gly Leu Thr Lys Gln Gly
        515                 520                 525

Ser Glu Gln Ile Arg Gln Arg Lys Leu Lys Pro Gln Leu Lys Lys Leu
    530                 535                 540

Gln Asp Asp Gln Ala Ala Leu Glu Gln Lys Gln Asp Arg Leu Asn Lys
545                 550                 555                 560

Leu Glu Gln Gln Val Ala Glu Pro Ser Gln Ser Gln Ala Gln Leu Lys
               565                 570                 575

Arg Leu Glu Pro Leu Lys Lys Glu Val Gln Lys Leu Gln Glu Lys Leu
            580                 585                 590

Ala Lys Pro Arg Pro Lys Leu Asp Cys Leu Gln Ala Ala Lys Pro Phe
        595                 600                 605

Asp Arg Pro Asp Arg Pro Leu Tyr Glu Gly Ile Pro Asn Ile Phe Val
    610                 615                 620

Gly Val Leu Leu Asp Leu Asp Lys His Leu Val Val Thr Val Val Asp
625                 630                 635                 640

Ala Met Arg Arg Lys Arg Leu Thr Leu Arg Ser Ala Arg Lys Val Ser
               645                 650                 655

Pro Glu Gly Tyr Arg Leu Leu Gln Lys Tyr Phe Arg Gln Arg Gln Glu
            660                 665                 670

His Ser Gln Gln Arg Gln Gln Asp Gln Lys Ala Gln Arg Arg Ser His
        675                 680                 685

Gln Thr Glu Ser Gly Leu Gly Gln Gln Val Ala Arg Leu Phe Ala Lys
```

690 695 700

Gly Leu Val Glu Leu Ala Gln Ala Tyr Lys Ala Ser Thr Ile Val Ile
705 710 715 720

Pro Ile Thr Gly Gly Trp Arg Glu Arg Leu Tyr Ser Gln Leu Val Ala
725 730 735

Arg Ala Lys Ile Arg Cys Asn Gly Ser Lys Gln Ala Val Asp Arg Tyr
740 745 750

Thr Lys Glu His Gly Glu Arg Leu His Gln Trp Asp Tyr Asn Arg Leu
755 760 765

Ser Gln Ala Ile Thr Asp Cys Ala Ala Thr His Gly Ile Glu Ile Val
770 775 780

Leu Gln Lys Thr Val Phe Glu Ala Asp Val Phe Gln Gln Ser Ala Asn
785 790 795 800

Leu Ala Ile Ala Ala Tyr Asp Ser Leu Asn Ser Glu Pro Pro
805 810

<210> 89
<211> 637
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-53 effector

<400> 89
Ser Met Ser Gln Ile Thr Ile Gln Cys Gln Leu Phe Ala Asp Glu Ala
1 5 10 15

Ala Arg Lys Tyr Leu Trp Gln Leu Met Thr Asp Arg Asn Thr Pro Leu
20 25 30

Val Asn Glu Leu Leu Arg Leu Ile Ala Leu His Pro Asp Phe Leu Lys
35 40 45

```
Trp Arg Ser Ser Gly Lys Leu Pro Thr Thr Glu Val Thr Lys Leu Ala
    50                  55                  60

Lys Thr Leu Lys Thr Asp Glu Arg Phe Asn Asn Gln Pro Ala Lys Phe
65                  70                  75                  80

His Ile Ser Ala Glu Lys Thr Ala Ile Tyr Thr Phe Lys Ser Trp Leu
                85                  90                  95

Ala Ile Gln Lys Arg Ala Gln Gln Gln Leu Glu Gly Lys Ile Ser Trp
            100                 105                 110

Leu Arg Met Leu Arg Thr Asn Glu Glu Leu Val Ala Asp Cys Gly Gln
        115                 120                 125

Asp Leu Ala Gln Ile Arg Thr Lys Ala Gln Ala Leu Ile Leu Gln Tyr
    130                 135                 140

Gln Ser Ser Glu Asp Ser Ala Glu Ser Gln Lys Asn Phe Arg Lys Ser
145                 150                 155                 160

Leu Tyr Gln Ala Tyr Asp Ile Ala Glu Pro Arg Leu Gly Ala Gln Asn
                165                 170                 175

Glu Arg Ser Leu Thr Arg Ser Ala Ile Ala Tyr Leu Leu Lys Asn Arg
            180                 185                 190

Cys His Ile Pro Ser Asp Thr Asp Glu Asp Leu Lys Lys Phe Leu Lys
        195                 200                 205

Tyr Arg Arg Lys Ile Glu Asn Gln Val Lys Arg Leu Thr Gln Gln Leu
    210                 215                 220

Glu Asn Arg Leu Pro Gln Gly Arg Asp Leu Thr Gly Lys Arg Phe Leu
225                 230                 235                 240

Lys Thr Leu Asp Ser Ala Thr Lys Ser Val Pro Val Asp Asn Thr Glu
                245                 250                 255
```

```
Ala Ser Arg Trp Gln Ser Gln Leu Leu Glu Arg Pro Asp Leu Val Pro
            260                 265                 270

Phe Pro Ile Thr Leu Glu Ser Asn Met Asp Leu Met Trp Phe Ile Ala
        275                 280                 285

Gln Gln Asp Lys Ile Gly Leu His Ile Gly Gly Ile Ser Glu His Glu
    290                 295                 300

Phe Thr Ile Gly Cys Gly Gln Arg Gln Leu His Tyr Phe Gln Arg Phe
305                 310                 315                 320

Leu Ser Asp Tyr Gln Thr Met Leu Ala Ser Lys Arg Gln His Thr Ser
                325                 330                 335

Ser Leu Phe Leu Leu Arg Ser Ala Lys Leu Ile Trp Ile Pro Ser Lys
            340                 345                 350

Asp Lys Gly Glu Pro Trp Asn Val His Gln Leu Tyr Leu Ser Cys Thr
        355                 360                 365

Leu Asp Thr Arg Leu Leu Thr Ser Glu Gly Thr Glu Leu Val Lys Gln
    370                 375                 380

Glu Val Ala Ala Gly Thr Thr Lys Lys Leu Val Thr Met Gln Glu Lys
385                 390                 395                 400

Leu Asp Arg Thr Asp Thr Gln Asp Asn Tyr Val Lys Arg Leu Gln Ser
                405                 410                 415

Thr Leu Asp Arg Leu Asp Arg Pro Phe Asn Arg Pro Ser Lys Pro Ser
            420                 425                 430

Tyr Gln Gly Gln Ser Asn Ile Ile Val Ala Val Ser Met Gly Leu Gln
        435                 440                 445

Ser Pro Val Thr Ala Ile Ala Ile Asp Ile Thr Thr Gln Lys Ile Leu
    450                 455                 460

Ala Tyr Arg Asn Thr Lys Gln Leu Leu Gly Asp Asp Tyr Arg Leu Val
465                 470                 475                 480
```

Asn Arg Gln Arg Asn Leu Gln Thr Gln Gln Arg His Leu Ser His Lys
485 490 495

Ala Gln Lys Gln Gly Leu Ser Arg Gln Cys Ser Asn Ser Glu Leu Gly
500 505 510

Glu His Leu Asp Arg Leu Phe Ala Lys Ala Ile Val Glu Leu Ala Gln
515 520 525

Thr Tyr Gln Ala Gly Ser Ile Ala Leu Pro Lys Leu Asp Gln Ile Arg
530 535 540

Leu Ser Ile Gln Ser Glu Ile Asp Thr Lys Ala Gln Gln Lys Ile Pro
545 550 555 560

Gly Tyr Val Glu Gly Gln Lys Lys Tyr Ala Lys Gln Ile Arg Ile Asn
565 570 575

Leu His Asn Trp Ser Tyr Asn Arg Val Ser Gln Leu Ile Ile Asn Lys
580 585 590

Ala Gly Gln Ser Gly Ile Ala Ile Glu Tyr Gly Asp Gln Pro Ala Arg
595 600 605

Ala Ser Pro Asn Val Arg Glu Ala Ser Arg Asn Glu Asn Arg Ala Arg
610 615 620

Glu Ile Ala Ile Ser Ala Tyr Ala Lys Arg Thr Ile Pro
625 630 635

<210> 90
<211> 376
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<220>
<223> MG64-2 effector intergenic region encoding tracrRNA

<400> 90
cacacaacag aaccttgata atagaatagt aaataattaa tagcgccgcc gttcatgctt 60

ctaggagcct ctgaaaggtg acaaatgcgg gttagtttgg ctgttgtcag acagtcttgc 120

tttctgaccc tggtagctgc ccaccccgaa gctgctgttc cttgtgaaca ggaattaggt 180

gcgcccccag taataagggt atgggtttac cacagtggtg gctactgaat cacctccgag 240

caaggaggaa cccactttaa ttttttttcg taaagccaag cgggagccaa aaccctaggg 300

ggtttacgaa agtctcacga ttcttacatt gagtaagttt cagtgttttt gggatggtta 360

acccttttatt tacaag 376

<210> 91
<211> 390
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-4 effector intergenic region encoding tracrRNA

<400> 91
tatacaaacg tttgaacctg gaaaatagaa taatgagaaa tagcgccgct tgttcatgcg 60

cgaaagcgtc tctgaacagt gtaaatgtgg gttagtttga ctgtcgtgaa gacggtcttg 120

ctttctgacc ctggtagctg cccaccttga agctgctgtc tcttgtagac aggaatcagg 180

tgcgccccca gtaatatagg tgcgggttta ccgcagtggt ggctacccaa tcacctccga 240

gcaaggagga acccacctta attatttttt ggcaaaccaa agtgggagca atttcactgg 300

gaggttcgcc aagttttcaa acaacttatt ttgtagaggt tttgccgttt tctgctaaga 360

aaagattttc ttgtccagat gcagtgataa 390

<210> 92
<211> 462
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>

<223> MG64-6 effector intergenic region encoding tracrRNA

<400> 92
gatttatcag taccgtacct ggaaaattga atataattga taacagcgcc gcaggtcatg 60

ccgtcaaaag cctctgaact gtgttaaatg ggggttagtt tgactgttga aagacagttg 120

tgctttctga ccctggtagc tgcccaccct gatgctgcta tctttcggga taggaataag 180

gtgcgctccc agtaataggg gtgtagatgt actacagtgg tggctactaa atcacctccg 240

accaaggagg aatccatcct taatttttta ttttttcgtg aacctaagcg agagcaaaat 300

ctctaggagg ttcacgaaaa agctgaatcc ttcatggaat atgagtttca gtttgttagt 360

gggatgatgg cttctcctga aaaggagaaa tggagtaaga attttagagg tttacgaaaa 420

tgacctttaa aagctacttc aggcaagtgt tacagcgctc gc 462

<210> 93
<211> 456
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-7 effector intergenic region encoding tracrRNA

<400> 93
tcaaaaagac cgtaccttga aaatataatg gtaacaaaca gcgccgcagt tcatgcgtct 60

tatggcgcct ctgtgctgtg caaaatgtgg gttagtttga ctgttggaag acagtcttgc 120

tttctgaccc tggtagctgc ccaccttgaa gctgctatcc cttgtggata ggaatcaggt 180

gcgcccccag taatagaggt gcgggtttac cgcagtggtg gctaccgaat cacctccgag 240

caaggaggaa cccaccttaa ttatttttt ggcatggcaa agcgggagcg attttaccgg 300

gactgatgcc aaagcttcaa atctttttat tgacaaggtt tctagacttt tgtttgtcac 360

ttgatttatt tttttaattg tcaactagca agtgattttg gtagttttgc caaaagtgtc 420

tctaggaatc ttgataaata aagggtttta ggcgcg 456

<210> 94
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-2 effector target CRISPR repeat

<400> 94
gtttcaacga ccatcccaac taggggtggg ttgaaag 37

<210> 95
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-4 effector target CRISPR repeat

<400> 95
gtttcaactt tccttccagc tagaggcggg ttgaaag 37

<210> 96
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-6 effector target CRISPR repeat

<400> 96
gtttcaacca ccatctcaac tagggatggg ttgaaag 37

<210> 97
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>

<223> MG64-7 effector target CRISPR repeat

<400> 97
gtttcaacgc cccttcaagc tttgggcggg ttgaaagc 38

<210> 98
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG108-1 effector target CRISPR repeat

<400> 98
gttgcgatcg ccgctccggt ggcgatgggg ttgaaag 37

<210> 99
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG110-1-6 effector target CRISPR repeat

<400> 99
gtgacctgcc gcataggcag ctggtaaa 28

<210> 100
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG110-2-6 effector target CRISPR repeat

<400> 100
tatgaactgc cgcataggca gccaaga 27

<210> 101
<211> 922
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-13-B transposition protein

<400> 101
Met Leu Ser Ser Gln Ala Phe Asn Asp Trp Cys Gln Arg Leu Gln Ile
1 5 10 15

Asn Ser Val Ala Arg Ala Val Ile Glu Gln Val Arg Asn Ala Ala Pro
20 25 30

Ser Arg Gln Val Gln Gly Arg Arg Lys Asn Val Cys Gly Ser Tyr Pro
35 40 45

Ser Arg Lys Met Gly Val Thr Ile Gln Phe Glu Ser His Arg Asn Glu
50 55 60

Leu Ala Arg Ile His Glu Leu Glu His Asp Gln Thr Val Leu Glu Tyr
65 70 75 80

Tyr Asp Gln Pro Pro Pro Ile Glu Leu Val Tyr Ser Ser Lys Thr Gly
85 90 95

Arg Arg Thr Arg His Gln Tyr Thr Pro Asp Phe Phe Val Leu Arg Thr
100 105 110

Asp Gly Val Glu Trp Glu Glu Cys Lys Thr Glu Thr Glu Leu Ile Lys
115 120 125

Leu Ala Gln Glu Asn Ser Asn Arg Tyr Cys Gln Asp Val Glu Gly Lys
130 135 140

Trp His Cys Pro Pro Gly Glu Ala Tyr Ala Gln Val Phe Gly Phe Gln
145 150 155 160

Phe Arg Val Trp Cys Asn Ser Glu Ile Asn Trp Val Leu Gln Asp Asn

```
                165                 170                 175

Trp Val Trp Leu Glu Asp Tyr Leu Gly Val Glu Val Pro Ser Val Ser
            180                 185                 190

Pro Ala Ala Leu Gln Ala Leu Gln Arg His Leu Gln Gln Thr Pro Gly
        195                 200                 205

Ile Cys Val Ser Glu Leu Leu Ala Val Glu Pro Thr Val Ser Pro Asp
    210                 215                 220

His Leu Tyr His Leu Ile Ala Thr Glu Gln Val Tyr Val Glu Leu Ala
225                 230                 235                 240

Ala Ala Arg Leu Ser Glu Pro Glu Gln Val Arg Val Phe Leu Asp Gln
                245                 250                 255

Asp Met Ala Gln Ala Phe His Arg Ile Gln Gly Glu Phe Glu Leu Val
            260                 265                 270

Gln Thr Asn Lys Gln Ser Met Leu Asn Phe Ser Ile Gly Thr Phe Leu
        275                 280                 285

His Trp Asp Glu Val Glu Trp Glu Val Ile Asn Ile Gly Ala Asn Thr
    290                 295                 300

Val Gly Leu Leu Arg Ser Glu Gly Glu Val Ile Glu Leu Ala Lys Pro
305                 310                 315                 320

Thr Leu Glu Asn Leu Val Ser Arg Gln Arg Ile Val Ala Ser Tyr Asn
                325                 330                 335

Ser Glu Asn Ile Glu Lys Gln Lys Ile Ile Asn Asp Lys Ile Lys Ala
            340                 345                 350

Ile Leu Gln Glu Ala Thr Ser Ala Glu Ile Ala Glu Ala Asn Arg Arg
        355                 360                 365

Tyr Asp Leu Ile Gln Ala Cys Ser Asn Thr Ser Arg His Pro Ser Ser
    370                 375                 380
```

```
Thr Val Arg Arg Trp Gln Asn Gln Tyr Gln Lys Ala Glu Lys Ile Tyr
385                 390                 395                 400

Gly Asn Gly Tyr Val Gly Leu Leu Pro Arg Tyr His Thr Lys Gly Asn
                405                 410                 415

Arg Thr Met Arg Ile Asp Asp Ile Ala Tyr Gln Phe Met Ile Glu Phe
            420                 425                 430

Ile Glu Gln His Tyr Glu Asn Pro Lys Gln Arg Arg Lys Gln Arg Val
        435                 440                 445

Tyr Glu Ala Phe Val Ser Ala Cys Glu Ala His Glu Pro Lys Leu Ile
    450                 455                 460

Pro Pro Ser Arg Ile Thr Phe Ser Gln Ala Ile Asp Lys Arg Gly Gly
465                 470                 475                 480

Gln Ala Gln Thr Leu Lys Arg Glu Gly Arg Arg Ala Ala Ile Ala Lys
                485                 490                 495

Glu Ala Phe Tyr Trp Gln Leu Glu Arg Thr Thr Pro Gln His Gly Ser
            500                 505                 510

Arg Pro Phe Glu Ile Val His Leu Asp His Thr Gln Leu Asp Ile Glu
        515                 520                 525

Leu Met Ser Ser Leu Ser Thr Leu Ala Val Gln Pro Leu Ala Ser Asn
    530                 535                 540

His Arg Thr Leu Ser Gln Asn Leu Gly Arg Pro Trp Ala Thr Phe Met
545                 550                 555                 560

Val Asp Ala Tyr Ser Arg Arg Leu Leu Ala Val Tyr Leu Thr Tyr Glu
                565                 570                 575

Glu Pro Ser Tyr Arg Ser Cys Met Met Ala Leu Arg Ile Cys Val Gln
            580                 585                 590

Arg Phe Glu Arg Leu Pro Gln Ile Leu Val Val Asp Asn Gly Val Glu
```

```
        595                 600                 605

Phe His Ser His Tyr Phe Glu Glu Leu Leu Ala Tyr Tyr Ala Cys Thr
    610                 615                 620

Lys Lys His Arg Pro Pro Ala Met Ala Arg Phe Gly Ser Val Val Glu
625                 630                 635                 640

Arg Leu Phe Gly Thr Ala Asn Thr Gln Phe Val His Glu Leu Gln Gly
                645                 650                 655

Asn Thr Gln Ile Thr Arg Gln Val Arg Gln Val Thr Gln Ser Ile Asn
            660                 665                 670

Pro Lys Arg Leu Ala Ile Trp Thr Leu Ser Asp Leu Tyr Pro Ala Leu
        675                 680                 685

Cys Glu Trp Ala Tyr Glu Ile Tyr Asp Gln Arg Glu His Pro Ala Leu
    690                 695                 700

Gly Gln Ser Pro His Asp Ile Phe Thr Gln Gly Leu Ala Leu Gly Gly
705                 710                 715                 720

Thr Arg Asp His Arg Arg Val Glu Asn Asn Glu Val Phe Gln Ile Leu
                725                 730                 735

Thr Leu Pro Ala Pro Asp Gln Arg Arg Arg Lys Val Gln Ala Gly Arg
            740                 745                 750

Gly Val Lys Ile His Asn Ile Tyr Tyr Trp Ser Asn Ala Phe Arg Asp
        755                 760                 765

Pro Lys Val Glu Gly Ser Gln Val Glu Val Lys Tyr Asp Pro Phe Asp
    770                 775                 780

Ala Ser Ile Ala Tyr Ala Phe Val His His Gln Trp Val Lys Cys Ile
785                 790                 795                 800

Ser Thr Tyr Cys Ala Asp Leu Gln Gly Arg Ser Glu Arg Glu Leu Arg
                805                 810                 815
```

Leu Val Ser Glu Glu Leu Arg Gln Arg Gln Arg Val Ala Gly Cys Arg
820 825 830

His Ser Leu Ser Asp Lys Glu Leu Val Glu Phe Leu Asn Ser Gln His
835 840 845

Ala Lys Glu Gly Gln Phe Leu Gln Arg Arg Leu Gln Ala Val Glu Gln
850 855 860

Leu Ala Val Leu Gln Gln Leu Asp Pro Pro Val Asp Asp Ala Asp Ala
865 870 875 880

Glu Ser Gln Gln Ile Glu Phe Gly Gly Asp Ala Gln Pro Ser Cys Asn
885 890 895

Gly Asn Ser Pro Gln Thr Ala Gln Thr Val Pro Ala Ile Val Asn Phe
900 905 910

Ala Val Pro Pro Glu Phe Tyr Gly Glu Phe
915 920

<210> 102
<211> 364
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-13-C transposition protein

<400> 102
Met Ala Val Gln Asp Gly Phe Pro Tyr Glu Leu Leu Asp Gln Pro Ile
1 5 10 15

Thr Asp Arg Val Ala Tyr Phe Arg His Tyr Thr Met Ala His Pro Val
20 25 30

Leu Leu Asn Ala Ala Gln Ile Leu Met Gln Thr Val Arg Glu Pro Ala
35 40 45

```
Gly Ile Ser Leu Val Phe Val Phe Gly Pro Thr Gly Val Gly Lys Ser
    50                  55                  60

Thr Leu Leu Arg Arg Val Ser Gln Lys Leu Thr Glu Ala Met Phe Ala
65                  70                  75                  80

Ala Leu Gln Gln Asp Gln Gly Arg Ile Pro Met Thr Gly Ile Glu Ala
                85                  90                  95

Ala Thr Pro Glu Phe Gly Gln Phe Asp Trp Lys Asp Phe Tyr Leu Arg
            100                 105                 110

Ala Leu Ala Ala Val Gln Glu Pro Phe Ile Tyr Pro Trp Ser Thr Ser
        115                 120                 125

Arg Leu Thr Thr Leu Lys Leu Arg Leu Ala Leu Glu His Ala Leu Arg
    130                 135                 140

Lys Arg Gln Leu Gln Val Phe Tyr Ile Asp Glu Ala Gln Asn Phe Ala
145                 150                 155                 160

Lys Val Ala Ser Ala Arg Lys Leu Ser Asp Gln Thr Asp Ala Ile Lys
                165                 170                 175

Ser Leu Ala Asn Leu Thr Gly Val Gln Phe Val Leu Thr Gly Thr Tyr
            180                 185                 190

Glu Val Leu Met Leu Arg Asn Leu Ser Ala Gln Leu Cys Arg Arg Ser
        195                 200                 205

Ile Asp Ile His Phe Pro Arg Tyr Arg Ala Glu Ile Pro Glu Asp Leu
    210                 215                 220

Gln Ser Phe Arg Gly Ile Val Gln Thr Phe Gln Arg His Leu Pro Leu
225                 230                 235                 240

Pro Gln Glu Pro Asn Leu Leu Glu Gln Trp Glu Phe Cys Tyr Glu Arg
                245                 250                 255

Ser Leu Gly Cys Val Gly Ile Leu Lys Asp Trp Leu Ser Arg Ala Leu
            260                 265                 270
```

```
Val Ala Thr Leu Glu Ala Gly Ala Lys Thr Leu Thr Pro Gln Met Leu
        275                 280                 285


Glu Gln His Ala Trp Ser Leu Glu Arg Cys Met Ile Met Leu Ala Glu
    290                 295                 300


Ala Lys Glu Glu Glu Lys Lys Leu Glu Val Gln Pro Leu Gln Ser Ala
305                 310                 315                 320


Leu Arg Ile Ala Leu Gly Leu Glu Thr Ala Val Val Glu Pro Ala Ser
                325                 330                 335


Pro Pro Thr Pro Ser Val Ser Pro Arg Arg Thr Ala Ile Ser Ala Lys
            340                 345                 350


Pro Lys Arg Tyr Pro Val Gly Gly Gly Gln His Asp
        355                 360


<210> 103
<211> 267
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
      MG64 transposition protein sequence

<220>
<223> MG64-13-Q transposition protein

<400> 103
Met Ile Glu Val Pro Ser Ile Phe Glu Ser Trp Ser Leu Glu Thr Phe
1               5                   10                  15


Gln Leu Pro Arg Arg Ser His Leu Tyr Ser Leu Ser Pro Ala Gly Leu
            20                  25                  30


Ala Thr Asp Ala Ile Glu Ser Leu Thr Ser Tyr Leu Thr Arg Leu Ala
        35                  40                  45


Glu Ala His Arg Val Ser Val Arg Lys Leu Leu Asp Glu Lys Ile Leu
    50                  55                  60
```

```
Leu Leu Leu Lys Pro Pro Arg Thr Phe Lys Tyr Glu Ile Gln Ser Ala
65                  70                  75                  80

Gln Ser Trp Leu Ala Ala Thr Pro Glu Val Ala Lys Ala Val Glu Cys
                85                  90                  95

Leu Arg Arg Leu Thr Leu Arg Gln Asp Leu Gln His Leu Thr Leu Leu
            100                 105                 110

Pro Trp Gln His Gln Leu Ser Thr Pro Ser Val Phe His Asp Val Gln
        115                 120                 125

Pro Phe Cys Pro Val Cys Phe Glu Glu Ala Arg Asn Thr Asp Ala Ala
    130                 135                 140

Val Tyr Glu Pro Leu Leu Trp Thr Leu Glu Ala Val Lys Val Cys Leu
145                 150                 155                 160

Arg His Arg Arg Tyr Leu Gln Phe Cys Cys Thr Tyr Cys Gly Arg Lys
                165                 170                 175

Gln Pro Phe Leu Lys Leu Glu Ala Arg Pro Gly His Cys Ala His Cys
            180                 185                 190

Gly Ser Trp Leu Gly Leu Glu Leu Pro Phe Phe Val Ser Thr Glu Pro
        195                 200                 205

Glu Glu Leu Asp Trp His Ile Asn Leu Ala Gln Met Leu Gly Gln Arg
    210                 215                 220

Leu Gln Ala Ala Trp Gln Arg Thr Asn Gln Ser Leu Ser Lys Ala Lys
225                 230                 235                 240

Arg Ser Lys Gln Ala Gln Lys Pro Ser Leu Met Ser Leu Leu Glu Ser
                245                 250                 255

Arg Tyr Glu Leu Asp Glu Ile Lys Leu Val Tyr
            260                 265
```

<210> 104

<211> 218
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 effector sequence

<220>
<223> MG64-54 effector

<400> 104
Met Thr Ile Ile Thr Val Gln Cys Arg Leu Lys Ala Pro Glu Ala Thr
1 5 10 15

Arg Arg Thr Leu Trp Glu Leu Met Ala Gly Arg Asn Thr Pro Leu Ile
20 25 30

Asn Lys Leu Ile Glu Leu Leu Arg Gly His Pro Asp Leu Asp Lys Trp
35 40 45

Arg Asp Lys Gly Lys Leu Pro Ala Gly Val Val Gly Glu Leu Cys Lys
50 55 60

Thr Leu Lys Ala Asp Pro Leu Tyr Asn Gly Gln Pro Ser Asn Ala Tyr
65 70 75 80

Lys Ser Ala Val Glu Thr Val Glu Glu Gln Phe Glu Ser Leu Leu Ala
85 90 95

Leu His Leu Glu Leu Lys Glu Lys Leu Arg Ile Lys Thr Arg Arg Leu
100 105 110

Ala Ile Phe Lys Ser Asn Ala Glu Leu Met Gln Leu Ser Gly Leu Thr
115 120 125

Leu Asp Asp Met Lys Lys Arg Ala Gly Glu Val Leu Asn Glu Leu Lys
130 135 140

Glu Asp Glu Arg Lys Lys Pro Ile Gly Thr Leu Tyr Arg Lys Tyr Glu
145 150 155 160

Leu Leu Phe Phe Asp Asp Pro Asp Arg Trp Ala Val Cys Tyr Leu Ile
165 170 175

Lys Asn Gly Gly Lys Leu Pro Lys Gly Lys Glu Asp Pro Glu Lys Phe
180 185 190

Glu Lys Ser Arg Arg Lys Leu Glu Ile Met Val Gln Lys Leu Thr Met
195 200 205

Arg Leu Asp Ser Cys Leu Pro Gly Gly Arg
210 215

<210> 105
<211> 541
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-54-B transposition protein

<400> 105
Met Asp Lys Thr Asn Asn Pro Thr Leu Thr Glu Gly Leu Glu Gly Glu
1 5 10 15

Ala Ala Leu Arg Leu Glu Val Ile Thr Asp Leu Leu Gln Pro Cys Asp
20 25 30

Arg Ser Thr Tyr Gly Gln Arg Leu Arg Glu Gly Ala Ala Lys Leu Gly
35 40 45

Ile Ser Val Arg Ser Leu Gln Arg Leu Phe Lys Arg Tyr Gln Glu Glu
50 55 60

Gly Leu Ser Ala Leu Met Ser Gly Pro Arg Arg Asp Arg Gly Ala His
65 70 75 80

Arg Ile Gly Glu Glu Trp Ile Gln Phe Ile Val Lys Thr Tyr Gln Asp
85 90 95

Gly Asn Lys Gly Ser Lys Arg Leu Thr Pro Lys Gln Val Ala Ile Arg
100 105 110

```
Val Arg Ala Gln Ala His Gln Lys Gly Gln Glu Ser Tyr Pro Ser Tyr
        115                 120                 125

Arg Thr Val Leu Arg Val Leu Glu Pro Ser Ile Glu Asn Lys Asn Lys
    130                 135                 140

Ser Ile Arg Ser Pro Gly Trp Arg Gly Glu Ala Leu Ser Leu Lys Thr
145                 150                 155                 160

Arg Asp Gly Asp Ala Leu Glu Ala Ile Arg Ser Asn Gln Val Trp Gln
                165                 170                 175

Cys Asp His Thr Arg Ala Asp Val Leu Leu Val Asp Arg Glu Gly Arg
            180                 185                 190

Leu Ile Gly Arg Pro Trp Leu Thr Thr Val Ile Asp Ser Tyr Ser Arg
        195                 200                 205

Cys Ile Val Gly Ile Asp Leu Ser Phe Asp Ala Pro Ser Ser Gln Lys
    210                 215                 220

Val Ala Leu Ala Leu Arg His Ala Ile Leu Pro Lys Leu Tyr Gly Ala
225                 230                 235                 240

Glu Tyr Lys Leu Asn Cys Glu Trp Asn Ala Tyr Gly Lys Pro Glu Cys
                245                 250                 255

Leu Phe Thr Asp Gly Gly Lys Asp Phe Arg Ser Asn His Leu Ala Glu
            260                 265                 270

Ile Ala Ala Gln Leu Gly Phe Ile Leu Lys Leu Arg Asp Arg Pro Ser
        275                 280                 285

Glu Gly Gly Ile Val Glu Arg Pro Phe Lys Thr Leu Asn Gln Ser Leu
    290                 295                 300

Phe Ser Thr Leu Pro Gly Tyr Thr Gly Pro Asn Val Lys Glu Arg Pro
305                 310                 315                 320

Glu Asp Ala Glu Lys Glu Ala Thr Leu Thr Leu Arg Glu Leu Glu Ile
```

```
                325                 330                 335

Leu Ile Val Arg Tyr Ile Val Asp Asn Tyr Asn Gln Lys Gln Asp Ala
            340                 345                 350

Arg Thr Gly Asp Gln Thr Arg Ala Gln Arg Trp Glu Gly Gly Ile Asp
        355                 360                 365

Leu Pro Glu Pro Leu Ser Glu Arg Glu Leu Asp Ile Cys Leu Leu Lys
    370                 375                 380

Ser Thr Arg Arg Lys Ile Gln Arg Gly Gly His Leu Gln Phe Glu Asn
385                 390                 395                 400

Ile Thr Tyr Arg Gly Glu His Leu Gly Gly Tyr Glu Gly Asp Tyr Met
                405                 410                 415

Asn Leu Arg Tyr Asp Pro Arg Asp Ile Thr Ser Ile Trp Val Tyr Arg
            420                 425                 430

Gln Glu Ser Ser Gly Asp Val Phe Leu Cys Arg Ala His Ala Val Gly
        435                 440                 445

Leu Gln Thr Glu Ile Leu Ser Leu Asn Asp Ala Gln Ala Ala Ala Lys
    450                 455                 460

Arg Leu Arg Asp Lys Gly Arg Gly Leu Asp Asn Asn Ser Ile Leu Gln
465                 470                 475                 480

Glu Ser Ile Glu Arg Glu Ala Thr Val Thr Lys Asn Leu Gln Arg Arg
                485                 490                 495

Lys Lys Glu Glu Gln Ala Tyr Lys Thr Pro Thr Pro Ala Pro Glu Glu
            500                 505                 510

Lys Val Glu Leu Thr Pro Glu Val Ser Lys Ile Glu Asp Ile Glu His
        515                 520                 525

Leu Glu Val Trp Asp Leu Asp Glu Leu Asn Gly Trp Gly
    530                 535                 540
```

<210> 106
<211> 268
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-54-C transposition protein

<400> 106
Met Glu Asp Lys Lys Leu Pro Ile Asp Gln Glu Trp Leu Asn Gly Glu
1 5 10 15

Ile Ala Arg Leu Asn Lys Lys Thr Val Ile Pro Ile Glu Gln Val Lys
20 25 30

Ser Leu His Asp Trp Leu Asp Glu Lys Arg Ala Ala Arg Gln Ser Cys
35 40 45

Arg Val Val Gly Glu Ser Arg Thr Gly Lys Thr Val Thr Cys Gln Ser
50 55 60

Tyr Ala Leu Arg Asn Thr Pro Arg Leu Ser Gly Asn Thr Pro Pro Asn
65 70 75 80

Leu Pro Val Val Tyr Val Val Pro Ala Gln Lys Cys Gly Pro Lys Asp
85 90 95

Phe Phe Gly Ala Ile Ile Gly Ala Leu Ile Tyr Arg Ala Val Lys Gly
100 105 110

Ser Thr Ser Glu Leu Arg Asn Arg Ala Thr Glu Ile Leu Lys Ser Cys
115 120 125

Gly Val Glu Met Leu Ile Ile Asp Glu Ala Asp Arg Leu Arg Pro Glu
130 135 140

Thr Phe Ser Glu Val Arg Asp Ile Tyr Asp Asn Leu Glu Ile Ala Val
145 150 155 160

```
Val Leu Val Gly Thr Asp Arg Leu Asp Ala Ala Ile Lys Arg Asp Glu
                165                 170                 175


Gln Val Tyr Asn Arg Phe Ser Ala Cys Arg Arg Phe Gly Leu Leu Thr
            180                 185                 190


Gly Glu Asp Phe Ile Arg Ala Val Ala Ile Trp Glu Asp Lys Ile Leu
        195                 200                 205


Lys Leu Pro Val Ala Ser Asn Leu Thr Ser Lys Glu Thr Leu Lys Ile
    210                 215                 220


Leu Thr Arg Gly Thr Gly Gly Tyr Ile Gly Lys Leu Asp Gln Ile Leu
225                 230                 235                 240


Lys Gly Ala Ala Ile Arg Ser Leu Arg Lys Gly His Lys Arg Val Glu
                245                 250                 255


Thr Glu Val Leu Lys Glu Val Val Lys Glu Phe Ser
            260                 265


<210> 107
<211> 178
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
      MG64 transposition protein sequence

<220>
<223> MG64-54-Q transposition protein

<400> 107
Met Ser Glu Asp Ile Lys Pro Trp Leu Leu Pro Val Gln Pro Glu Pro
1               5                   10                  15


Gly Glu Ser Leu Ser Ser Phe Leu Gly Arg Val Arg Gln Arg Asn His
            20                  25                  30


Leu Thr Ala Gly Ala Leu Gly Asn Ile Thr Arg Leu Gly Ala Thr Ile
        35                  40                  45


Ala Arg Trp Glu Lys Phe Tyr Leu Asn Pro Phe Pro Ser Ala Ala Gln
```

50 55 60

Leu Glu Ala Met Ala Ala Ile Val Gly Leu Thr Gly Glu Phe Ile Arg
65 70 75 80

Asp Met Leu Pro Pro Thr Gly Glu Ala Met Lys Cys Glu Pro Ile Arg
85 90 95

Leu Cys Pro Glu Cys Tyr Ala Glu Ser Pro Tyr His Arg Leu Gln Trp
100 105 110

Gln Phe Gln Ser Val Trp Lys Cys Thr Ser His Gly Asn Lys Leu Leu
115 120 125

Ser Lys Cys Pro Val Cys Ser Ala Ser Phe Gly Ile Pro Ser Asn Trp
130 135 140

Asn Gly Tyr Cys Arg Arg Cys Gly Thr Pro Phe Glu Leu Met Arg Ala
145 150 155 160

Arg Ser Glu Ser Ile Pro Ala Leu Ser Pro Val Asp Asp Lys Gly Tyr
165 170 175

Asp Arg

<210> 108
<211> 638
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG108 effector sequence

<220>
<223> MG108-2 effector

<400> 108
Met Ser Gln Ile Thr Ile Gln Cys Arg Leu Val Ala Ser Glu Ser Thr
1 5 10 15

Arg Arg Glu Leu Trp Glu Leu Met Ala Glu Lys Asn Thr Pro Leu Ile
20 25 30

```
Asn Glu Leu Leu Glu Gln Ile Gly His His Pro Asp Phe Glu Thr Trp
        35                  40                  45

Arg Glu Lys Gly Lys Leu Pro Ala Asp Val Val Lys Gln Leu Ala Glu
    50                  55                  60

Pro Leu Lys Thr Asp Pro Arg Phe Ile Gly Gln Pro Gly Arg Phe Tyr
65                  70                  75                  80

Thr Ser Ala Ile Ala Leu Val Lys Tyr Ile Tyr Lys Ser Trp Phe Ala
                85                  90                  95

Val Met Met Gln Leu Gln Tyr Gln Leu Lys Gly Lys Ile Arg Trp Leu
            100                 105                 110

Glu Met Leu Lys Ser Asp Glu Glu Leu Val Glu Thr Ser Gly Ile Thr
        115                 120                 125

Leu Asp Asn Leu Arg Ile Lys Ala Ala Glu Ile Leu Ala Gln Cys Thr
    130                 135                 140

Pro Gln Pro Asp Ser Val Glu Ser Gln Glu Lys Lys Ser Lys Lys Arg
145                 150                 155                 160

Lys Lys Gly Lys Lys Thr Lys Lys Ser Asp Ser Gln Pro Ser Leu Ser
                165                 170                 175

Lys Asn Leu Phe Asp Thr Tyr Asp Ser Thr Glu Asp Ile Ile Ser Lys
            180                 185                 190

Ser Ala Ile Val Tyr Leu Leu Lys Asn Gly Cys Lys Leu Thr Asp Lys
        195                 200                 205

Glu Glu Asp Pro Lys Glu Phe Ala Lys Arg Arg Arg Thr Thr Glu Ile
    210                 215                 220

Gln Ile Gln Arg Leu Thr Glu Gln Leu Ser Ala Arg Ile Pro Lys Gly
225                 230                 235                 240
```

```
Arg Asp Leu Thr Asn Ala Lys Trp Leu Glu Thr Leu Ala Ile Ala Thr
                245                 250                 255

Glu Asn Ala Pro Glu Asn Glu Thr Glu Ala Lys Ser Trp Gln Asn Lys
            260                 265                 270

Leu Leu Arg Lys Ser Ala Ser Val Pro Phe Pro Val Ser Tyr Glu Thr
        275                 280                 285

Asn Glu Asp Leu Thr Trp Phe Lys Asn Asn Lys Gly Arg Leu Cys Val
    290                 295                 300

Tyr Phe Asn Gly Leu Ser Glu His Ile Phe Gln Ile Tyr Cys Asp Ser
305                 310                 315                 320

Gln His Leu His Trp Phe Gln Arg Phe Leu Ala Asp Gln Glu Val Lys
                325                 330                 335

Glu Asn Ser Lys Asn Gln His Ser Ala Ser Leu Phe Thr Leu Arg Ala
            340                 345                 350

Gly Lys Ile Val Trp Gln Glu Gly Glu Gly Lys Gly Glu Pro Trp Asn
        355                 360                 365

Ile Asn His Leu Thr Leu His Cys Ser Val Asp Thr Arg Leu Trp Thr
    370                 375                 380

Ala Glu Gly Thr Gln Gln Val Arg Lys Glu Lys Glu Val Glu Ile Ile
385                 390                 395                 400

Lys Phe Leu Asn Lys Asn Lys Asp Arg Ser Asp Leu Ser Lys Asn Gln
                405                 410                 415

Lys Ala Asn Ile Lys Arg Gln Asn Ser Thr Leu Val Lys Ile Ala Asn
            420                 425                 430

Pro Phe Pro Arg Pro Ser Gln Pro Leu Tyr Gln Gly Gln Gly Gln Ile
        435                 440                 445

Leu Val Gly Val Ser Ile Gly Leu Glu Ser Pro Ala Thr Leu Ala Val
    450                 455                 460
```

```
Val Asp Ala Ile Thr Lys Lys Val Ile Thr Tyr Arg Ser Ile Arg Gln
465                 470                 475                 480

Leu Leu Gly Glu Asn Tyr Arg Leu Leu Asn Arg Gln Arg Arg Gln Lys
                485                 490                 495

Gln Ser Leu Ser His Gln Arg Gln Lys Ala Gln Arg Leu Ala Asp Gly
            500                 505                 510

Asn Gln Leu Gly Glu Ser Glu Leu Gly Gln Tyr Ile Asp Arg Leu Ile
        515                 520                 525

Ala Lys Lys Ile Val Val Ile Ala Ala Ala Tyr Asn Ala Gly Ser Ile
    530                 535                 540

Val Leu Pro Lys Leu Gly Asn Val Arg Gln Ile Ile Gln Ser Glu Met
545                 550                 555                 560

Asp Ala Leu Ala Glu Gln Lys Cys Pro Glu Tyr Lys Glu Gly Gln Lys
                565                 570                 575

Lys Tyr Ala Lys Gln Tyr Arg Ile Asn Ile His Gln Trp Ser Tyr Gly
            580                 585                 590

Arg Leu Ile Glu Cys Val Lys Thr Gln Ala Ala Lys Ile Gly Ile Val
        595                 600                 605

Ile Glu Glu Ala Lys Gln Ser Ile Arg Gly Ser Pro Gln Glu Lys Ala
    610                 615                 620

Lys Glu Leu Ala Leu Ala Ala Tyr Asp Ser Arg Ser Asn Asp
625                 630                 635
```

<210> 109
<211> 572
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG108 transposition protein sequence

<220>
<223> MG108-2-B transposition protein

<400> 109
Met Tyr Met Gln Asp Asp Arg Thr Arg Ser Ala Glu Ile Pro Ala Pro
1   5   10   15

Ser Ser Gly Glu Gly Asn Ser Glu Val Glu Thr Ser Glu Val Ile Thr
            20   25   30

Glu Leu Pro Leu Glu Ala Arg Leu Lys Gln Glu Ile Ile Gln Ser Leu
        35   40   45

Leu Glu Pro Cys Asp Arg Ile Thr Tyr Gly Gln Arg Gln Arg Glu Ala
    50   55   60

Ala Glu Lys Leu Gly Val Ser Val Arg Thr Val Arg Arg Leu Val Lys
65   70   75   80

Lys Trp Glu Glu Lys Gly Leu Glu Ala Leu Thr Gln Thr Ala Arg Thr
                85   90   95

Asp Lys Gly Lys His Arg Ile Asp Glu Glu Trp Gln Glu Phe Ile Ile
            100   105   110

Lys Thr Tyr Lys Glu Gly Asn Lys Gly Ser Lys Arg Ile Thr Pro Lys
        115   120   125

Gln Val Ala Val Arg Val Gln Val Arg Ala Ala Gln Leu Gly Glu Glu
    130   135   140

Lys Tyr Pro Ser Tyr Arg Thr Val Tyr Arg Val Leu Glu Pro Ile Ile
145   150   155   160

Glu Gln Gln Glu Gln Thr Gln Asn Ile Arg Ser Arg Gly Trp Arg Gly
                165   170   175

Ser Arg Leu Ser Val Lys Thr Arg Thr Gly Gln Asp Leu Ser Val Glu
            180   185   190

Tyr Ser Asn His Val Trp Gln Cys Asp His Thr Arg Val Asp Val Leu

```
        195                 200                 205

Leu Val Asp Arg His Ser Lys Ile Leu Gly Arg Pro Trp Leu Thr Thr
    210                 215                 220

Val Ile Asp Thr Tyr Ser Arg Cys Ile Val Gly Ile Asn Leu Gly Phe
225                 230                 235                 240

Asp Ala Pro Ser Ser Gln Val Val Ala Leu Ala Leu Arg His Ala Ile
                245                 250                 255

Leu Pro Lys Gln Tyr Gly Leu Glu Tyr Lys Leu His Cys Glu Trp Gly
            260                 265                 270

Thr Phe Gly Lys Pro Glu His Phe Tyr Thr Asp Gly Gly Lys Asp Phe
        275                 280                 285

Arg Ser Asn His Leu Gln Gln Ile Gly Val Gln Leu Gly Phe Val Cys
    290                 295                 300

His Leu Arg Asp Arg Pro Ser Glu Gly Gly Ile Val Glu Arg Pro Phe
305                 310                 315                 320

Gly Thr Phe Asn Thr Glu Leu Phe Ser Thr Leu Pro Gly Tyr Thr Gly
                325                 330                 335

Ser Asn Val Gln Glu Arg Pro Glu Glu Ala Glu Lys Glu Ala Cys Leu
            340                 345                 350

Thr Leu Arg Glu Leu Glu Gln Leu Leu Val Arg Tyr Ile Val Asp Lys
        355                 360                 365

His Asn Gln Ser Ile Asp Ala Arg Met Gly Asp Gln Ser Arg Phe Gln
    370                 375                 380

Arg Trp Glu Ala Gly Leu Ile Ala Val Pro Ser Leu Met Gly Glu Arg
385                 390                 395                 400

Glu Leu Asp Ile Cys Leu Met Lys Gln Thr Arg Arg Thr Ile Tyr Arg
                405                 410                 415
```

```
Gly Gly Tyr Leu Gln Phe Glu Asn Leu Thr Tyr Arg Gly Glu His Leu
            420                 425                 430

Ala Gly Tyr Ala Gly Glu Ser Val Val Leu Arg Tyr Asp Pro Arg Asp
        435                 440                 445

Ile Thr Thr Ile Leu Val Tyr Arg Ser Glu Glu Gly Lys Glu Val Phe
    450                 455                 460

Leu Ala Arg Ala Tyr Ala Gln Asp Leu Glu Thr Glu Gln Leu Ala Leu
465                 470                 475                 480

Asp Glu Ala Lys Ala Ser Ser Arg Lys Ile Arg Glu Ala Gly Lys Thr
                485                 490                 495

Val Ser Asn Arg Ser Ile Leu Ala Glu Val Gln Glu Arg Glu Thr Phe
            500                 505                 510

Leu Thr Gln Lys Lys Thr Lys Lys Glu Arg Gln Lys Glu Glu Gln Val
        515                 520                 525

Gln Leu Arg Lys Ala Lys Gln Pro Leu Pro Ile Glu Pro Glu Glu Ile
    530                 535                 540

Glu Val Ala Ser Val Arg Asp Glu Ala Glu Ser Glu Met Pro Glu Val
545                 550                 555                 560

Phe Asp Tyr Asp Gln Met Arg Glu Asp Tyr Gly Trp
                565                 570
```

<210> 110
<211> 240
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG108 transposition protein sequence

<220>
<223> MG108-2-C transposition protein

<400> 110

```
Met Thr Thr Gln Glu Ala Lys Thr Val Ala Gln Gln Leu Gly Asp Phe
1               5                  10                  15

Pro Val Asn Asp Thr Lys Leu Gln Ala Glu Ile Gln Arg Leu Asn Arg
            20                  25                  30

Lys Thr Phe Val Pro Leu Glu Gln Val Glu Asn Leu His Glu Trp Leu
        35                  40                  45

Glu Glu Lys Arg Gln Ala Arg Gln Ser Gly Arg Val Val Gly Glu Ser
    50                  55                  60

Arg Thr Gly Lys Thr Met Gly Cys Asn Ala Tyr Arg Leu Arg His Lys
65                  70                  75                  80

Pro Lys Gln Glu Leu Gly Lys Pro Pro Ile Val Pro Val Val Tyr Phe
                85                  90                  95

Lys Val Thr Gln Glu Cys Ala Ala Lys Glu Phe Phe Gly Leu Ile Ile
            100                 105                 110

Gln Asn Leu Lys Tyr Gln Val Val Lys Gly Thr Val Pro Glu Leu Arg
        115                 120                 125

Ala Arg Ala Val Gln Val Leu Lys Gly Cys Arg Val Glu Met Leu Ile
    130                 135                 140

Leu Asp Glu Ala Asp Arg Leu Lys Pro Lys Thr Phe Ala Glu Val Gln
145                 150                 155                 160

Asp Ile Phe Asp Asp Leu Glu Leu Ala Val Ile Leu Val Gly Thr Asp
                165                 170                 175

Arg Leu Asp Ala Val Ile Lys Arg Asp Glu Gln Val Tyr Asn Arg Phe
            180                 185                 190

Arg Ala Cys His Arg Phe Gly Lys Leu Ser Gly Glu Glu Phe Lys Arg
        195                 200                 205

Thr Val Glu Ile Trp Glu Lys Lys Val Leu Gln Leu Pro Val Ala Ser
    210                 215                 220
```

Asn Leu Phe Ser Lys Thr Ile Glu Arg Gly Lys Cys Ser Gln Asn Gly
225 230 235 240

<210> 111
<211> 261
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<220>
<223> MG64-6 active effector sgRNA

<220>
<221> modified_base
<222> (239)..(261)
<223> a, c, u, g, unknown or other

<400> 111
auaacagcgc cgcaggucau gccgucaaaa gccucugaac uguguuaaau ggggguuagu 60

uugacuguug aaagacaguu gugcuuucug acccugguag cugcccaccc ugaugcugcu 120

aucuuucggg auaggaauaa ggugcgcucc caguaauagg gguguagaug uacuacagug 180

guggcuacua aaucaccucc gaccaaggag gaauccaucc gaaaggaugg guugaaagnn 240

nnnnnnnnnn nnnnnnnnnn n 261

<210> 112
<211> 263
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<220>
<223> MG64-2 effector sgRNA

<220>
<221> modified_base
<222> (241)..(263)
<223> a, c, u, g, unknown or other

<400> 112
gaauuaauag cgccgccguu caugcuucua ggagccucug aaaggugaca aaugcggguu 60

aguuuggcug uugucagaca gucuugcuuu cugacccugg uagcugccca ccccgaagcu 120

gcuguuccuu gugaacagga auuaggugcg cccccaguaa uaaggguaug gguuuaccac 180

agugguggcu acugaaucac cuccgagcaa ggaggaaccc acugaaaggu ggguugaaag 240

nnnnnnnnnn nnnnnnnnnn nnn 263

<210> 113
<211> 265
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-4 effector sgRNA

<220>
<221> modified_base
<222> (243)..(265)
<223> a, c, u, g, unknown or other

<400> 113
gaaauagcgc cgcuuguuca ugcgcgaaag cgucucugaa caguguaaau gugggguuagu 60

uugacugucg ugaagacggu cuugcuuucu gacccuggua gcugcccacc uugaagcugc 120

ugucucuugu agacaggaau caggugcgcc cccaguaaua uaggugcggg uuuaccgcag 180

ugguggcuac ccaaucaccu ccgagcaagg aggaacccac cuugaaagag gcgggguugaa 240

agnnnnnnnn nnnnnnnnnn nnnnn 265

<210> 114
<211> 276
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-7 effector sgRNA

<220>
<221> modified_base
<222> (254)..(276)
<223> a, c, u, g, unknown or other

<400> 114
guaacaaaca gcgccgcagu ucaugcgucu uauggcgccu cugugcugug caaaaugugg 60

guuaguuuga cuguuggaag acagucuugc uuucugaccc ugguagcugc ccaccuugaa 120

gcugcuaucc cuuguggaua ggaaucaggu gcgcccccag uaauagaggu gcggguuuac 180

cgcaguggug gcuaccgaau caccuccgag caaggaggaa cccaccuuga aaaagcuuug 240

ggcggguuga aagnnnnnnn nnnnnnnnnn nnnnnn 276

<210> 115
<211> 246
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG108-1 effector sgRNA 1

<220>
<221> modified_base
<222> (224)..(246)
<223> a, c, u, g, unknown or other

<400> 115
gaugcuaagu cgcgccuagc aucaaggagc uaugucuuga uugucuuggg uguccgcccu 60

ggaugaguug agguguagau gcuucuauca uggcagcuac uaaacgcccc aagcaagggg 120

aacccaucuu uaauuuuggc aaaccgaagc gggggcaaaa ucuccaggag guucgccaaa 180

accuuugaaa cuccuuagcu ggaaauggcg auggggguuga aagnnnnnnn nnnnnnnnnn 240

nnnnnn 246

<210> 116
<211> 274
<212> RNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG108-1 effector sgRNA 2

<220>
<221> modified_base
<222> (252)..(274)
<223> a, c, u, g, unknown or other

<400> 116
cuuguucgag uucguuugaa cugaacaagg guaaguaugg gccaguuuaa uugcuuuccg 60

ucccaggaua gcugccagcu ucuaccguag guucguccug caagugaugc uaagucgcgc 120

cuagcaucaa ggagcuaugu cuugauuguc uuggugucc gcccuggaug aguugaggug 180

uagaugcuuc uaucauggca gcuacuaaac gccccaagca aggggaaccc aucgaaagau 240

gggguugaaa gnnnnnnnnn nnnnnnnnnn nnnn 274

<210> 117
<211> 550
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-13 effector intergenic region encoding tracrRNA

<400> 117
gttgaggtgg atttgtgtac tgggtcaaaa gatttcgcta ggtttgcctt ttcgcgcgct 60

gctgaagcag tgaaaagttc cgacagttta aaagttctaa cacttttact gctttccgtc 120

gggagtagtt gtccgcttct gcgttgttga gaccgaagtt tgtccaggtg ataggcagaa 180

cgttgatgca acaggtagat agtcgcgctc tatcaggaag ctgttctagt tattagtgtt 240

cgtgcactaa taaagaggat acagggatac atgtggttgt gtccagcaat ggcacaacag 300

ggccactact cgaagccccg agcaagggtt gagcctaccc aaattttgta ctaaaaattt 360

ggcaaaccga agcgaggtca gtttgaccca gcaccttgtc aatcagtcga aatcttgcac 420

cagtctgcgt tttgagttgg caacagcctc acgctgaacc acttcaagca gaaagttatg 480

gcaatgtgat gacctttgcc aaaataggtg tgcaaagccg actgtagcaa gaatttgagc 540

actaggagtt 550

<210> 118
<211> 457
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG108-2 effector intergenic region encoding tracrRNA

<400> 118
tactagaaaa atttgcacct tgataataga ataatagtca caatagcgcc gtagttcatg 60

cttgctaaag cctctgaatt gcgaaaagtc cgggttagtg ctgtcggcag acagcgttgc 120

tttctgaccc tggtagctgc ccaccccgat gctgctgtcc cttgcagaca ggaaccaggt 180

gcgcccccag taataagggt gtgggtttac cacagtggtg gctactgaat cacctccgag 240

caaggaggaa tccaccttaa ctatttttc gtgaatccaa gcgggagcaa aattcccaag 300

ggatagacga atttggtgaa ataaatagcc agacttaagt ttgagctttt tctgggatgg 360

tcaaccctcc tgttacgaca gtttccaaag taaaattttc gagttttacg agttttgttt 420

ctggaagcta gtacacaaaa gggttgaagc gtttaaa 457

<210> 119
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-13 effector target CRISPR repeat

<400> 119
tggcaattgc ccttccagtg ttgggtgggt tgaaag 36

<210> 120
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG108-2 effector target CRISPR repeat

<400> 120
gtttcaacga ccatcccgac aaggggtggg ttgaaag 37

<210> 121
<211> 60
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG110-1-6 effector crRNA

<220>
<221> modified_base
<222> (9)..(40)
<223> a, c, u, g, unknown or other

<400> 121
cugguaaann nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn gugaccugcc gcauaggcag 60

<210> 122
<211> 59
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG110-2-6 effector crRNA

<220>
<221> modified_base
<222> (7)..(38)
<223> a, c, u, g, unknown or other

<400> 122
ccaagannnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnua ugaacugccg cauaggcag 59

<210> 123
<211> 385
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-6 active transposon end LE

<400> 123
gggaggaaag gttaattggg gggttgattt attttggttg gttgtgttat atttgttatt 60

cgcgggcgat tagcacagtg gtagcgcact tccttcacac ggaaggggtc acaggttcaa 120

atcctgtatc gcccatttgt gtacagtgac acattaattg tcatcaatga cagattgctg 180

tcgtggagcc aaattatgtg tcgctgagac aaattaatgt cgtttaacta tcagtgacaa 240

atttttgtcg cttttcacaa caatagtgtg aaggaagtgc gcctttcaat ccatcctaga 300

aattataatt ccaatcccta cttacctaga atggtggttg aaactgtaag attcgcgccg 360

ctaataaaac tttcagcgat ttgga 385

<210> 124
<211> 313
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-6 active transposon end RE

<400> 124
aacacataat gctatcctag cttacaaaaa agaaagcgac aatcaatctg tcactcctca 60

aattctcttc tttgagaacg acgacagcta aattgtcact gattgggagg acgacactta 120

atttgtcact aacggctagc gatctttaat caaggaaaac agctagaata tagagaaatc 180

aatagttttc atcagcgaca acaatttggc atcacgtcaa ataattagtc actgtacagg 240

ttttatttt aatgtacagt agccctcttg caaaagtcca agaattagga atgttagagg 300

aggtttgata act 313

<210> 125
<211> 300
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-2 putative transposon end LE

<400> 125
ttgtaaaacc tcagatgagc aaggtattga tgaataatat tttccgccaa aatttacctc 60

tgtttgggca ataattgtat gatagaattt atatattagt tgcttcaata gcaaaatgcg 120

ggtgtagttc agtggtagaa cgtcaccttc ccaaggtgaa tgtcgtgggt tcgagtccca 180

tctcccgctt gtgttgttgt tgtacagtgt tgtcgattgc caaattattt gtccgcgttg 240

acaaattgat gtccgtttgc caaattattt gtcatttaca aaactcgaac ccacgacatt 300

<210> 126
<211> 400
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-2 putative transposon end RE

<400> 126
atctacaaaa ttgttttata caactatgtt tgtattttag ttctcaggac aaagattgtg 60

tcaattcttc acaaaagttg taattctcgc ctttttcaaa aaaggacaaa gaatttgtca 120

actttcctaa accggacaac taatttgtca aaaagctttt tcaatatata aaggggtat 180

aatagctgta acccttgcgt agagtaggtt agtgagctac taggtttggc ggtcaataat 240

ttggcaagtg gtcaaataat ttggcaatcg acaattattt atcgttgggg acgaatcagt 300

gtcacggagt caaattatgt gtcgttgtga cagattgatg tcgtctaagc accagtgaca 360

aattgatgtt gctttaataa tttaaaatcg tcacgaaata 400

<210> 127
<211> 489
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-4 putative transposon end LE

<400> 127
tttcgggttc tccagttgtg ctcacaaggc tgcccgtttg gcttcatccc gttggcaata 60

cccatacgtt ttttttacg gctgtgtcca atcttccgta gtgccctcga aatcgtgcgc 120

tgactaatct ccccctccca tagctgtgcc atctcgcttt gggttttatc cccatgctct 180

ttgacgaaag cccgaaaatt ctcccaatcg ctgattttgc ctgtctgttg gcatgcttgc 240

cttggctttg gcttgacatc gcccgtttcg gcttggcgct caaaccataa gttaatcgta 300

ttgcgactga tattgaacag ttgactcgcc tcacttttct tcagaccgtc caactcaatc 360

gcttgcatta ctttttgccg gaagtcatca ctgtaaggtt tagccattga aatgaggtag 420

acgcaatggg tattcttcta gcttacgtcc taacctgcat ggctatagct atactatcca 480

cgacttgga 489

<210> 128
<211> 602
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-4 putative transposon end RE

<400> 128
aagtgttcaa tcaaaacaca tatgtattat aaagttaaaa ggacatttat tttgtcaatg 60

acagtaaaaa gtcacaaaaa tattgatttc agaaaaaaga catatatttt gtcgaattta 120

caatatagga cggctaattc gtcaataact tcgtgctagg aattagcgat agaaaaacaa 180

taaaaccctt gcatagcaaa ggctttagag taaataataa gttgtcaatt aatttggcaa 240

acggtcagat aatgtgacac tcgacagaaa tcctttctag ggaaaggttt tgttgtttat 300

atatctaact ttgttcgagc gctcattctc aataagcagc ctagaaatgg gtacagatcg 360

agatatctag agccggaatt gagataggat tgagataatt taatcagggc cattaagcat 420

atggcaaggg tttcggacgc tacgaagaaa ccacatgaga taaacccaag ctctcaatct 480

ttatagctcg taagctcaac aatttcttaa gtcacattga aagcactggt tctaacgatt 540

tcaatctcta taggcagtca gaaaagaatc caagtgaaga taaacctttg tttgtgctag 600

tg 602

<210> 129
<211> 681
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-7 putative transposon end LE

<400> 129
agttaggagt ttgaagttag gagtttgaag ttaccactca gcactcaata ctcagcactt 60

tctactcatt tccctttctt tgtggtacac tgcaatatag gctagtcaaa ctatgccggg 120

atgtagcgca gcttggtagc gcacttcgtt cgggacgaag gggtcgctgg ttcgaatcca 180

gtcatcccga taattattta gttatttata cgacatatta cacgtaatga aagctgtaca 240

ttaactaatt atttgtcact gttaacagat tgatgtcaaa tttcacaaat tgatgtcact 300

attaacagat tagtgtgctt gagctagccg ttaagactca gcttttgatt gtagctgaaa 360

tttatttgtt tgctatgtta caggctcacc cttgaaatta attataactg gacatgactt 420

tacttgttca aatacgaact gttctggtat ataatgattt ttgtttgggg tcatacttca 480

aactgttagg attattttgc aatatacatt tcggcagttt ttgacccttc ctaaaatctt 540

gtgagaaatc cgggacaact ggattcgctt tcaacccacc actaataatt ttaaattatt 600

ctttccactg ttgattaatc catttaattt cagctaaggc ttaaaacatc aaaatattca 660

aagtatctca ctgccaacag a 681

<210> 130
<211> 415

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-7 putative transposon end RE

<400> 130
atttgtgttt tataaaaaca tagttgtatt atactgacac ctggacaaag aatttgttaa 60

attttgcaaa atcattggtt tgtttgtttt ttggaaaatg acaattaatg tgttatattt 120

ctgagattgg acagttaata tgttaattaa ggtttgaaag gctttaaatc tgaaaataat 180

aaaagcccca tcagaaacct ttctatgaca gggctttcaa tatttgatag agttgtgaca 240

ataatttgtt aagttgacaa ataattagtt aatgtacaaa atatttaatg tatagttaaa 300

atattgtaat agtaacggtt ttgtctatga attatagcaa aggacgggca aggatgcccg 360

ttccacaaga ttttgatttg tagtttactt tggaattttt ctttttatta tttaa 415

<210> 131
<211> 260
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-13 putative transposon end LE

<400> 131
tgtcactgta tccataaggt gagccattgt gaaagcataa gttgagccaa aaatgcttga 60

aatgctgatt tttgcgttgc ttgatttagt ttaatccaaa cggtttgaaa gcataacgtg 120

agccaaagtt ttattttgaa agcataagct gagcaaacat ttgcggcaac ggcttgaaca 180

gcctcaaaat tgtgattagc tgaaatcaga agcgatgagt tccggtttca cggttgtaac 240

tcttcaggag gctgcgaaga 260

<210> 132
<211> 1979
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-13 putative transposon end RE

<400> 132
agagcctcac tttatcgcat tgccacacaa gagcgcacac ataaaacatt aggggtaggt 60

tgaaagaggc aggggtctgc aaattaaata ccaagcttga cctgaagcag cattagcttt 120

aggaactgcg atcgccatca tgcgaactgc tttattaatg cagtgagcag aaattcaagt 180

ttcgacgagt aattgtcgat tattagaagg agttacaaag tatggagcca gaagatacgg 240

aattgagagt gagcaaacca atccaagtgg gagcaagaaa tcaattctgg caaatagatt 300

gcactgtttt tgatgtcagt ctggtcgaga aaactactcc agattattgg caaatggatt 360

ttactgctct tgatgtaaat ctggtcgggg aaactacttc agatgaccct ctttaaccct 420

agctactgtg cgcgacacac aatcgcaaac tatgttggga taccaactta catctgaaga 480

ggtacatggg agtgcgatcg caacactact tcatgctgcc gctcgaaaaa ctatctatca 540

caatatgaat tatttaatct gcgttttcat ggaggcatga tggaagaaca taaatttgga 600

atcactaatg cagttatcct taattcaata tcagggtacg tcaaactcac aagaagaaga 660

cccagcaccc tcaaggatat tgaagagatg gtgaatcgat attttcagca aagctcaggt 720

gactagaaac aatattcatc cacgaatgct tgataagtat agcttgattc tcgtctgagc 780

gagaatcaag ctggcatata tcgttagaac ttcgtggaca acagcacaaa aaacaagaga 840

caaactgctc gttaataaac tcaagtttcg cattaaagct atttagaaat gggtgaagac 900

cgttttatct tcatctattt ttagtctagc tcataaacct tttggaacag tacgttacaa 960

attaattaca ttagcagccg gggacttagc gctacagata aacgtgaatc tgggataagc 1020

tgaaacgcat ttaacccgag ataaagttag ctcatctagg tatagacaca ctagtccaat 1080

taatttggag gctttcttat tcaggttata catactgaaa ctcttttcct ctctatctgc 1140

caagcaccgc aactaatttg atagaaggca accaggaaat ttattggata aattggcgcg 1200

gtaatttgac gatgtgttag ggcagtgatg cactgcttct tgtgaagaag cgtttaaaga 1260

gtaatttgtc tctcatattt cttgctcaac gcagcaaatg tgatgctagt gcaaaatagc 1320

gtgacgggtg acacttacta aattgagtca gccgagcatc ggatggtgat atggagcgac 1380

agcaaaccag agcatcggtt gccaatccag caatctagct gactaagcag tgcaagtaat 1440

ttgacagcag agcaaaacag cgtgactgtt gacaaggaga aatttcattt ctagatcgtg 1500

ccattaattt gcgaatttag gtgaattttg tgctatcaat ttgcgaacct agaatttgtg 1560

ccattgattt aagaacaaag gatcgtgcag tagaccgcca gatcggttgc aatgcactag 1620

gttcgattcc ttcaatgtgc tttcagatga acaagcaagt atctgtgcca ctgatttgcg 1680

aatcggtcga agatttgtgc cacagattta caaatgacta ttcgacccta acaagccgat 1740

tgtgccagca atttgcgaat cgtgccatat tttttgatta ttcacacgtg cgactttccc 1800

aaaagtcagc caaaaatcgc tcaaaacgac attagcgtgc gacaagcgac atttatcgtg 1860

cgaccgtaca tttgctttta caaaggcggc acccagattc gaactggggg tggaggtttt 1920

gcagacctct gccttaccac ttggcgatgc cgccgaattg catcatggac attgtatca 1979

<210> 133
<211> 1444
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-54 putative transposon end LE

<400> 133
ccgccggagg gtcgagcgac aaactgagca aaagtagtga taaattttgt ctaaaatcaa 60

gcgacctcat tctctgtaca gtgccagatt aaatgacgtg atgccaaatt agtgtcgtcg 120

gttgataccc ttgctacatc aacactctac agcctcataa ccattcttag acctatcact 180

ttagtgccaa attgtgagtc gtttaatcaa ataatgacag attaatcgtc tttaaagaat 240

taatctcagt gacgaattaa ctgtcattat ccgacgatag ccccaagata atatcatgtt 300

gcctctaaca tgaatatgta atggggagaa tatcctaaac cattagtatc ctttcctaaa 360

taaacacttt cattcgcgcg aagagttata gcgccgttat tgccgataaa ctcaagcgcc 420

gttattgccg ataaactcaa tggtgttatt ctacaacggc ggtcaattaa tatttatagt 480

gaaactcctc catttttagt aaaaaatagc aacgtaacaa catagggcta tcgtttatta 540

taaatacgta attcccgaaa attagtggga aggtcacacc ttccatccgt tccatatatc 600

tgagagctga attgatttga attcatcaaa ctccaagtcg catacattct agtagtacag 660

ccatttcttc tatcctgacg tgtataagtt acagaggctt gattataact cttatggcgg 720

cctttaattt gatgatcgaa atctcctgaa tatagttcag cattaatatt aactccatca 780

tcactagaat taataactac tgcgtcctgt tctgaagctt gccacatacc attaaatgct 840

ccaggtgctc tattttgagc aaaagcagaa gacaatccag aaatagctat tacgcaaata 900

gaagaagcaa gtacaagtac gagtttttta ggtctcatgg tttttaattc taaaagctag 960

tcaatttgtg agattgaagg gtgttaaact aaagcagagg tagctacaaa taatgtagaa 1020

gctacgaaat tacgtatgaa actttttca aatccagcgt ggtcattatg gatagtgatt 1080

tttacgccta tagcttctta ttctaaagga atttcctatt aacgatcgcc gctattgtag 1140

cagtggcata attgtcacac taaggattga ctcccaaccg tataattaaa cactatttga 1200

aaggtagggg cagattttga tgtttgctac tccacggtat aagttttctt cggggagagc 1260

taaaacgtcc tttattcgcc gatgcaggcg aaaccataga gttttgaatt acagggtgga 1320

agacccaaac aaaacgtgcc gctcatgttc tgttgtctcc gccagtgcta aatgccaaac 1380

taattgtcgc attatcttag tatgttgcta ggataacaat atgttttatt aaaattttta 1440

ttta 1444

<210> 134
<211> 307
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG108-1 putative transposon end LE

<400> 134
tgtccattaa agaattattt gtcgccttta aagaattaat gtccaatttc aggaattttg 60

gatgtttcag gctttggcaa gaaatctctt gctcaaaatt ttggtttaaa gaattaatgt 120

ccaaaatttc gacatacttg cgttgaaata catcactcat cgttagagtg cgatcgcagc 180

gctcctttca acccatctac attcacacaa aagcactctt accaaagact cttactcttt 240

cagtgcaatc aaccttggaa gggcgattac aaagcagtct tccaaattca gcgtagtatt 300

gacattc 307

<210> 135
<211> 619
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG108-1 putative transposon end RE

<400> 135
ggtgtgagca gttaacaacc atggcttaaa taatgacact ttttggttta aatattcaac 60

cagcgaatag taaaatttat gacaaatttt gctttaaaca aagctatggc tattttggtg 120

aagtgaatca gcctgaagat caatagcagc aagggttgca tcacctagca gtgtttagag 180

ccaagatgac aataattctt taaagtcggt caaacaattc tttaatggac accatgtgtg 240

atgtacacta acgctcatct gtcactctag gcagatgtcc attaaggagc tatttgtcag 300

ttttaaacag ctcttgtcaa aaaattttag tactcaacgt attttcaagg gttttaagcg 360

acacccttg atttgctgct tttttgccct gaaattccag ttttaaggag atgttgtcac 420

agatgattca tgccttgcta ccctgtatcg gtatatggac agagtcatta cagacgatcg 480

cctgtagtgc gatcgcagcg ctcctttcta tccacattaa ttcacatgtt ttaacgctca 540

ccactaatcg actgctgagc ggttgctgta atagcgtcga gacgtatgta ggccgaaatc 600

gatcgcatcc gcaacagcg 619

<210> 136
<211> 709
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG110-1-6 putative transposon end LE

<400> 136
tattgttatc cctatttatt gttattagac aaattgttta tttaaacaac aataaataga 60

taccatcaaa atacaatttg tacagtgata aaatcgaaaa aatagataac aatttgttat 120

atgcagcgcg aaaactggaa tagagaactg agcggcgtgg tttaatatat gagcggcaat 180

agtggaagtt taaggggaa tgtatcgctg acgggtatag cttgttgatt gaagcataaa 240

ttgacaaatt cacaaacgaa aaatcaatga gttacgcttc tctttgaacc ataagatggc 300

aaaaaattaa accataagct gacaaatttc ttagtgaacg tgttacacca cccaaaacgt 360

ggtaagccac tgttgtaaaa ttacaagata tctcatgtac ttcatagtta cgaagtcggt 420

ttagctaagg gcgagtcact cacgaccagt tcaacagtgt tacgggcact gtttaccagc 480

gacctatgac gaagaccttc gatactttca ataatactac tgtagagcat ttcaaaggcc 540

atcgaaatat agaggtgtag ttatatttta agcgtctcag agcaaaaaag ccctctccat 600

agcgctttca cattttcatc cattctcgat ctcgacaatg aaacttaact cccaaatcaa 660

cgagccagcg gagaggagca atttagctag ctccgcgtgc attggcttg 709

<210> 137
<211> 479
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG110-1-6 putative transposon end RE

<400> 137
aaccgagacc tgtaagttgt aagtaaataa caacttaaat ataggttatg ccagctatta 60

ctgcaaattt gacaacttat ggttcaaatt tgtcacttta tgcttcaatt tacatagctt 120

taagttggtg gcatctctgc cagcctcacc ccggcacccg agtctaccgc tgtggctgct 180

tccttccgga cctgaccaga ttcacggttt atcgttgcgg ggggaccagc agagacacca 240

tagaggaggt ctcgttgacc ttggtggggc attatacgca ctctgtttcg atgtgcaacc 300

tacgctatgg ggtttattat ctttttcta ttagtgagaa tactttaacg aaattcgctg 360

gttcgccgct cccgttgctg aagcttttag cgctctcggc tatacttgac taactttatc 420

ggtttttgcc gctgtctctg tatcacagag ccggcgatgt attttatggg agcgcgcag 479

<210> 138
<211> 474
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG110-2-6 putative transposon end LE

<400> 138
tgaggttgcc tccagtacgt ccacaccacg tggatgacca gacaggcttt gaggttaact 60

aattgatgtt gttacaacca taagttgaca taacggccat aataaaatca acaggttaca 120

aaacgatttt tacccataag ttgacataaa aatgaagcat aagctgacat agagtgacaa 180

gaccacacat ggacacaaaa tcggcgttgt aagtttacaa gatcgacttg tagcatccac 240

tataggcaaa cacacctacg gcagccatcc tctcattaag tgcaagtggc actatcgaac 300

acaacgctca ccatttttgc gtagggcaaa gtcgatccgc cccgaatttt ctagaccatc 360

cgaacctttg gctagcaggc tacatgaatt tctccatgtg ccggatatcg agccattgtc 420

tgcataatta taaaacgctc cctattgctt gaaacgactt gcattgtttt gagg 474

<210> 139
<211> 334
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG110-2-6 putative transposon end RE

<400> 139
agcgcctcca acctgtagtt gcttttctac aggtttaggc tatgtcaact tatgatgcaa 60

gtttatgtca gcttttggtt ccactatgtc aacttatggt tccaacgaca attgatttgc 120

cagaaatggt gggggctctg ccagccacat cccggcactc gaaggactcg ccttggctgc 180

taccttccgg tcctgaccag gtcgacgaga ttccgatgcg ggaggaccaa cagagccccc 240

ataacaacgc cggcgagtct acctagcctg accggcaatt gcaatcgctt taccggtcag 300

ctgagttcaa gtgtctgaaa actcgacaag tagc 334

<210> 140
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<220>
<223> Nucleoplasmin NLS

<400> 140
Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

<210> 141
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<220>
<223> SV40 2x NLS

<400> 141
Pro Lys Lys Lys Arg Lys Val Asp Gly Ser Pro Lys Lys Lys Arg Lys
1 5 10 15

Val Asp Ser

<210> 142
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<220>
<223> P2A

<400> 142
Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn

1 5 10 15

Pro Gly Pro

<210> 143
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic peptide

<220>
<223> T2A

<400> 143
Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro
1 5 10 15

Gly Pro

<210> 144
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic peptide

<220>
<223> HA

<400> 144
Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
1 5

<210> 145
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic peptide

<220>
<223> Myc

<400> 145
Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
1   5   10

<210> 146
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      peptide

<220>
<223> FLAG

<400> 146
Asp Tyr Lys Asp Asp Asp Asp Lys
1   5

<210> 147
<211> 640
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
      MG64 effector sequence

<220>
<223> MG64-56 effector

<400> 147
Met Ser Val Ile Thr Ile Gln Cys Arg Leu Val Ala Asp Asp Lys Ala
1   5   10   15

Leu Arg His Leu Trp Glu Leu Met Ala Glu Lys Asn Thr Pro Leu Val
            20   25   30

Asn Glu Leu Leu Asp Arg Leu Gly Lys His Thr Asp Phe Glu Ala Trp
        35   40   45

Val Gln Ala Gly Lys Val Pro Lys Thr Thr Ile Lys Ala Leu Cys Asp
    50   55   60

```
Ser Leu Lys Thr Gln Glu Pro Phe Ile Gly Gln Pro Gly Arg Phe Tyr
65                  70                  75                  80

Thr Ser Ala Thr Thr Leu Val Ala Tyr Ile Tyr Lys Ser Trp Leu Ala
                85                  90                  95

Leu His Lys Arg Arg Gln Arg Lys Ile Glu Gly Lys Glu Arg Trp Leu
            100                 105                 110

Glu Met Leu Lys Ser Asp Val Glu Leu Glu Gln Glu Ser Asn Ser Ser
        115                 120                 125

Leu Glu Leu Ile Arg Thr Ile Ala Thr Glu Ile Leu Ser Lys Phe Ser
    130                 135                 140

Ala Ser Ser Thr Asp Gly Ile Asn Gln Lys Ser Lys Gly Lys Lys Ser
145                 150                 155                 160

Lys Lys Leu Lys Lys Asp Lys Ala Asp Glu Pro Met Ser Ile Lys Pro
                165                 170                 175

Gly Val Leu Phe Glu Ala Tyr Gln Lys Thr Glu Asp Ile Leu Arg Arg
            180                 185                 190

Ser Ala Leu Val Tyr Leu Ile Lys Asn Asn Cys Gln Val Asn Phe Ala
        195                 200                 205

Glu Glu Asp Pro Asp Lys Tyr Ala Lys Met Arg Arg Lys Lys Glu Ile
    210                 215                 220

Glu Ile Glu Arg Leu Lys Glu Gln Leu Lys Ser Arg Val Pro Lys Gly
225                 230                 235                 240

Arg Asp Leu Thr Gly Lys Lys Trp Leu Glu Thr Leu Glu Lys Ala Val
                245                 250                 255

Asn Ser Ile Pro Gln Asp Glu Asn Glu Ala Lys Ser Trp Gln Ala Gly
            260                 265                 270

Leu Leu Arg Lys Ser Ser Thr Val Pro Phe Pro Val Ala Tyr Glu Thr
        275                 280                 285
```

```
Asn Glu Asp Met His Trp Glu Ile Ser Asp Lys Gly Arg Ile Phe Val
    290                 295                 300

Ser Phe Asn Gly Leu Ser Lys Leu Lys Leu Glu Val Tyr Cys Asp Gln
305                 310                 315                 320

Arg His Leu Pro Trp Phe Gln Arg Phe Val Glu Asp Gln Glu Thr Lys
                325                 330                 335

Arg Lys Gly Lys Asn Gln His Ser Ser Gly Leu Phe Thr Leu Arg Ser
            340                 345                 350

Gly Arg Leu Ser Trp Leu Lys Gln Glu Gly Lys Gly Glu Pro Trp Ser
        355                 360                 365

Val Asn Arg Leu Ile Leu Phe Cys Ser Val Asp Thr Arg Met Trp Thr
    370                 375                 380

Val Glu Gly Thr Gln Gln Val Ala Ile Glu Lys Ile Ala Asp Val Glu
385                 390                 395                 400

Gln Asn Leu Thr Lys Ala Lys Glu Lys Gly Glu Leu Asn Ser Asn Gln
                405                 410                 415

Gln Ala Phe Val Thr Arg Gln Gln Ser Thr Leu Ala Lys Ile Asn Thr
            420                 425                 430

Pro Phe Pro Arg Pro Ser Lys Pro Leu Tyr Glu Gly Lys Ser His Ile
        435                 440                 445

Leu Val Gly Val Ser Leu Gly Leu Glu Asn Pro Ala Thr Val Ala Val
    450                 455                 460

Phe Asp Ala Val Asn Asn Lys Val Leu Ala Tyr Arg Ser Val Lys Gln
465                 470                 475                 480

Leu Leu Gly Asn Asn Tyr Asn Leu Leu Asn Arg Gln Gln Gln Gln Lys
                485                 490                 495
```

Gln Arg Leu Ser His Asp Arg His Lys Ala Gln Lys Asp Phe Ala Arg
500 505 510

Asn Asp Phe Gly Glu Ser Glu Leu Gly Gln Tyr Val Asp Arg Leu Leu
515 520 525

Ala Lys Glu Ile Val Ala Ile Ala Val Thr Tyr Phe Ala Gly Ser Ile
530 535 540

Val Leu Pro Lys Leu Gly Asp Met Arg Glu Ile Ile Gln Ser Glu Val
545 550 555 560

Gln Ala Arg Ala Glu Lys Lys Ile Pro Gly Phe Lys Glu Gly Gln Gln
565 570 575

Lys Tyr Ala Lys Glu Tyr Arg Lys Gln Val His Asn Trp Ser Tyr Gly
580 585 590

Arg Leu Ile Glu Asn Ile Gln Ser Gln Ala Ala Lys Val Gly Ile Leu
595 600 605

Ile Glu Thr Gly Gln Gln Pro Ile Arg Gly Ser Pro Gln Glu Gln Ala
610 615 620

Arg Asp Leu Ala Leu Phe Ala Tyr Gln Cys Arg Ile Ala Ser Ser Ile
625 630 635 640

<210> 148
<211> 579
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-56-B transposition protein

<400> 148
Met Gly Glu Thr Leu Asn Ser Asn Glu Val Asp Glu Ser Leu Val Leu
1 5 10 15

Tyr Asp Gly Ser Asp Glu Val Asp Glu Ile Ser Glu Ser Glu Asp Thr

```
            20                  25                  30

Lys Gln Asn Asn Val Ile Val Thr Glu Leu Ser Glu Glu Ala Lys Leu
        35                  40                  45

Arg Met Glu Val Leu Gln Ser Leu Ile Glu Pro Cys Asp Arg Lys Thr
    50                  55                  60

Tyr Gly Ile Lys Leu Lys Gln Ala Ala Glu Lys Leu Gly Lys Thr Val
65                  70                  75                  80

Arg Thr Val Gln Arg Leu Val Lys Lys Tyr Gln Glu Gln Gly Leu Ser
                85                  90                  95

Gly Val Thr Glu Val Glu Arg Ser Asp Lys Gly Gly Tyr Arg Ile Asp
            100                 105                 110

Asp Asp Trp Gln Asp Phe Ile Val Lys Thr Tyr Lys Glu Gly Asn Lys
        115                 120                 125

Gly Gly Arg Lys Met Thr Pro Ala Gln Val Ala Ile Arg Val Gln Val
    130                 135                 140

Arg Ala Gly Gln Leu Gly Leu Glu Lys Tyr Pro Cys His Met Thr Val
145                 150                 155                 160

Tyr Arg Val Leu Asn Pro Ile Ile Glu Arg Lys Glu Gln Lys Gln Lys
                165                 170                 175

Val Arg Asn Ile Gly Trp Arg Gly Ser Arg Val Ser His Gln Thr Arg
            180                 185                 190

Asp Gly Gln Thr Leu Asp Val His His Ser Asn His Val Trp Gln Cys
        195                 200                 205

Asp His Thr Lys Leu Asp Val Met Leu Val Asp Gln Tyr Gly Glu Thr
    210                 215                 220

Leu Ala Arg Pro Trp Leu Thr Lys Ile Thr Asp Ser Tyr Ser Arg Cys
225                 230                 235                 240
```

```
Ile Met Gly Ile His Leu Gly Phe Asp Ala Pro Ser Ser Leu Val Val
                245                 250                 255

Ala Leu Ala Met Arg His Ala Met Leu Arg Lys Gln Tyr Ser Ser Glu
            260                 265                 270

Tyr Lys Leu His Cys Glu Trp Gly Thr Tyr Gly Val Pro Glu Asn Leu
        275                 280                 285

Phe Thr Asp Gly Gly Lys Asp Phe Arg Ser Glu His Leu Lys Gln Ile
    290                 295                 300

Gly Phe Gln Leu Gly Phe Glu Cys His Leu Arg Asp Arg Pro Pro Glu
305                 310                 315                 320

Gly Gly Ile Glu Glu Arg Gly Phe Gly Thr Ile Asn Thr Asp Phe Leu
                325                 330                 335

Ser Gly Phe Tyr Gly Tyr Leu Gly Ser Asn Val Gln Glu Arg Ala Glu
            340                 345                 350

Gly Ala Glu Glu Glu Ala Cys Ile Thr Leu Arg Glu Leu His Leu Leu
        355                 360                 365

Ile Val Arg Tyr Ile Val Asp Asn Tyr Asn Gln Arg Ile Asp Ala Arg
    370                 375                 380

Ser Gly Asn Gln Thr Arg Phe Gln Arg Trp Glu Ala Gly Leu Pro Ala
385                 390                 395                 400

Leu Pro Asn Leu Val Asn Glu Arg Glu Leu Asp Ile Cys Leu Met Lys
                405                 410                 415

Lys Thr Arg Arg Ser Ile Tyr Lys Gly Gly Tyr Val Ser Phe Glu Asn
            420                 425                 430

Ile Met Tyr Arg Gly Asp Tyr Leu Ser Ala Tyr Ala Gly Glu Ser Val
        435                 440                 445

Leu Leu Arg Tyr Asp Pro Arg Asp Ile Ser Thr Val Phe Val Tyr Arg
```

```
    450                 455                 460

Gln Asp Ser Gly Lys Glu Val Leu Leu Ser Gln Ala His Ala Ile Asp
465                 470                 475                 480

Leu Glu Thr Glu Gln Ile Ser Leu Glu Glu Thr Lys Ala Ala Ser Arg
                485                 490                 495

Lys Ile Arg Asn Ala Gly Lys Gln Leu Ser Asn Lys Ser Ile Leu Ala
            500                 505                 510

Glu Val Gln Asp Arg Asp Thr Phe Ile Lys Gln Lys Lys Lys Ser His
        515                 520                 525

Lys Glu Arg Lys Lys Glu Glu Gln Ala Gln Val Asn Phe Val Lys Pro
    530                 535                 540

Pro Gln Thr Asn Glu Pro Val Glu Thr Val Glu Glu Ile Pro Gln Pro
545                 550                 555                 560

Gln Lys Arg Arg Pro Arg Val Phe Asp Tyr Glu Gln Leu Arg Lys Asp
                565                 570                 575

Tyr Asp Asp
```

<210> 149
<211> 278
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-56-C transposition protein

<400> 149

```
Met Ala Glu Asp Tyr Leu Arg Lys Trp Val Gln Asn Leu Trp Gly Asp
1               5                   10                  15

Asp Pro Ile Pro Glu Glu Leu Leu Pro Ile Ile Glu Arg Leu Ile Thr
            20                  25                  30
```

```
Pro Ser Val Val Glu Leu Glu His Ile Gln Lys Ile His Asp Trp Leu
        35                  40                  45

Asp Ser Leu Arg Leu Ser Lys Gln Cys Gly Arg Ile Val Ala Pro Pro
    50                  55                  60

Arg Ala Gly Lys Ser Val Thr Cys Asp Val Tyr Lys Leu Leu Asn Lys
65                  70                  75                  80

Pro Gln Lys Arg Thr Gly Lys Arg Asp Ile Val Pro Val Leu Tyr Met
                85                  90                  95

Gln Val Pro Gly Glu Cys Ser Ala Gly Glu Leu Leu Thr Leu Ile Leu
            100                 105                 110

Glu Ser Leu Lys Tyr Asp Ala Ile Ser Gly Lys Leu Thr Asp Leu Arg
        115                 120                 125

Arg Arg Val Leu Arg Leu Leu Lys Glu Ser Lys Val Glu Met Leu Val
    130                 135                 140

Ile Asp Glu Ala Asn Phe Leu Lys Leu Asn Thr Phe Ser Glu Ile Ala
145                 150                 155                 160

Arg Ile Tyr Asp Leu Leu Lys Ile Ser Ile Val Leu Val Gly Thr Asp
                165                 170                 175

Gly Leu Asp Asn Leu Ile Lys Lys Glu Pro Tyr Ile His Asp Arg Phe
            180                 185                 190

Ile Glu Cys Tyr Arg Leu Pro Leu Val Ser Glu Lys Lys Phe Pro Glu
        195                 200                 205

Phe Val Gln Ile Trp Glu Asp Glu Val Leu Cys Leu Pro Val Pro Ser
    210                 215                 220

Asn Leu Thr Lys Arg Glu Thr Leu Met Pro Leu Tyr Gln Lys Thr Ser
225                 230                 235                 240
```

Gly Lys Ile Gly Leu Val Asp Arg Val Leu Arg Arg Ala Ala Ile Leu
245 250 255

Ser Leu Arg Lys Gly Leu Lys Asn Ile Asp Lys Ala Thr Leu Asp Glu
260 265 270

Val Leu Glu Trp Phe Glu
275

<210> 150
<211> 176
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
MG64 transposition protein sequence

<220>
<223> MG64-56-Q transposition protein

<400> 150
Met Glu Ile Pro Ala Glu Gln Pro Arg Phe Phe Gln Val Glu Pro Leu
1 5 10 15

Glu Gly Glu Ser Leu Ser His Phe Leu Gly Arg Phe Arg Arg Glu Asn
20 25 30

Tyr Leu Thr Ala Thr Gln Leu Gly Lys Leu Thr Gly Ile Gly Ala Val
35 40 45

Ile Ser Arg Trp Glu Lys Phe Tyr Leu Asn Pro Phe Pro Thr Pro Gln
50 55 60

Glu Leu Glu Ala Leu Ala Ala Val Val Glu Val Lys Val Asp Arg Leu
65 70 75 80

Ile Glu Met Leu Pro Pro Arg Gly Val Thr Met Lys Pro Arg Pro Ile
85 90 95

Arg Leu Cys Ser Ala Cys Tyr Gln Glu Ser Pro Cys His Arg Val Glu
100 105 110

Trp Gln Phe Lys Asp Val Met Val Cys Asp Cys Leu Arg His Cys Pro

115 120 125

Leu Asn Asn Arg His Gln Leu Ala Leu Leu Thr Lys Cys Thr Asn Cys
130 135 140

Glu Thr Pro Phe Pro Ile Pro Ala Asp Trp Val Gln Gly Glu Cys Pro
145 150 155 160

His Cys Phe Leu Pro Phe Thr Lys Met Ala Arg Arg Gln Lys Arg Tyr
165 170 175

<210> 151
<211> 485
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<220>
<223> MG64-56 effector intergenic region encoding tracrRNA

<400> 151
caaaataccg aaccttgaaa acttaatatg aaagtaacag cgccgcagtt catgctcttc 60

tgagtctctg tactgtgata aatctgggtt agtttaacgg ttgaaagacc gttttgcttt 120

ctgaccctgg tagctgctcg ctcttgatgc tgctgtcttt tgacaggata ggtgcgctcc 180

cagcaataaa gagttaaagc tgataaagct tgagccgttg taaaacggtg gggtttacct 240

cagtggtggc tactgaatca ccccccttcgt cgggggaacc ctcctaaata ttttttttgg 300

cgtgtcaaag cgggggcaaa aatcctggag tcccgccaaa atctcaaaac ctttgtccta 360

tcttgacttg ataaactagc atgtcagtta atttagtttt ttgatgtcaa gtaggagatg 420

cttttaggca gtcctgccaa agatgtgtat ggaaagctct aatagcaagg gttctagacg 480

gatcg 485

<210> 152
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

oligonucleotide

<220>
<223> MG64-56 effector target CRISPR repeat

<400> 152
gtttcaacaa ccatcccagc taggggtggg ttgaaag 37

<210> 153
<211> 595
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-56 putative transposon end LE

<400> 153
tatgcgactt gagcgatttg tgtaagtttt gtgttgcaca aaccatccta agcgacatta 60

gtttgcaaaa aacgacatta atttacgaat cgcgaccttt aatttgcgaa tatacaacag 120

attttgtcga ttaactaatt atttgtcgtc ttaacaaatt aatgtcgccc aaatcttcaa 180

gactataatc cttatgtatc aaaggttata gccttttgaa cttatattgg ctatcatcaa 240

atatttaact aattaagtgt cgtcttttaa ttaattaaca ttttaaatgt cgttttttca 300

aaaaacacct ttccaaattt ttcttttgct cataacaaaa taactgtcgt cttttggaag 360

tgagtgaaaa atataaaatt aaatgtcgct ttttggaaca aagtagtatg atatttatta 420

ggcaatagta gctatgtaac aacaaaaaca tagttagatt gaagtcttct tttttgtctc 480

tagctacgaa gtcattaccc ttgctgcgat taaatttaga cgcaagctaa tttcgctctt 540

agacttgctg taccgtattg cctaaccaac tagtttcaag cgatgaagtt tgttt 595

<210> 154
<211> 493
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-56 putative transposon end RE

<400> 154
taagttctat taactccaag ttttaataat tgcatggcaa taacaatcct ttttagaaag 60

gatttaagag ggttgaaagg aatgtcacct tcccaagaat acttttcaaa agctattttg 120

ggttagggaa gaataatcac agataactaa tatgcacaag taagtctaaa atagggataa 180

gtctgtcgat tagtccaata gcaaggcatc ttgttagacg acattaattt gttaacgtta 240

gttggaacta attcgacgac attaattcgt taacagcgac attaatttgt taatgacgac 300

attaatctgt taacgacgac attaatctgt taacgacgac aaataatctg ttaattgaca 360

gatttgaaag cgggtgatgg gactcgaacc cacgacgttc accttgggaa ggtgacattc 420

taccactgaa ttacacccgc aaatggagtt taggctcaat aaagctaacc cccattataa 480

cacgatttgc cca 493

<210> 155
<211> 568
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-2 active transposon end RE

<400> 155
cctcgtgata cgcctatttt tataggttaa tgtcatgata ataatggttt cttagacgtc 60

aggtggcact tttcggggaa atgtgttcta tctacaaaat tgttttatac aactatgttt 120

gtattttagt tctcaggaca aagattgtgt caattcttca caaaagttgt aattctcgcc 180

tttttcaaaa aaggacaaag aatttgtcaa ctttcctaaa ccggacaact aatttgtcaa 240

aaagcttttt caatatataa agggggtata atagctgtaa cccttgcgta gagtaggtta 300

gtgagctact aggtttggcg gtcaataatt tggcaagtgg tcaaataatt tggcaatcga 360

caattattta tcgttgggga cgaatcagtg tcacggagtc aaattatgtg tcgttgtgac 420

agattgatgt cgtctaagca ccagtgacaa attgatgttg ctttaataat ttaaaatcgt 480

cacgaaataa gcctaacgtg agttttcgtt ccactgagcg tcagaccacc gttgatgata 540

ccgctgcctt actgggtgca ttagccag 568

<210> 156
<211> 227
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-4 effector putative tracrRNA

<400> 156
aatagcgccg cttgttcatg cgcgaaagcg tctctgaaca gtgtaaatgt gggttagttt 60

gactgtcgtg aagacggtct tgctttctga ccctggtagc tgcccacctt gaagctgctg 120

tctcttgtag acaggaatca ggtgcgcccc cagtaatata ggtgcgggtt taccgcagtg 180

gtggctaccc aatcacctcc gagcaaggag gaacccacct taattat 227

<210> 157
<211> 228
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-7 effector putative tracrRNA

<400> 157
aacagcgccg cagttcatgc gtcttatggc gcctctgtgc tgtgcaaaat gtgggttagt 60

ttgactgttg gaagacagtc ttgctttctg accctggtag ctgcccacct tgaagctgct 120

atcccttgtg gataggaatc aggtgcgccc ccagtaatag aggtgcgggt ttaccgcagt 180

ggtggctacc gaatcacctc cgagcaagga ggaacccacc ttaattat 228

<210> 158
<211> 328
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-9 effector putative tracrRNA

<400> 158
gctgaagtta gtctctggat actgaagtca gtgtctaggg actgaagtaa ggggagcttt 60

acccgcaagg gatactttcg acccccctgta gctgcccgct cctggtgggg tgccctgaca 120

cccgcctcat tacagcaatg tatgactgtc tgggctaatg aaagaaggat tagggacgca 180

ggttcacgac ctacttcaat catagtctgt gcaacccaga taagtgagta tgacccgcaa 240

gggtctaaac gcctttagta gaatggttct tccagtttgt taagggcagg gcttggttct 300

ccaaggtggc tacgaattct tctcgatt 328

<210> 159
<211> 187
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-10 effector putative tracrRNA

<400> 159
ggtagcgctg gagccacata gttcataagc tcacgcttct tggacttcct gtgttctcta 60

aaacgggttc tgttttaccc ttaccaaggg atactttcag atccgagtag ctgcaagctc 120

atggcggagt gtcccctgac gctttgccac cgtcatagcg atgtgatggc cgtctggcgt 180

atgaacg 187

<210> 160
<211> 326
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-12 effector putative tracrRNA

<400> 160
aatagcgccc ctaatcaaat cgaaggcgtt tgaggacggg gaaaatgggt aagtttccat 60

cgaaaggtgg ttcttttcag ccctatgtag tcccactccc ttgtggagtg ttagcatctg 120

aggtgcctag cacagcaaag tctctcagga gacggagtca aagctgggga acaagtatta 180

cgaacgaatc tcgcgattct attagtatac gtaggtcgct cccatgcaac aagatgctca 240

tctctcagca atgagggtag ggaacgtatc ccaaacacta catgaatagc tcgtcgggct 300

tgcccggtga gtatccacca agcctt 326

<210> 161
<211> 311
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-13 effector putative tracrRNA

<400> 161
cgctaggttt gccttttcgc gcgctgctga agcagtgaaa agttccgaca gtttaaaagt 60

tctaacactt ttactgcttt ccgtcgggag tagttgtccg cttctgcgtt gttgagaccg 120

aagtttgtcc aggtgatagg cagaacgttg atgcaacagg tagatagtcg cgctctatca 180

ggaagctgtt ctagttatta gtgttcgtgc actaataaag aggatacagg gatacatgtg 240

gttgtgtcca gcaatggcac aacagggcca ctactcgaag ccccgagcaa gggttgagcc 300

tacccaaatt t 311

<210> 162
<211> 342
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-14 effector putative tracrRNA

<400> 162
aacagcgccg cagtttaagc tctatagccg ctgaactgtg aaaaatgtgg gtcagtttgg 60

tcgttgcaag acgatcgtgc tttccgaccc tagtaactgt ccgctcactg actgccatcc 120

tggggcaaat cttcaaattt tgtgtatttg tgtggggatg gaaagctgca ttagtcgatt 180

ctcttcctcc aatgtagcgt aggtgcgcac ccagcagaag tgagttaagc cttcacaatg 240

tggaggtaca gaagcatcat ctctccattt tttggtgtag atggtgtgac tgaagtggta 300

gttaccgaat cgcccctgat caaggggggag ccctccataa tt 342

<210> 163
<211> 297
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-15 effector putative tracrRNA

<400> 163
aatgacgttt gtttctgtgc gcgctgctga agcagtgaaa aatacggaca gtctaatttg 60

ctttccgtcc ggagtagttg tccgcttctg cgttgttgga accgaagttt gtccaggtga 120

taggcggagt gtttcggcaa cagatggaga gtcgcacctc tatcaggggg cagttttagt 180

ggttagtgtt tagcgcacta tcactaggat acagggatac atgtggtagt gaccagagat 240

ggcactgcgg gaccgctact cgaagccccg aacaagggtt gagcttaccc aaatctt 297

<210> 164
<211> 226
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-16 effector putative tracrRNA

<400> 164
aatagcgccg cagtttaagc tcagcaagcc tctggactgc gaaaagtatg gggtagtttg 60

accgtcggta aacggttgtg ctttctgccc ctggcgactg cccaccccga tgctgtcgat 120

ttcttaactg ggaatcgaga tgaggtgcgc ccccagcaaa agggaacggg tttactggag 180

tggtggtcgc cgaatcaccc ccgagcaagg gggactcgtc ctttgc 226

<210> 165
<211> 227
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-17 effector putative tracrRNA

<400> 165
aatagcgccg cagtttaagc tcaatgagcc tcttgactgc gaacagtatg gggtagtttg 60

accggcggta accggttgtg ctttctgccc ctggcgactg cccaccccga tgctgtcgat 120

ttctcaaccg ggaatcgaga atcaggtgcg cccccagcaa gagggaacgg gtttactgga 180

gtggtggtcg ccgaatcacc cccgagcaag ggggactcgt cctttgc 227

<210> 166
<211> 227
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-19 effector putative tracrRNA

<400> 166
aatagcgccg cagtttaagc tcagcaagcc tctggactgc gaaaagtatg gggtagtttg 60

accgtcggta aacggttgtg ctttctgccc ctggcgactg cccaccccga tgctgtcgat 120

ttctcaaccg ggaatcgaga atcaggtgcg cccccagcaa gagggaacgg gtttactgga 180

gtggtggtcg ccgaatcacc cccgagcaag ggggactcgt cctttgc 227

<210> 167
<211> 211
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-20 effector putative tracrRNA

<400> 167
aatagcgccg cagtttaagc tcaatgagcc tcttgactgc gaacagtatg gggtagtttg 60

accggcggta accggttgtg ctttctgccc ctggcgactg cccaccccga tgctgtcgat 120

ttctcaaccg ggaatcgaga atcaggtgcg cccccagcaa gagggaacgg gtttactgga 180

gtggtggtcg ccgaatcacc cccgagcaag g 211

<210> 168
<211> 222
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-21 effector putative tracrRNA

<400> 168
aatagcgccg cagttcatgc ttctttgaag cctctgtgct gtgcaaaatg tgggttagtt 60

tggctgttga agaaacagcc ttgctttctg accctggtag ctgtccaccc tgaagctgct 120

atcccctgtg gataggatag gtgcgccccc agcaataggg gagcgggtat accgcagtgg 180

tggctactga atcacctcca agcaaggagg aatccacttt at 222

<210> 169
<211> 222
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-22 effector putative tracrRNA

<400> 169
aatagcgccg cagttcatgc ttctttgaag cctctgtgct gtgcaaaatg tgggttagtt 60

tggctgttga agaaacagcc ttgctttctg accctggtag ctgtccaccc tgaagctgct 120

atcccctgtg gataggatag gtgcgccccc agcaataggg gagcgggtat accgcagtgg 180

tggctactga atcacctcca agcaaggagg aatccacttt at 222

<210> 170
<211> 248
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-24 effector putative tracrRNA

<400> 170
aaaagcgccg tagaacatgc tcacgcctct gttctgcgaa aaattagggt ttgtttggct 60

gtctgacagc agtcttactt tctgtcccta gaatctgacc actccgatgc tgctgttgta 120

agtgaacttg attgagctcg gacaccatgc aacaggataa ggggcgcacc cagcaagaga 180

ggacggactt accgtagtgt tggcttctga agcaactccg accaaggagt agtccatgca 240

ttcatcat 248

<210> 171
<211> 277
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-25 effector putative tracrRNA

<400> 171
aatcgcgccg cacattcatg ttccttgaga acctctgaat tgcgaaagtg tgggctagtt 60

tgttcgcttg atgcgaatgt gctttctggc cctggtagct gtccgccctg atgctgattt 120

ctacgggtaa ctgtaggaat gattaactcg ttctatagac aggtttcgtg ctttctattg 180

taacggggtc ggtgcgctcc cagcaatagg ggtgtgggtc tactacagtg atggctactg 240

aatcacctcc gagcaaggag gaatccacct taacttt 277

<210> 172
<211> 226
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-27 effector putative tracrRNA

<400> 172
aatagcgccg cagtttaagc tcagcaagcc tctggactgc gaaaagtatg gggtagtttg 60

accgtcggta aacggttgtg ctttctgccc ctggcgactg cccaccccga tgctgtcgat 120

ttcttaactg ggaatcgaga tgaggtgcgc ccccagcaag agggaacggg tttactggag 180

tggtggtcgc cgaatcaccc ccgagcaagg gggactcgtc ctttgc 226

<210> 173
<211> 226
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-28 effector putative tracrRNA

<400> 173
aatagcgccg cagtttaagc tcagcaagcc tctggactgc gaaaagtatg gggtagtttg 60

accgtcggta aacggttgtg ctttctgccc ctggcgactg cccaccccga tgctgtcgat 120

ttcttaactg ggaatcgaga tgaggtgcgc ccccagcaag agggaacggg tttactggag 180

tggtggtcgc cgaatcaccc ccgagcaagg gggactcgtc ctttgc 226

<210> 174
<211> 255
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-29 effector putative tracrRNA

<400> 174

tagcgcacga acctgactag cgtcattacg atgcgagttc tggaaatggg acagtttcat 60

tgctttccgt ccctggcact gcctgcttac tcacgaccac aaggagacaa ttcagcgctg 120

tgagttcgcg attctctact gtttatgaca aactaacgtt tgtctggcag tacagcaaga 180

ttgtgtatca gaccatgttt ggcgaagata catggcaggc cgaatcgcca atgaaaccag 240

cggacttccc ctaat 255

<210> 175
<211> 226
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-32 effector putative tracrRNA

<400> 175
aatagcgccg cagtttaagc tcaatgagcc tcttgactgc gaacagtatg gggtagtttg 60

accggcggta accggttgtg ctttctgccc ctggcgactg cccaccccga tgctgtcgat 120

ttcttaactg ggaatcgaga tgaggtgcgc ccccagcaag agggaacggg tttactggag 180

tggtggtcgc cgaatcaccc ccgagcaagg gggactcgtc ctttgc 226

<210> 176
<211> 203
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-44 effector putative tracrRNA

<400> 176
tctagcgccg cagctcatgt cagcaatggc caatgtgttg tgctaaatgc gagctagttt 60

gactgcctgc taagcagtct tgctttctgg ctcaggtgac tatccaccca aaggtcgttg 120

gtgcgctggc gatttgaggg cacgggttcc ggagtgatag ttaccattac acctccggcc 180

aaggaggaat ccaccccacc ccc 203

<210> 177
<211> 203
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-46 effector putative tracrRNA

<400> 177
tctagcgccg cagctcatgt cagcaatggc caatgtgttg tgctaaatgc gagctagttt 60

gactgcctgc taagcagtct tgctttctgg ctcaggtgac tatccaccca aaggtcgttg 120

gtgcgctggc gatttgaggg cacgggttcc ggagtgatag ttaccattac acctccggcc 180

aaggaggaat ccaccccacc ccc 203

<210> 178
<211> 310
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-49 effector putative tracrRNA

<400> 178
gattgcgcct cgatcgatgc tctatgagcc gctcgatcgt agaaaaatgg gtgagtttga 60

ttatctactt cgttagataa tgctgctttc cgaccctggc attctgtccg cccttgaagc 120

tgcttctcat ggactagcgt aagctcgttg gtaagaagga aaagtcataa tttaaagtca 180

cgtctttcta gtatgacata ggtgcgctcc cacgcaatat agggttcagc ttttatttta 240

taaaagtaga gactttcctc tagtgacagt gccgaaatga ccccgtgcga ggggtaacta 300

cctaagtttt 310

<210> 179
<211> 227
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-51 effector putative tracrRNA

<400> 179
aacagcgccg cagttcatgt ttgttataaa cctctgtact gcgataaatg cgggttagtt 60

tgactgttgt gagacagtct tgctttctga ccctagtagc tgcccacctt gatgctgctg 120

ttcccagtga acaggaataa ggtgcgcccc cagtaataga ggtgcgggtt taccgcagtg 180

gtggctactg aatcacctcc gactaaggag gaatccacct taattat 227

<210> 180
<211> 380
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-52 effector putative tracrRNA

<400> 180
accggggatc gtcctggtga agggagagtt gctttggatc ggtcacaagc ttttccttaa 60

ctaattctca ctgactgact aggatgatcg agggggttat gttttaccac tgcaaggtgg 120

atactttcaa acccctgtgg tagctgctcg ctcctggtga ggtgccctga cacttcaccc 180

cactacagca atgtgtgtgg ctgtctggtt atgagagaag ttagggcata ggttcgttac 240

ctgcattcaa tcataaatta cgcgaccaga taagtgagta tgatccgcaa ggatctatat 300

gtctttagca aagaagtgct tctgctttgt tactggcgta gggcatggtt ctctaaagtg 360

gctaccgaac cttcccaatt 380

<210> 181
<211> 257
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64_56 effector putative tracrRNA

<400> 181
aacagcgccg cagttcatgc tcttctgagt ctctgtactg tgataaatct gggttagttt 60

aacggttgaa agaccgtttt gctttctgac cctggtagct gctcgctctt gatgctgctg 120

tcttttgaca ggataggtgc gctcccagca ataaagagtt aaagctgata aagcttgagc 180

cgttgtaaaa cggtggggtt tacctcagtg gtggctactg aatcaccccc ttcgtcgggg 240

gaaccctcct aaatatt 257

<210> 182
<211> 246
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG108-1 effector putative tracrRNA

<400> 182
tttagcgcac ttgttcgagt tcgtttgaac tgaacaaggg taagtatggg ccagtttaat 60

tgctttccgt cccaggatag ctgccagctt ctaccgtagg ttcgtcctgc aagtgatgct 120

aagtcgcgcc tagcatcaag gagctatgtc ttgattgtct tgggtgtccg ccctggatga 180

gttgaggtgt agatgcttct atcatggcag ctactaaacg ccccaagcaa ggggaaccca 240

tcttta 246

<210> 183
<211> 223
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG108-2 effector putative tracrRNA

<400> 183
aatagcgccg tagttcatgc ttgctaaagc ctctgaattg cgaaaagtcc gggttagtgc 60

tgtcggcaga cagcgttgct ttctgaccct ggtagctgcc caccccgatg ctgctgtccc 120

ttgcagacag gaaccaggtg cgcccccagt aataagggtg tgggtttacc acagtggtgg 180

ctactgaatc acctccgagc aaggaggaat ccaccttaac tat 223

<210> 184
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-9 effector target CRISPR repeat

<400> 184
gtctttcatc ctatctcgcg ccagatcgct tcctgcaacc c 41

<210> 185
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-12 effector target CRISPR repeat

<400> 185
gttgcaagcg cctccttggc tgttggtggg tggaaag 37

<210> 186
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-14 effector target CRISPR repeat

<400> 186
gttgcaatcg ccttcccaga gatgggtggg ctgaaag 37

<210> 187
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-15 effector target CRISPR repeat

<400> 187
gttacaatta ccctcccagc gttgggtggg ttgaaagg 38

<210> 188
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-16 effector target CRISPR repeat

<400> 188
aagttgcatc cgctttccag caaccagggc gggtgaaag 39

<210> 189
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-17 effector target CRISPR repeat

<400> 189
gttgcatccg ctttccagca accagggcgg gtgaaag 37

<210> 190
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-19 effector target CRISPR repeat

<400> 190
gtacccaaag ccttttttcc ttaagcctat ccg 33

<210> 191
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-25 effector target CRISPR repeat

<400> 191
gtttcaaccg ccatcccagc taggggt 27

<210> 192
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-27 effector target CRISPR repeat

<400> 192
agttgcatcc gctttccagc aaccagggcg ggtgaaag 38

<210> 193
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-28 effector target CRISPR repeat

<400> 193
gttgcatctg cttttcagca actagggcgg gggaaagc 38

<210> 194
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-29 effector target CRISPR repeat

<400> 194
ggcgcgatcg cctttatggg tacgggcaag ttgaaag 37

<210> 195
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-32 effector target CRISPR repeat

<400> 195
aagttgcatc cgctttccag caaccagggc gggtgaaagt t 41

<210> 196
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-44 effector target CRISPR repeat

<400> 196
gttgcctccc gcttcgaggc acgggaacga ttgaaag 37

<210> 197

<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-46 effector target CRISPR repeat

<400> 197
gttgcctccc gcttcgaggc acgggaacga ttgaaag 37

<210> 198
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-49 effector target CRISPR repeat

<400> 198
gttgcaacac tccctgactg cctgacacaa atgcctcgaa agc 43

<210> 199
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-52 effector target CRISPR repeat

<400> 199
gtcgcaatga ctattttggc ttggggcgga atga 34

<210> 200
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

oligonucleotide

<220>
<223> MG64-51 effector target CRISPR repeat

<400> 200
gtttcaacac ccctcccgaa gtggggcggg ttgaaag 37

<210> 201
<211> 261
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-6 active effector sgRNA

<220>
<221> modified_base
<222> (239)..(261)
<223> a, c, u, g, unknown or other

<400> 201
auaacagcgc cgcaggucau gccgucaaaa gccucugaac uguguuaaau ggggguuagu 60

uugacuguug aaagacaguu gugcuuucug acccugguag cugcccaccc ugaugcugcu 120

aucuuucggg auaggaauaa ggugcgcucc caguaauagg gguguagaug uacuacagug 180

guggcuacua aaucaccucc gaccaaggag gaauccaucc gaaaggaugg guugaaagnn 240

nnnnnnnnnn nnnnnnnnnn n 261

<210> 202
<211> 263
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-2 effector sgRNA

<220>
<221> modified_base

<222> (241)..(263)
<223> a, c, u, g, unknown or other

<400> 202
gaauuaauag cgccgccguu caugcuucua ggagccucug aaaggugaca aaugcggguu 60

aguuuggcug uugucagaca gucuugcuuu cugacccugg uagcugccca ccccgaagcu 120

gcuguuccuu gugaacagga auuaggugcg cccccaguaa uaaggguaug gguuuaccac 180

agugguggcu acugaaucac cuccgagcaa ggaggaaccc acugaaaggu ggguugaaag 240

nnnnnnnnnn nnnnnnnnnn nnn 263

<210> 203
<211> 265
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-4 effector sgRNA

<220>
<221> modified_base
<222> (243)..(265)
<223> a, c, u, g, unknown or other

<400> 203
gaaauagcgc cgcuuguuca ugcgcgaaag cgucucugaa caguguaaau gugggguuagu 60

uugacugucg ugaagacggu cuugcuuucu gacccuggua gcugcccacc uugaagcugc 120

ugucucuugu agacaggaau caggugcgcc cccaguaaua uaggugcggg uuuaccgcag 180

ugguggcuac ccaaucaccu ccgagcaagg aggaacccac cuugaaagag gcgggguugaa 240

agnnnnnnnn nnnnnnnnnn nnnnn 265

<210> 204
<211> 276
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-7 effector sgRNA

<220>
<221> modified_base
<222> (254)..(276)
<223> a, c, u, g, unknown or other

<400> 204
guaacaaaca gcgccgcagu ucaugcgucu uauggcgccu cugugcugug caaaaugugg 60

guuaguuuga cuguuggaag acagucuugc uuucugaccc ugguagcugc ccaccuugaa 120

gcugcuaucc cuuguggaua ggaaucaggu gcgccccag uaauagaggu gcggguuuac 180

cgcaguggug gcuaccgaau caccuccgag caaggaggaa cccaccuuga aaaagcuuug 240

ggcggguuga aagnnnnnnn nnnnnnnnnn nnnnnn 276

<210> 205
<211> 246
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG108-1 effector sgRNA 1

<220>
<221> modified_base
<222> (224)..(246)
<223> a, c, u, g, unknown or other

<400> 205
gaugcuaagu cgcgccuagc aucaaggagc uaugucuuga uugucuuggg uguccgcccu 60

ggaugaguug agguguagau gcuucuauca uggcagcuac uaaacgcccc aagcaagggg 120

aacccaucuu uaauuuuggc aaaccgaagc gggggcaaaa ucuccaggag guucgccaaa 180

accuuugaaa cuccuuagcu ggaaauggcg auggggguuga aagnnnnnnn nnnnnnnnnn 240

nnnnnn 246

<210> 206
<211> 274
<212> RNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG108-1 effector sgRNA 2

<220>
<221> modified_base
<222> (252)..(274)
<223> a, c, u, g, unknown or other

<400> 206
cuuguucgag uucguuugaa cugaacaagg guaaguaugg gccaguuuaa uugcuuuccg 60

ucccaggaua gcugccagcu ucuaccguag guucguccug caagugaugc uaagucgcgc 120

cuagcaucaa ggagcuaugu cuugauuguc uuggugucc gcccuggaug aguugaggug 180

uagaugcuuc uaucauggca gcuacuaaac gccccaagca aggggaaccc aucgaaagau 240

gggguugaaa gnnnnnnnnn nnnnnnnnnn nnnn 274

<210> 207
<211> 60
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG110-1-6 effector crRNA

<220>
<221> modified_base
<222> (9)..(40)
<223> a, c, u, g, unknown or other

<400> 207
cugguaaann nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn gugaccugcc gcauaggcag 60

<210> 208
<211> 59
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG110-2-6 effector crRNA

<220>
<221> modified_base
<222> (7)..(38)
<223> a, c, u, g, unknown or other

<400> 208
ccaagannnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnua ugaacugccg cauaggcag 59

<210> 209
<211> 227
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-4 effector putative tracrRNA

<400> 209
aauagcgccg cuuguucaug cgcgaaagcg ucucugaaca guguaaaugu ggguuaguuu 60

gacugucgug aagacggucu ugcuuucuga cccugguagc ugcccaccuu gaagcugcug 120

ucucuuguag acaggaauca ggugcgcccc caguaauaua ggugcggguu uaccgcagug 180

guggcuaccc aaucaccucc gagcaaggag gaacccaccu uaauuau 227

<210> 210
<211> 228
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-7 effector putative tracrRNA

<400> 210
aacagcgccg caguucaugc gucuuauggc gccucugugc ugugcaaaau gugggguuagu 60

uugacuguug gaagacaguc uugcuuucug acccugguag cugcccaccu ugaagcugcu 120

aucccuugug gauaggaauc aggugcgccc ccaguaauag aggugcgggu uuaccgcagu 180

gguggcuacc gaaucaccuc cgagcaagga ggaacccacc uuaauuau 228

<210> 211
<211> 328
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<220>
<223> MG64-9 effector putative tracrRNA

<400> 211
gcugaaguua gucucuggau acugaaguca gugucuaggg acugaaguaa ggggagcuuu 60

acccgcaagg gauacuuucg acccccugua gcugcccgcu ccuggugggg ugcccugaca 120

cccgccucau uacagcaaug uaugacuguc ugggcuaaug aaagaaggau uagggacgca 180

gguucacgac cuacuucaau cauagucugu gcaacccaga uaagugagua ugacccgcaa 240

gggucuaaac gccuuuagua gaaugguucu uccaguuugu uaagggcagg gcuugguucu 300

ccaagguggc uacgaauucu ucucgauu 328

<210> 212
<211> 187
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<220>
<223> MG64-10 effector putative tracrRNA

<400> 212
gguagcgcug gagccacaua guucauaagc ucacgcuucu uggacuuccu guguucucua 60

aaacggguuc uguuuuaccc uuaccaaggg auacuuucag auccgaguag cugcaagcuc 120

auggcggagu guccccugac gcuuugccac cgucauagcg augugauggc cgucuggcgu 180

augaacg 187

<210> 213
<211> 326
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-12 effector putative tracrRNA

<400> 213
aauagcgccc cuaaucaaau cgaaggcguu ugaggacggg gaaaaugggu aaguuuccau 60

cgaaaggugg uucuuuucag cccuauguag ucccacuccc uuguggagug uuagcaucug 120

aggugccuag cacagcaaag ucucucagga gacggaguca aagcugggga acaaguauua 180

cgaacgaauc ucgcgauucu auuaguauac guaggucgcu cccaugcaac aagaugcuca 240

ucucucagca augaggguag ggaacguauc ccaaacacua caugaauagc ucgucgggcu 300

ugcccgguga guauccacca agccuu 326

<210> 214
<211> 311
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-13 effector putative tracrRNA

<400> 214
cgcuagguuu gccuuuucgc gcgcugcuga agcagugaaa aguuccgaca guuuaaaagu 60

ucuaacacuu uuacugcuuu ccgucgggag uaguuguccg cuucugcguu guugagaccg 120

aaguuugucc aggugauagg cagaacguug augcaacagg uagauagucg cgcucuauca 180

ggaagcuguu cuaguuauua guguucgugc acuaauaaag aggauacagg gauacaugug 240

guugugucca gcaauggcac aacagggcca cuacucgaag ccccgagcaa ggguugagcc 300

uacccaaauu u 311

<210> 215

<211> 342
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-14 effector putative tracrRNA

<400> 215
aacagcgccg caguuuaagc ucuauagccg cugaacugug aaaaaugugg gucaguuugg 60

ucguugcaag acgaucgugc uuuccgaccc uaguaacugu ccgcucacug acugccaucc 120

uggggcaaau cuucaaauuu uguguauuug ugugggggaug gaaagcugca uuagucgauu 180

cucuuccucc aauguagcgu aggugcgcac ccagcagaag ugaguuaagc cuucacaaug 240

uggagguaca gaagcaucau cucuccauuu uuugguguag auggugugac ugaaguggua 300

guuaccgaau cgccccugau caaggggggag cccuccauaa uu 342

<210> 216
<211> 297
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-15 effector putative tracrRNA

<400> 216
aaugacguuu guuucugugc gcgcugcuga agcagugaaa aauacggaca gucuaauuug 60

cuuuccgucc ggaguaguug uccgcuucug cguuguugga accgaaguuu guccagguga 120

uaggcggagu guuucggcaa cagauggaga gucgcaccuc uaucaggggg caguuuuagu 180

gguuaguguu uagcgcacua ucacuaggau acagggauac auguggguagu gaccagagau 240

ggcacugcgg gaccgcuacu cgaagccccg aacaaggguu gagcuuaccc aaaucuu 297

<210> 217
<211> 226
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-16 effector putative tracrRNA

<400> 217
aauagcgccg caguuuaagc ucagcaagcc ucuggacugc gaaaaguaug ggguaguuug 60

accgucggua aacgguugug cuuucugccc cuggcgacug cccaccccga ugcugucgau 120

uucuuaacug ggaaucgaga ugaggugcgc ccccagcaaa agggaacggg uuuacuggag 180

ugguggucgc cgaaucaccc ccgagcaagg gggacucguc cuuugc 226

<210> 218
<211> 227
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-17 effector putative tracrRNA

<400> 218
aauagcgccg caguuuaagc ucaaugagcc ucuugacugc gaacaguaug ggguaguuug 60

accggcggua accgguugug cuuucugccc cuggcgacug cccaccccga ugcugucgau 120

uucucaaccg ggaaucgaga aucaggugcg cccccagcaa gagggaacgg guuuacugga 180

gugguggucg ccgaaucacc cccgagcaag ggggacucgu ccuuugc 227

<210> 219
<211> 227
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-19 effector putative tracrRNA

<400> 219
aauagcgccg caguuuaagc ucagcaagcc ucuggacugc gaaaaguaug ggguaguuug 60

accgucggua aacgguugug cuuucugccc cuggcgacug cccaccccga ugcugucgau 120

uucucaaccg ggaaucgaga aucaggugcg cccccagcaa gagggaacgg guuuacugga 180

gugguggucg ccgaaucacc cccgagcaag ggggacucgu ccuuugc 227

<210> 220
<211> 211
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-20 effector putative tracrRNA

<400> 220
aauagcgccg caguuuaagc ucaaugagcc ucuugacugc gaacaguaug ggguaguuug 60

accggcggua accgguugug cuuucugccc cuggcgacug cccaccccga ugcugucgau 120

uucucaaccg ggaaucgaga aucaggugcg cccccagcaa gagggaacgg guuuacugga 180

gugguggucg ccgaaucacc cccgagcaag g 211

<210> 221
<211> 222
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-21 effector putative tracrRNA

<400> 221
aauagcgccg caguucaugc uucuuugaag ccucugugcu gugcaaaaug uggguuaguu 60

uggcuguuga agaaacagcc uugcuuucug acccugguag cuguccaccc ugaagcugcu 120

aucccugug gauaggauag gugcgccccc agcaauaggg gagcggguau accgcagugg 180

uggcuacuga aucaccucca agcaaggagg aauccacuuu au 222

<210> 222
<211> 222
<212> RNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-22 effector putative tracrRNA

<400> 222
aauagcgccg caguucaugc uucuuugaag ccucugugcu gugcaaaaug uggguuaguu 60

uggcuguuga agaaacagcc uugcuuucug acccugguag cuguccaccc ugaagcugcu 120

aucccugug gauaggauag gugcgccccc agcaauaggg gagcggguau accgcagugg 180

uggcuacuga aucaccucca agcaaggagg aauccacuuu au 222

<210> 223
<211> 248
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-24 effector putative tracrRNA

<400> 223
aaaagcgccg uagaacaugc ucacgccucu guucugcgaa aaauuagggu uuguuuggcu 60

gucugacagc agucuuacuu ucugucccua gaaucugacc acuccgaugc ugcuguugua 120

agugaacuug auugagcucg gacaccaugc aacaggauaa ggggcgcacc cagcaagaga 180

ggacggacuu accguagugu uggcuucuga agcaacuccg accaaggagu aguccaugca 240

uucaucau 248

<210> 224
<211> 277
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-25 effector putative tracrRNA

<400> 224
aaucgcgccg cacauucaug uuccuugaga accucugaau ugcgaaagug ugggcuaguu 60

uguucgcuug augcgaaugu gcuuucuggc ccugguagcu guccgcccug augcugauuu 120

cuacggguaa cuguaggaau gauuaacucg uucuauagac agguuucgug cuuucuauug 180

uaacgggguc ggugcgcucc cagcaauagg gguguggguc uacuacagug auggcuacug 240

aaucaccucc gagcaaggag gaauccaccu uaacuuu 277

<210> 225
<211> 226
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-27 effector putative tracrRNA

<400> 225
aauagcgccg caguuuaagc ucagcaagcc ucuggacugc gaaaaguaug ggguaguuug 60

accgucggua aacgguugug cuuucugccc cuggcgacug cccaccccga ugcugucgau 120

uucuuaacug ggaaucgaga ugaggugcgc ccccagcaag agggaacggg uuuacuggag 180

ugguggucgc cgaaucaccc ccgagcaagg gggacucguc cuuugc 226

<210> 226
<211> 226
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-28 effector putative tracrRNA

<400> 226
aauagcgccg caguuuaagc ucagcaagcc ucuggacugc gaaaaguaug ggguaguuug 60

accgucggua aacgguugug cuuucugccc cuggcgacug cccaccccga ugcugucgau 120

uucuuaacug ggaaucgaga ugaggugcgc ccccagcaag agggaacggg uuuacuggag 180

ugguggucgc cgaaucaccc ccgagcaagg gggacucguc cuuugc 226

<210> 227
<211> 255
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-29 effector putative tracrRNA

<400> 227
uagcgcacga accugacuag cgucauuacg augcgaguuc uggaaauggg acaguuucau 60

ugcuuuccgu cccuggcacu gccugcuuac ucacgaccac aaggagacaa uucagcgcug 120

ugaguucgcg auucucuacu guuuaugaca aacuaacguu ugucuggcag uacagcaaga 180

uuguguauca gaccauguuu ggcgaagaua cauggcaggc cgaaucgcca augaaaccag 240

cggacuuccc cuaau 255

<210> 228
<211> 226
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-32 effector putative tracrRNA

<400> 228
aauagcgccg caguuuaagc ucaaugagcc ucuugacugc gaacaguaug ggguaguuug 60

accggcggua accgguugug cuuucugccc cuggcgacug cccaccccga ugcugucgau 120

uucuuaacug ggaaucgaga ugaggugcgc ccccagcaag agggaacggg uuuacuggag 180

ugguggucgc cgaaucaccc ccgagcaagg gggacucguc cuuugc 226

<210> 229
<211> 203
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-44 effector putative tracrRNA

<400> 229
ucuagcgccg cagcucaugu cagcaauggc caauguguug ugcuaaaugc gagcuaguuu 60

gacugccugc uaagcagucu ugcuuucugg cucaggugac uauccaccca aaggucguug 120

gugcgcuggc gauuugaggg cacggguucc ggagugauag uuaccauuac accuccggcc 180

aaggaggaau ccaccccacc ccc 203

<210> 230
<211> 203
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-46 effector putative tracrRNA

<400> 230
ucuagcgccg cagcucaugu cagcaauggc caauguguug ugcuaaaugc gagcuaguuu 60

gacugccugc uaagcagucu ugcuuucugg cucaggugac uauccaccca aaggucguug 120

gugcgcuggc gauuugaggg cacggguucc ggagugauag uuaccauuac accuccggcc 180

aaggaggaau ccaccccacc ccc 203

<210> 231
<211> 310
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-49 effector putative tracrRNA

<400> 231
gauugcgccu cgaucgaugc ucuaugagcc gcucgaucgu agaaaaaugg gugaguuuga 60

uuaucuacuu cguuagauaa ugcugcuuuc cgacccuggc auucuguccg cccuugaagc 120

ugcuucucau ggacuagcgu aagcucguug guaagaagga aaagucauaa uuuaaaguca 180

cgucuuucua guaugacaua ggugcgcucc cacgcaauau aggguucagc uuuuauuuua 240

uaaaaguaga gacuuuccuc uagugacagu gccgaaauga ccccgugcga gggguaacua 300

ccuaaguuuu 310

<210> 232
<211> 227
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-51 effector putative tracrRNA

<400> 232
aacagcgccg caguucaugu uuguuauaaa ccucuguacu gcgauaaaug cggguuaguu 60

ugacuguugu gagacagucu ugcuuucuga cccuaguagc ugcccaccuu gaugcugcug 120

uucccaguga acaggaauaa ggugcgcccc caguaauaga ggugcggguu uaccgcagug 180

guggcuacug aaucaccucc gacuaaggag gaauccaccu uaauuau 227

<210> 233
<211> 380
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-52 effector putative tracrRNA

<400> 233
accggggauc guccugguga agggagaguu gcuuuggauc ggucacaagc uuuuccuuaa 60

cuaauucuca cugacugacu aggaugaucg agggggguau guuuuaccac ugcaaggugg 120

auacuuucaa accccugugg uagcugcucg cuccugguga ggugcccuga cacuucaccc 180

cacuacagca augugugugg cugucugguu augagagaag uuagggcaua gguucguuac 240

cugcauucaa ucauaaauua cgcgaccaga uaagugagua ugauccgcaa ggaucuauau 300

gucuuuagca aagaagugcu ucugcuuugu uacuggcgua gggcaugguu cucuaaagug 360

gcuaccgaac cuucccaauu 380

<210> 234
<211> 257
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG64-56 effector putative tracrRNA

<400> 234
aacagcgccg caguucaugc ucuucugagu cucuguacug ugauaaaucu ggguuaguuu 60

aacgguugaa agaccguuuu gcuuucugac ccugguagcu gcucgcucuu gaugcugcug 120

ucuuuugaca ggauaggugc gcucccagca auaaagaguu aaagcugaua aagcuugagc 180

cguuguaaaa cggugggguu uaccucagug guggcuacug aaucaccccc uucgucgggg 240

gaacccuccu aaauauu 257

<210> 235
<211> 246
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG108-1 effector putative tracrRNA

<400> 235
uuuagcgcac uuguucgagu ucguuugaac ugaacaaggg uaaguauggg ccaguuuaau 60

ugcuuuccgu cccaggauag cugccagcuu cuaccguagg uucguccugc aagugaugcu 120

aagucgcgcc uagcaucaag gagcuauguc uugauugucu uggguguccg cccuggauga 180

guugaggugu agaugcuucu aucauggcag cuacuaaacg ccccaagcaa ggggaaccca 240

ucuuua 246

<210> 236
<211> 223
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<220>
<223> MG108-2 effector putative tracrRNA

<400> 236
aauagcgccg uaguucaugc uugcuaaagc cucugaauug cgaaaagucc ggguuagugc 60

ugucggcaga cagcguugcu uucugacccu gguagcugcc caccccgaug cugcuguccc 120

uugcagacag gaaccaggug cgcccccagu aauaagggug uggguuuacc acaguggugg 180

cuacugaauc accuccgagc aaggaggaau ccaccuuaac uau 223

<210> 237
<211> 41
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-9 effector crRNA

<400> 237
gucuuucauc cuaucucgcg ccagaucgcu uccugcaacc c 41

<210> 238
<211> 37
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-12 effector crRNA

<400> 238
guugcaagcg ccuccuuggc uguuugguggg uggaaag 37

<210> 239
<211> 37
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-14 effector crRNA

<400> 239
guugcaaucg ccuucccaga gauggguggg cugaaag 37

<210> 240
<211> 38
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-15 effector crRNA

<400> 240
guuacaauua cccucccagc guuggguggg uugaaagg 38

<210> 241
<211> 39
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-16 effector crRNA

<400> 241
aaguugcauc cgcuuuccag caaccagggc gggugaaag 39

<210> 242
<211> 37
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-17 effector crRNA

<400> 242
guugcauccg cuuuccagca accagggcgg gugaaag 37

<210> 243
<211> 33
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-19 effector crRNA

<400> 243
guacccaaag ccuuuuuucc uuaagccuau ccg 33

<210> 244
<211> 27
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-25 effector crRNA

<400> 244
guuucaaccg ccaucccagc uaggggu 27

<210> 245
<211> 38
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-27 effector crRNA

<400> 245
aguugcaucc gcuuuccagc aaccagggcg ggugaaag 38

<210> 246
<211> 38
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-28 effector crRNA

<400> 246
guugcaucug cuuuucagca acuagggcgg gggaaagc 38

<210> 247
<211> 37
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-29 effector crRNA

<400> 247
ggcgcgaucg ccuuuauggg uacgggcaag uugaaag 37

<210> 248
<211> 41
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-32 effector crRNA

<400> 248
aaguugcauc cgcuuuccag caaccagggc gggugaaagu u 41

<210> 249

<211> 37
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-44 effector crRNA

<400> 249
guugccuccc gcuucgaggc acgggaacga uugaaag 37

<210> 250
<211> 37
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-46 effector crRNA

<400> 250
guugccuccc gcuucgaggc acgggaacga uugaaag 37

<210> 251
<211> 43
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-49 effector crRNA

<400> 251
guugcaacac ucccugacug ccugacacaa augccucgaa agc 43

<210> 252
<211> 34
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

oligonucleotide

<220>
<223> MG64-52 effector crRNA

<400> 252
gucgcaauga cuauuuuggc uuggggcgga auga 34

<210> 253
<211> 37
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-51 effector crRNA

<400> 253
guuucaacac cccucccgaa gugggggcggg uugaaag 37

<210> 254
<211> 37
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-56 effector crRNA

<400> 254
guuucaacaa ccaucccagc uagggguggg uugaaag 37

<210> 255
<211> 37
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-2 effector crRNA

<400> 255

guuucaacga ccaucccaac uagggguggg uugaaag 37

<210> 256
<211> 37
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-4 effector crRNA

<400> 256
guuucaacuu uccuuccagc uagaggcggg uugaaag 37

<210> 257
<211> 37
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-6 effector crRNA

<400> 257
guuucaacca ccaucucaac uagggauggg uugaaag 37

<210> 258
<211> 38
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-7 effector crRNA

<400> 258
guuucaacgc cccuucaagc uuugggcggg uugaaagc 38

<210> 259
<211> 36
<212> RNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG64-13 effector crRNA

<400> 259
uggcaauugc ccuuccagug uuggggugggu ugaaag 36

<210> 260
<211> 37
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG108-2 effector crRNA

<400> 260
guuucaacga ccaucccgac aaggggguggg uugaaag 37

<210> 261
<211> 37
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<223> MG108-1 effector crRNA

<400> 261
guugcgaucg ccgcuccggu ggcgaugggg uugaaag 37