



(12) **DEMANDE DE BREVET CANADIEN**
CANADIAN PATENT APPLICATION

(13) **A1**

(86) Date de dépôt PCT/PCT Filing Date: 2019/08/28
(87) Date publication PCT/PCT Publication Date: 2020/03/05
(85) Entrée phase nationale/National Entry: 2021/02/17
(86) N° demande PCT/PCT Application No.: US 2019/048607
(87) N° publication PCT/PCT Publication No.: 2020/047124
(30) Priorités/Priorities: 2018/08/28 (US62/723,886);
2018/08/31 (US62/725,778); 2019/05/21 (US62/850,883);
2019/06/21 (US62/864,924)

(51) Cl.Int./Int.Cl. *C12N 15/10* (2006.01),
C12N 9/12 (2006.01)
(71) Demandeur/Applicant:
FLAGSHIP PIONEERING INNOVATIONS VI, LLC, US
(72) Inventeurs/Inventors:
RUBENS, JACOB ROSENBLUM, US;
VON MALTZAHN, GEOFFREY A., US;
CITORIK, ROBERT JAMES, US;
STEINBERG, BARRETT ETHAN, US;
LI, DONGHUI, US
(74) Agent: SMART & BIGGAR LLP

(54) Titre : PROCEDES ET COMPOSITIONS POUR MODULER UN GENOME
(54) Title: METHODS AND COMPOSITIONS FOR MODULATING A GENOME

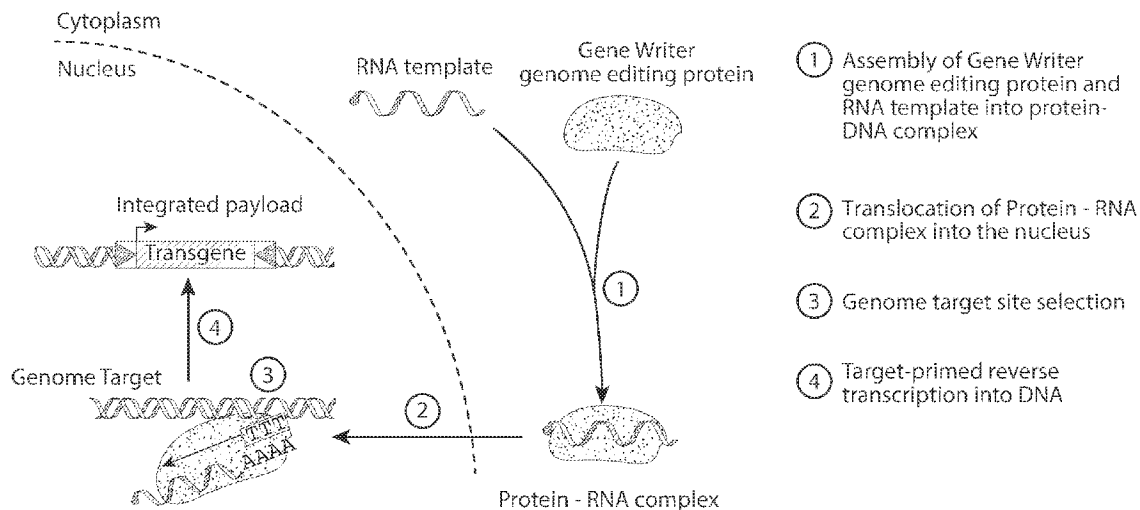


Figure 1

(57) **Abrégé/Abstract:**

Methods and compositions for modulating a target genome are disclosed.

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

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

05 March 2020 (05.03.2020)



(10) International Publication Number

WO 2020/047124 A1

(51) International Patent Classification:

C12N 15/10 (2006.01) C12N 9/12 (2006.01)

(21) International Application Number:

PCT/US2019/048607

(22) International Filing Date:

28 August 2019 (28.08.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/723,886	28 August 2018 (28.08.2018)	US
62/725,778	31 August 2018 (31.08.2018)	US
62/850,883	21 May 2019 (21.05.2019)	US
62/864,924	21 June 2019 (21.06.2019)	US

(71) Applicant: FLAGSHIP PIONEERING, INC. [US/US];

55 Cambridge Parkway, 8th Floor, Suite 800E, Cambridge, MA 02142 (US).

(72) Inventors: RUBENS, Jacob, Rosenblum; 55 Cambridge

Parkway, 8th Floor, Suite 800E, Cambridge, MA 02142

(US). VON MALTZAHN, Geoffrey, A.; 55 Cambridge

Parkway, 8th Floor, Suite 800E, Cambridge, MA 02142

(US). CITORIK, Robert, James; 55 Cambridge Park-

way, 8th Floor, Suite 800E, Cambridge, MA 02142 (US).

STEINBERG, Barrett, Ethan; 55 Cambridge Parkway,

8th Floor, Suite 800E, Cambridge, MA 02142 (US). LI,

Donghui; 55 Cambridge Parkway, 8th Floor, Suite 800E,

Cambridge, MA 02142 (US).

(74) Agent: KOYFMAN, Hannah, R. et al.; Lando & Anastasi,

LLP, Riverfront Office Park, One Main Street, Suite 1100,

Cambridge, MA 02142 (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, 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, SM, ST, SV, SY, TH, TJ, TM, TN,

TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ,

UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,

(54) Title: METHODS AND COMPOSITIONS FOR MODULATING A GENOME

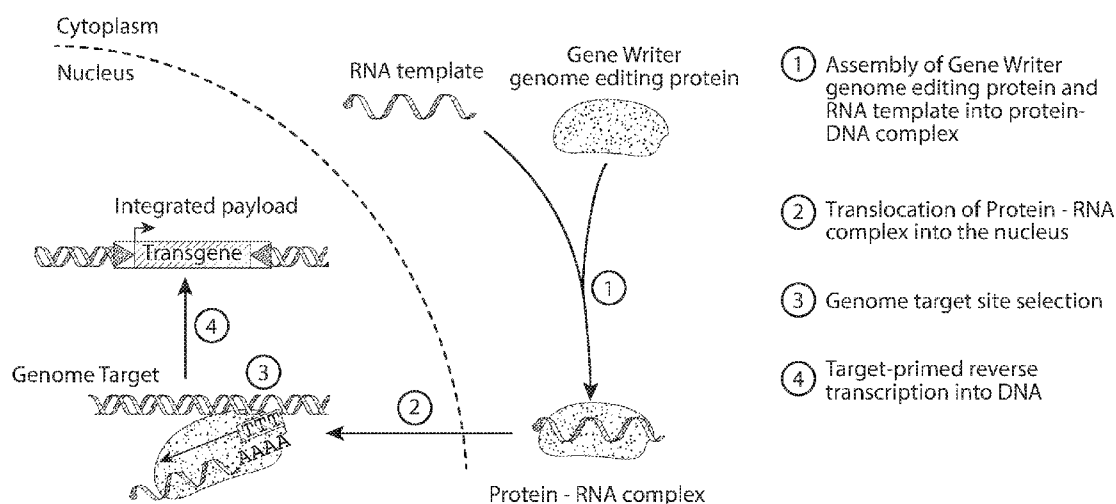


Figure 1

(57) Abstract: Methods and compositions for modulating a target genome are disclosed.

WO 2020/047124 A1

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))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
- *with sequence listing part of description (Rule 5.2(a))*

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 303

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 303

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHIER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

METHODS AND COMPOSITIONS FOR MODULATING A GENOME

This application claims priority to U.S. Serial No.: 62/723,886 filed August 28, 2018, U.S. Serial No.: 62/725,778 filed August 31, 2018, U.S. Serial No.: 62/850,883 filed May 21, 2019, and U.S. Serial No.: 62/864,924 filed June 21, 2019, the entire contents of each of which is incorporated herein by reference.

SEQUENCE LISTING

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 28, 2019, is named V2065-7000WO_SL.txt and is 4,004,548 bytes in size.

BACKGROUND

Integration of a nucleic acid of interest into a genome occurs at low frequency and with little site specificity, in the absence of a specialized protein to promote the insertion event. Some existing approaches, like CRISPR/Cas9, are more suited for small edits and are less effective at integrating longer sequences. Other existing approaches, like Cre/loxP, require a first step of inserting a loxP site into the genome and then a second step of inserting a sequence of interest into the loxP site. There is a need in the art for improved proteins for inserting sequences of interest into a genome.

SUMMARY OF THE INVENTION

This disclosure relates to novel compositions, systems and methods for altering a genome at one or more locations in a host cell, tissue or subject, *in vivo* or *in vitro*. In particular, the invention features compositions, systems and methods for the introduction of exogenous genetic elements into a host genome.

Features of the compositions or methods can include one or more of the following enumerated embodiments.

1. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain; and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence that encodes a therapeutic polypeptide or that encodes a mammalian (e.g., human) polypeptide, or a fragment or variant thereof.

2. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain; and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence, wherein one or more of:

- i. the heterologous object sequence encodes a protein, e.g. an enzyme (e.g., a lysosomal enzyme) or a blood factor (e.g., Factor I, II, V, VII, X, XI, XII or XIII);
- ii. the heterologous object sequence comprises a tissue specific promoter or enhancer;
- iii. the heterologous object sequence encodes a polypeptide of greater than 250, 300, 400, 500, or 1,000 amino acids, and optionally up to 7,500 amino acids;
- iv. the heterologous object sequence encodes a fragment of a mammalian gene but does not encode the full mammalian gene, e.g., encodes one or more exons but does not encode a full-length protein;
- v. the heterologous object sequence encodes one or more introns;
- vi. the heterologous object sequence is other than a GFP, e.g., is other than a fluorescent protein or is other than a reporter protein.
- vii. the heterologous object sequence is other than a T cell chimeric antigen receptor

3. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain; and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence.

4. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a target DNA binding domain, (ii) a reverse transcriptase domain and (iii) an endonuclease domain; and

5 (b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence.

5. A system for modifying DNA comprising:

10 (a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain, wherein one or both of (i) or (ii) are derived from an avian retrotransposase, e.g., have a sequence of Table 2 or 3 or at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto; and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence.

15

6. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain, wherein the polypeptide has an activity at 37°C that is no less than 70%, 75%, 80%, 85%, 90%, or 95% of its activity at 25°C under otherwise similar conditions; and

20

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence.

7. The system of embodiment 6, wherein the polypeptide is derived from an avian

25 retrotransposase, e.g., an avian retrotransposase of column 8 of Table 3, or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto.

8. The system of embodiment 6, wherein the avian retrotransposase is a retrotransposase from *Taeniopygia guttata*, *Geospiza fortis*, *Zonotrichia albicollis*, or *Tinamus guttatus*, or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto.

30

9. The system of embodiment 6, wherein the polypeptide is derived from a retrotransposase of column 8 of Table 3, or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto.

5

10. The system of any of the preceding embodiments, wherein the template RNA comprises a sequence of Table 3 (e.g., one or both of a 5' untranslated region of column 6 of Table 3 and a 3' untranslated region of column 7 of Table 3), or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto.

10

11. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain; and

15 (b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence, wherein one or more of:

- i. the nucleic acid encoding the polypeptide and the template RNA or a nucleic acid encoding the template RNA are separate nucleic acids;
- ii. the template RNA does not encode an active reverse transcriptase, e.g., comprises an inactivated mutant reverse transcriptase, e.g., as described in Examples 1-2, or does
- 20 not comprise a reverse transcriptase sequence; or
- iii. the template RNA does not encode an active endonuclease, e.g., comprises an inactivated endonuclease or does not comprise an endonuclease; or
- iv. the template RNA comprises one or more chemical modifications.

20

25 12. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain; and

30 (b) a template RNA (or DNA encoding the template RNA) comprising (i) a 5' untranslated sequence that binds the polypeptide, (ii) a 3' untranslated sequence that binds the polypeptide, (iii) a heterologous object sequence, and (iv) a promoter operably linked to the heterologous object sequence,

wherein the promoter is disposed between the 5' untranslated sequence that binds the polypeptide and the heterologous sequence, or

wherein the promoter is disposed between the 3' untranslated sequence that binds the polypeptide and the heterologous sequence.

5

13. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain; and

10 (b) a template RNA (or DNA encoding the template RNA) comprising (i) a 5' untranslated sequence that binds the polypeptide, (ii) a 3' untranslated sequence that binds the polypeptide, and (iii) a heterologous object sequence, and

wherein the heterologous object sequence comprises an open reading frame (or the reverse complement thereof) in a 5' to 3' orientation on the template RNA; or

15 wherein the heterologous object sequence comprises an open reading frame (or the reverse complement thereof) in a 3' to 5' orientation on the template RNA.

14. A system for modifying DNA comprising:

20 (a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain, wherein at least one of (i) or (ii) is heterologous, and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence.

15. A system for modifying DNA comprising:

25 (a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a target DNA binding domain, (i) a reverse transcriptase domain and (iii) an endonuclease domain, wherein at least one of (i), (ii) or (iii) is heterologous, and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence.

30

16. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a sequence at least 80% identical (e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identical) to a reverse transcriptase domain of a purinic/aprimidinic endonuclease (APE)-type non-LTR retrotransposon and (ii) a sequence at least 80% identical (e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identical) to an endonuclease domain of an APE-type non-LTR retrotransposon; and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence.

17. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a sequence at least 80% identical (e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identical) to a reverse transcriptase domain of a restriction enzyme-like endonuclease (RLE)-type non-LTR retrotransposon, (ii) a sequence at least 80% identical (e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identical) to an endonuclease domain of a RLE-type non-LTR retrotransposon, and (iii) a heterologous target DNA binding domain (e.g., a heterologous zinc-finger DNA binding domain); and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence.

18. The system of any of the preceding embodiments, wherein the template RNA comprises (iii) a promoter operably linked to the heterologous object sequence.

19. The system of any of the preceding embodiments, wherein the polypeptide further comprises (iii) a DNA-binding domain.

20. The system of embodiment 17, wherein the polypeptide comprises a sequence at least 80% identical (e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identical) to the sequence of SEQ ID NO: 1016.

21. The system of any of the preceding embodiments, wherein the polypeptide comprises a sequence at least 80% identical (e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identical) to a sequence in column 8 of Table 3.

22. The system of any of the preceding embodiments, wherein the nucleic acid encoding the polypeptide and the template RNA or the nucleic acid encoding the template RNA are covalently linked, e.g., are part of a fusion nucleic acid.

23. The system of embodiment 22, wherein the fusion nucleic acid comprises RNA.

24. The system of embodiment 22, wherein the fusion nucleic acid comprises DNA.

25. The system of any of the preceding embodiments, wherein (b) comprises template RNA.

26. The system of embodiment 25, wherein the template RNA further comprises a nuclear localization signal.

27. The system of any of the preceding embodiments, wherein (a) comprises RNA encoding the polypeptide.

28. The system of embodiment 27, wherein the RNA of (a) and the RNA of (b) are separate RNA molecules.

29. The system of embodiment 28, wherein the RNA of (a) and the RNA of (b) are present at a ratio of between 10:1 and 5:1, 5:1 and 2:1, 2:1 and 1:1, 1:1 and 1:2, 1:2 and 1:5, or 1:5 and 1:10.

30. The system of embodiment 28, wherein the RNA of (a) does not comprise a nuclear localization signal.

31. The system of any of the preceding embodiments, wherein the polypeptide further comprises a nuclear localization signal and/or a nucleolar localization signal.

32. The system of any of the preceding embodiments, wherein (a) comprises an RNA that encodes: (i) the polypeptide and (ii) a nuclear localization signal and/or a nucleolar localization signal.

33. The system of any of the preceding embodiments, wherein the RNA comprises a pseudoknot sequence, e.g., 5' of the heterologous object sequence.

34. The system of embodiment 33, wherein the RNA comprises a stem-loop sequence or a helix, 5' of the pseudoknot sequence.

35. The system of embodiment 33 or 34, wherein the RNA comprises one or more (e.g., 2, 3, or more) stem-loop sequences or helices 3' of the pseudoknot sequence, e.g. 3' of the pseudoknot sequence and 5' of the heterologous object sequence.

36. The system of any of embodiments 33-35, wherein the template RNA comprising the pseudoknot has catalytic activity, e.g., RNA-cleaving activity, e.g. cis-RNA-cleaving activity.

37. The system of any of the preceding embodiments, wherein the RNA comprises at least one stem-loop sequence or helix, e.g., 3' of the heterologous object sequence, e.g. 1, 2, 3, 4, 5 or more stem-loop sequences, hairpins or helices sequences.

38. Any above-numbered system, wherein the polypeptide comprises a sequence of at least 50 amino acids (e.g., at least 100, 150, 200, 300, 500 amino acids) having at least 80% identity (e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identity) to a sequence of a polypeptide listed in Table 1, or a reverse transcriptase domain or endonuclease domain thereof.

39. Any above-numbered system, wherein the polypeptide comprises a sequence of at least 50 amino acids (e.g., at least 100, 150, 200, 300, 500 amino acids) having at least 80% identity

(e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identity) to a sequence of a polypeptide listed in any of Tables 2-3 or a reverse transcriptase domain, endonuclease domain, or DNA binding domain thereof.

40. Any above-numbered system, wherein the polypeptide comprises a sequence of at least 50 amino acids (e.g., at least 100, 150, 200, 300, 500 amino acids) having at least 80% identity (e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identity) to the amino acid sequence of column 8 of Table 3, or a reverse transcriptase domain, endonuclease domain, or DNA binding domain thereof.

41. Any above-numbered system, wherein the template RNA comprises a sequence of Table 3 (e.g., one or both of a 5' untranslated region of column 6 of Table 3 and a 3' untranslated region of column 7 of Table 3), or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto.

42. The system of embodiment 41, wherein the template RNA comprises a sequence of about 100-125 bp from a 3' untranslated region of column 7 of Table 3, e.g., wherein the sequence comprises nucleotides 1-100, 101-200, or 201-325 of the 3' untranslated region of column 7 of Table 3, or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto.

43. Any above-numbered system, wherein (a) comprises RNA and (b) comprises RNA.

44. Any above-numbered system, which comprises only RNA, or which comprises more RNA than DNA by an RNA:DNA ratio of at least 10:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, or 100:1.

45. Any above-numbered system, which does not comprise DNA, or which does not comprise more than 10%, 5%, 4%, 3%, 2%, or 1% DNA by mass or by molar amount.

46. Any above-numbered system, which is capable of modifying DNA by insertion of the heterologous object sequence without an intervening DNA-dependent RNA polymerization of (b).

47. Any above-numbered system, which is capable of modifying DNA by insertion of a heterologous object sequence in the presence of an inhibitor of a DNA repair pathway (e.g., SCR7, a PARP inhibitor), or in a cell line deficient for a DNA repair pathway (e.g., a cell line deficient for the nucleotide excision repair pathway or the homology-directed repair pathway).

48. Any above-numbered system, which does not cause formation of a detectable level of double stranded breaks in a target cell.

49. Any above-numbered system, which is capable of modifying DNA using reverse transcriptase activity, and optionally in the absence of homologous recombination activity.

50. Any above-numbered system, wherein the template RNA has been treated to reduce secondary structure, e.g., was heated, e.g., to a temperature that reduces secondary structure, e.g., to at least 70, 75, 80, 85, 90, or 95 C.

51. The system of embodiment 50, wherein the template RNA was subsequently cooled, e.g., to a temperature that allows for secondary structure, e.g., to less than or equal to 30, 25, or 20 C

52. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain; and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide, (ii) a heterologous object sequence, (iii) a first homology domain having at least 10 bases of 100% identity to a target DNA strand, at the 5' end of the template RNA, and (iv) a second homology domain having at least 10 bases of 100% identity to a target DNA strand, 5' end of the template RNA.

53. The system of any of the preceding embodiments, wherein (a) and (b) are part of the same nucleic acid.

54. The system of any of embodiments 1-52, wherein (a) and (b) are separate nucleic acids.

5

55. The system of any of the preceding embodiments, wherein the template RNA comprises at least 10 bases of 100% identity to a target DNA strand (e.g., wherein the target DNA strand is a human DNA sequence), at the 5' end of the template RNA.

10 56. The system of any of the preceding embodiments, wherein the template RNA comprises at least 10 bases of 100% identity to a target DNA strand (e.g., wherein the target DNA strand is a human DNA sequence), at the 3' end of the template RNA.

15 57. A host cell (e.g., a mammalian cell, e.g., a human cell) comprising any preceding numbered system.

58. A method of modifying a target DNA strand in a cell, tissue or subject, comprising administering any preceding numbered system to the cell, tissue or subject, wherein the system reverse transcribes the template RNA sequence into the target DNA strand, thereby modifying
20 the target DNA strand.

59. The method of embodiment 58, wherein the cell, tissue or subject is a mammalian (e.g., human) cell, tissue or subject.

25 60. The method of any of the preceding embodiments, wherein the cell is a fibroblast.

61. The method of any of the preceding embodiments, wherein the cell is a primary cell.

62. The method of any of the preceding embodiments, wherein the cell is not immortalized.

30

63. A method of modifying the genome of a mammalian cell, comprising contacting the cell with:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain, (ii) an endonuclease domain, and optionally (iii) a DNA-binding domain; and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence.

64. The method of embodiment 63, wherein the polypeptide does not comprise a target DNA binding domain.

65. The method of embodiment 63, wherein the polypeptide is derived from an APE-type transposon reverse transcriptase.

66. The method of embodiment 63, wherein the (i) a reverse transcriptase domain (ii) an endonuclease domain, or both of (i) and (ii), have a sequence of Table 1 or a sequence having at least 80%, 85%, 90%, 95%, 97%, 98%, 99%, 100% identity thereto.

67. The method of embodiment 63, wherein the polypeptide further comprises a target DNA binding domain.

68. A method of modifying the genome of a mammalian cell, comprising contacting the cell with:

(a) an RNA encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain, (ii) an endonuclease domain, and optionally (iii) a DNA-binding domain; and

(b) a template RNA comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence,

wherein the method does not comprise contacting the mammalian cell with DNA, or wherein the compositions of (a) and (b) do not comprise more than 1%, 0.5%, 0.2%, 0.1%, 0.05%, 0.02%, or 0.01% DNA by mass or by molar amount of nucleic acid.

69. The method of embodiment 68, which results in the addition of at least 5, 10, 20, 50, 100, 200, 500, 1,000, 2,000, or 5,000 base pairs of exogenous DNA sequence to the genome of the mammalian cell.

5

70. The method of embodiment 68 or 69, which results in the addition of a protein coding sequence to the genome of the mammalian cell.

71. A method of inserting DNA into the genome of a mammalian cell, comprising
10 contacting the cell with an RNA composition, wherein the RNA composition comprises:
(a) a first RNA that directs insertion of a template RNA into the genome, and
(b) a template RNA comprising a heterologous sequence,
wherein the method does not comprise contacting the mammalian cell with DNA, or
wherein the compositions of (a) and (b) do not comprise more than 1%, 0.5%, 0.2%, 0.1%,
15 0.05%, 0.02%, or 0.01% DNA by mass or by molar amount of nucleic acid,

wherein the method results in the addition of at least 5, 10, 20, 50, 100, 200, 500, 1,000, 2,000, or 5,000 base pairs of DNA (e.g., exogenous DNA) sequence to the genome of the mammalian cell.

20 72. The method of embodiment 71, wherein the first RNA encodes a polypeptide (e.g., a polypeptide of any of Tables 1, 2, or 3 herein), wherein the polypeptide directs insertion of the template RNA into the genome.

73. The method of embodiments 72, wherein the template RNA further comprises a sequence
25 that binds the polypeptide.

74. A method of adding at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 500, 1000 bp of exogenous DNA to the genome of a mammalian cell, without delivery of DNA to the cell.

30

75. A method of adding at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 500, 1000 bp of exogenous DNA to the genome of a mammalian cell, wherein the method does not comprise contacting the mammalian cell with DNA, or wherein the method comprises contacting the mammalian cell with a composition comprising less than 1%,
5 0.5%, 0.2%, 0.1%, 0.05%, 0.02%, or 0.01% DNA by mass or by molar amount of nucleic acid.

76. A method of adding at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 500, 1000 bp of exogenous DNA to the genome of a mammalian cell, comprising delivering only RNA to the mammalian cell.

10 77. A method of adding at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 500, 1000 bp of exogenous DNA to the genome of a mammalian cell, comprising delivering RNA and protein to the mammalian cell.

15 78. The method of any one of embodiments 68-77, wherein the template RNA serves as the template for insertion of the exogenous DNA.

79. The method of any one of embodiments 68-78, which does not comprise DNA-dependent RNA polymerization of exogenous DNA.

20 80. The method of any of embodiments 58-79, which results in the addition of at least 5, 10, 20, 50, 100, 200, 500, 1,000, 2,000, or 5,000 base pairs of DNA to the genome of the mammalian cell.

25 81. The methods of any of embodiments 68-80, wherein the RNA of (a) and the RNA of (b) are covalently linked, e.g., are part of the same transcript.

82. The methods of any of embodiments 68-80, wherein the RNA of (a) and the RNA of (b) are separate RNAs.

83. The method of any of embodiments 58-82, which does not comprise contacting the mammalian cell with a template DNA.

84. A method of modifying the genome of a human cell, comprising contacting the cell with:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain, (ii) an endonuclease domain, and optionally (iii) a DNA-binding domain; and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence,

wherein the method results in insertion of the heterologous object sequence into the human cell's genome,

wherein the human cell does not show upregulation of any DNA repair genes and/or tumor suppressor genes, or wherein no DNA repair gene and/or tumor suppressor gene is upregulated by more than 10%, 5%, 2%, or 1%, e.g., wherein upregulation is measured by RNA-seq, e.g., as described in Example 14.

85. A method of adding an exogenous coding region to the genome of a cell (e.g., a mammalian cell), comprising contacting the cell with an RNA comprising the non-coding strand of the exogenous coding region, wherein optionally the RNA does not comprise a coding strand of the exogenous coding region, wherein optionally the delivery comprises non-viral delivery.

86. A method of expressing a polypeptide in a cell (e.g., a mammalian cell), comprising contacting the cell with an RNA, wherein the RNA comprises a non-coding strand that is the reverse complement of a sequence that would encoding the polypeptide, wherein optionally the RNA does not comprise a coding strand encoding the polypeptide, wherein optionally the delivery comprises non-viral delivery.

87. The method of any of embodiments 58-86, wherein the sequence that is inserted into the mammalian genome is a sequence that is exogenous to the mammalian genome.

88. The method of any of embodiments 58-87, which operates independently of a DNA template.

89. The method of any of embodiments 58-88, wherein the cell is part of a tissue.

5

90. The method of any of embodiments 58-89, wherein the mammalian cell is euploid, is not immortalized, is part of an organism, is a primary cell, is non-dividing, is a hepatocyte, or is from a subject having a genetic disease.

10 91. The method of any of embodiments 58-90, wherein the contacting comprises contacting the cell with a plasmid, virus, viral-like particle, virosome, liposome, vesicle, exosome, or lipid nanoparticle.

15 92. The method of any of embodiments 58-91, wherein the contacting comprises using non-viral delivery.

93. The method of any of embodiments 58-92, which comprises comprising contacting the cell with the template RNA (or DNA encoding the template RNA), wherein the template RNA comprises the non-coding strand of an exogenous coding region, wherein optionally the template
20 RNA does not comprise a coding strand of the exogenous coding region, wherein optionally the delivery comprises non-viral delivery, thereby adding the exogenous coding region to the genome of the cell.

25 94. The method of any of embodiments 58-93, which comprises contacting the cell with the template RNA (or DNA encoding the template RNA), wherein the template RNA comprises a non-coding strand that is the reverse complement of a sequence that would encoding the polypeptide, wherein optionally the template RNA does not comprise a coding strand encoding the polypeptide, wherein optionally the delivery comprises non-viral delivery, thereby expressing the polypeptide in the cell.

30

95. The method of any of embodiments 63-94, wherein the contacting comprises administering (a) and (b) to a subject, e.g., intravenously.

96. The method of any of embodiments 63-95, wherein the contacting comprises administering a dose of (a) and (b) to a subject at least twice.

97. The method of any of embodiments 63-96, wherein the polypeptide reverse transcribes the template RNA sequence into the target DNA strand, thereby modifying the target DNA strand.

98. The method of any embodiments 63-97, wherein (a) and (b) are administered separately.

99. The method of any of embodiments 63-97, wherein (a) and (b) are administered together.

100. The method of any of embodiments 63-99, wherein the nucleic acid of (a) is not integrated into the genome of the host cell.

101. Any preceding numbered method, wherein the sequence that binds the polypeptide has one or more of the following characteristics:

- (a) is at the 3' end of the template RNA;
- (b) is at the 5' end of the template RNA;
- (b) is a non-coding sequence;
- (c) is a structured RNA; or
- (d) forms at least 1 hairpin loop structures.

102. Any preceding numbered method, wherein the template RNA further comprises a sequence comprising at least 20 nucleotides of at least 80% identity (e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identity) to a target DNA strand.

103. Any preceding numbered method, wherein the template RNA further comprises a sequence comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50,

60, 70, 80, 90, 100, 110, 120, 130, 140, 150 nucleotides of at least 80% identity (e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identity) to a target DNA strand.

104. Any preceding numbered method, wherein the sequence comprising at least 2, 3, 4, 5, 6,
 5 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140,
 150 nucleotides, or about: 2-10, 10-20, 20-30, 30-40, 40-50, 50-60, 60-70, 70-80, 80-90, 90-100,
 10-100, or 2-100 nucleotides, of at least 80% identity to a target DNA strand is at the 3' end of
 the template RNA.

10 105. Any preceding numbered method, wherein the template RNA further comprises a sequence
 comprising at least 100 nucleotides of at least 80% identity (e.g., at least 85%, 90%, 95%, 97%,
 98%, 99%, 100% identity) to a target DNA strand, e.g., at the 3' end of the template RNA.

106. The method of embodiment 104 or 105, wherein the site in the target DNA strand to
 15 which the sequence comprises at least 80% identity is proximal to (e.g., within about: 0-10, 10-
 20, 20-30, 30-50, or 50-100 nucleotides of) a target site on the target DNA strand that is
 recognized (e.g., bound and/or cleaved) by the polypeptide comprising the endonuclease.

107. Any preceding numbered method, wherein the sequence comprising at least 2, 3, 4, 5, 6,
 20 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140,
 150 nucleotides, or about: 2-10, 10-20, 20-30, 30-40, 40-50, 50-60, 60-70, 70-80, 80-90, 90-100,
 10-100, or 2-100 nucleotides, of at least 80% identity to a target DNA strand is at the 3' end of
 the template RNA;

optionally wherein the site in the target DNA strand to which the sequence comprises at
 25 least 80% identity is proximal to (e.g., within about: 0-10, 10-20, or 20-30 nucleotides of) a
 target site on the target DNA strand that is recognized (e.g., bound and/or cleaved) by the
 polypeptide comprising the endonuclease.

108. The method of embodiment 107, wherein the target site is the site in the human genome
 30 that has the closest identity to a native target site of the polypeptide comprising the

endonuclease, e.g., wherein the target site in the human genome has at least about: 16, 17, 18, 19, or 20 nucleotides identical to the native target site.

109. Any preceding numbered method, wherein the template RNA has at least 3, 4, 5, 6, 7, 8,
5 9, or 10 bases of 100% identity to the target DNA strand.

110. Any preceding numbered method, wherein the at least 3, 4, 5, 6, 7, 8, 9, or 10 bases of 100% identity to the target DNA strand are at the 3' end of the template RNA.

10 111. Any preceding numbered method, wherein the at least 3, 4, 5, 6, 7, 8, 9, or 10 bases of 100% identity to the target DNA strand are at the 5' end of the template RNA.

112. Any preceding numbered method, wherein the template RNA comprises at least 3, 4, 5, 6, 7, 8, 9, or 10 bases of 100% identity to the target DNA strand at the 5' end of the template
15 RNA and at least 3, 4, 5, 6, 7, 8, 9, or 10 bases of 100% identity to the target DNA strand at the 3' end of the template RNA.

113. Any preceding numbered method, wherein the heterologous object sequence is between 50-50,000 base pairs (e.g., between 50-40,000 bp, between 500-30,000 bp between 500-20,000
20 bp, between 100-15,000 bp, between 500-10,000 bp, between 50-10,000 bp, between 50-5,000 bp).

114. Any preceding numbered method, wherein the heterologous object sequence is at least 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, or 700 bp.
25

115. Any preceding numbered method, wherein the heterologous object sequence is at least 715, 750, 800, 950, 1,000, 2,000, 3,000, or 4,000 bp.

116. Any preceding numbered method, wherein the heterologous object sequence is less than
30 5,000, 10,000, 15,000, 20,000, 30,000, or 40,000 bp.

117. Any preceding numbered method, wherein the heterologous object sequence is less than 700, 600, 500, 400, 300, 200, 150, or 100 bp.

118. Any preceding numbered method, wherein the heterologous object sequence comprises:

5 (a) an open reading frame, e.g., a sequence encoding a polypeptide, e.g., an enzyme (e.g., a lysosomal enzyme), a membrane protein, a blood factor, an exon, an intracellular protein (e.g., a cytoplasmic protein, a nuclear protein, an organellar protein such as a mitochondrial protein or lysosomal protein), an extracellular protein, a structural protein, a signaling protein, a regulatory protein, a transport protein, a sensory protein, a motor protein, a defense protein, or a storage
10 protein;

(b) a non-coding and/or regulatory sequence, e.g., a sequence that binds a transcriptional modulator, e.g., a promoter, an enhancer, an insulator;

(c) a splice acceptor site;

(d) a polyA site;

15 (e) an epigenetic modification site; or

(f) a gene expression unit.

119. Any preceding numbered method, wherein the target DNA is a genomic safe harbor (GSH) site.

20 120. Any preceding numbered method, wherein the target DNA is a genomic Natural Harbor™ site.

121. Any preceding numbered method, which results in insertion of the heterologous object
25 sequence into the a target site in the genome at an average copy number of at least 0.01, 0.025, 0.05, 0.075, 0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.5, 3, 4, or 5 copies per genome.

122. Any preceding numbered method, which results in about 25-100%, 50-100%, 60-100%,
30 70-100%, 75-95%, 80%-90%, of integrants into a target site in the genome being non-truncated, as measured by an assay described herein, e.g., an assay of Example 6.

123. Any preceding numbered method, which results in insertion of the heterologous object sequence only at one target site in the genome of the cell.

5 124. Any preceding numbered method, which results in insertion of the heterologous object sequence into a target site in a cell, wherein the inserted heterologous sequence comprises less than 10%, 5%, 2%, 1%, 0.5%, 0.2%, or 0.1% mutations (e.g., SNPs or one or more deletions, e.g., truncations or internal deletions) relative to the heterologous sequence prior to insertion, e.g., as measured by an assay of Example 12.

10 125. Any preceding numbered method, which results in insertion of the heterologous object sequence into a target site in a plurality of cells, wherein less than 10%, 5%, 2%, or 1% of copies of the inserted heterologous sequence comprise a mutation (e.g., a SNP or a deletion, e.g., a truncation or an internal deletion), e.g., as measured by an assay of Example 12.

15 126. Any preceding numbered method, which results in insertion of the heterologous object sequence into a target cell genome, and wherein the target cell does not show upregulation of p53, or shows upregulation of p53 by less than 10%, 5%, 2%, or 1%, e.g., wherein upregulation of p53 is measured by p53 protein level, e.g., according to the method described in Example 30,
20 or by the level of p53 phosphorylated at Ser15 and Ser20.

127. Any preceding numbered method, which results in insertion of the heterologous object sequence into a target cell genome, and wherein the target cell does not show upregulation of any DNA repair genes and/or tumor suppressor genes, or wherein no DNA repair gene and/or tumor
25 suppressor gene is upregulated by more than 10%, 5%, 2%, or 1%, e.g., wherein upregulation is measured by RNA-seq, e.g., as described in Example 14.

128. Any preceding numbered method, which results in insertion of the heterologous object sequence into the target site (e.g., at a copy number of 1 insertion or more than one insertion) in
30 about 1-80% of cells in a population of cells contacted with the system, e.g., about: 1-10%, 10-

20%, 20-30%, 30-40%, 40-50%, 50-60%, 60-70%, or 70-80% of cells, e.g., as measured using single cell ddPCR, e.g., as described in Example 17.

129. Any preceding numbered method, which results in insertion of the heterologous object sequence into the target site at a copy number of 1 insertion in about 1-80% of cells in a population of cells contacted with the system, e.g., about: 1-10%, 10-20%, 20-30%, 30-40%, 40-50%, 50-60%, 60-70%, or 70-80% of cells, e.g., as measured using colony isolation and ddPCR, e.g., as described in Example 18.

130. Any preceding numbered method, which results in insertion of the heterologous object sequence into the target site (on-target insertions) at a higher rate than insertion into a non-target site (off-target insertions) in a population of cells, wherein the ratio of on-target insertions to off-target insertions is greater than 10:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, 100:1, 200:1, 500:1, or 1,000:1, e.g., using an assay of Example 11.

131. Any above-numbered method, results in insertion of a heterologous object sequence in the presence of an inhibitor of a DNA repair pathway (e.g., SCR7, a PARP inhibitor), or in a cell line deficient for a DNA repair pathway (e.g., a cell line deficient for the nucleotide excision repair pathway or the homology-directed repair pathway).

132. Any preceding numbered system, formulated as a pharmaceutical composition.

133. Any preceding numbered system, disposed in a pharmaceutically acceptable carrier (e.g., a vesicle, a liposome, a natural or synthetic lipid bilayer, a lipid nanoparticle, an exosome).

134. A method of making a system for modifying the genome of a mammalian cell, comprising:

a) providing a template RNA as described in any of the preceding embodiments, e.g., wherein the template RNA comprises (i) a sequence that binds a polypeptide comprising a reverse transcriptase domain and an endonuclease domain, and (ii) a heterologous object sequence; and

b) treating the template RNA to reduce secondary structure, e.g., heating the template RNA, e.g., to at least 70, 75, 80, 85, 90, or 95 C, and

c) subsequently cooling the template RNA, e.g., to a temperature that allows for secondary structure, e.g, to less than or equal to 30, 25, or 20 C.

5

135. The method of embodiment 134, which further comprises contacting the template RNA with a polypeptide that comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain, or with a nucleic acid (e.g., RNA) encoding the polypeptide.

10 136. The method of embodiment 134 or 135, which further comprises contacting the template RNA with a cell.

137. The system or method of any of the preceding embodiments, wherein the heterologous object sequence encodes a therapeutic polypeptide.

15

138. The system or method of any of the preceding embodiments, wherein the heterologous object sequence encodes a mammalian (e.g., human) polypeptide, or a fragment or variant thereof.

20 139. The system or method of any of the preceding embodiments, wherein the heterologous object sequence encodes an enzyme (e.g., a lysosomal enzyme), a blood factor (e.g., Factor I, II, V, VII, X, XI, XII or XIII), a membrane protein, an exon, an intracellular protein (e.g., a cytoplasmic protein, a nuclear protein, an organellar protein such as a mitochondrial protein or lysosomal protein), an extracellular protein, a structural protein, a signaling protein, a regulatory
25 protein, a transport protein, a sensory protein, a motor protein, a defense protein, or a storage protein.

140. The system or method of any of the preceding embodiments, wherein the heterologous object sequence comprises a tissue specific promoter or enhancer.

30

141. The system or method of any of the preceding embodiments, wherein the heterologous object sequence encodes a polypeptide of greater than 250, 300, 400, 500, or 1,000 amino acids, and optionally up to 1300 amino acids.

5 142. The system or method of any of the preceding embodiments, wherein the heterologous object sequence encodes a fragment of a mammalian gene but does not encode the full mammalian gene, e.g., encodes one or more exons but does not encode a full-length protein.

143. The system or method of any of the preceding embodiments, wherein the heterologous
10 object sequence encodes one or more introns.

144. The system or method of any of the preceding embodiments, wherein the heterologous object sequence is other than a GFP, e.g., is other than a fluorescent protein or is other than a reporter protein.

15 145. The system or method of any of the preceding embodiments, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain, wherein one or both of (i) or (ii) are derived from an avian retrotransposase, e.g., have a sequence of Table 2 or 3 or at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto.

20 146. The system or method of any of the preceding embodiments, wherein the polypeptide has an activity at 37°C that is no less than 70%, 75%, 80%, 85%, 90%, or 95% of its activity at 25°C under otherwise similar conditions.

25 147. The system or method of any of the preceding embodiments, wherein the nucleic acid encoding the polypeptide and the template RNA or a nucleic acid encoding the template RNA are separate nucleic acids.

148. The system or method of any of the preceding embodiments, wherein the template RNA
30 does not encode an active reverse transcriptase, e.g., comprises an inactivated mutant reverse

transcriptase, e.g., as described in Example 1 or 2, or does not comprise a reverse transcriptase sequence.

149. The system or method of any of the preceding embodiments, wherein the template RNA
5 comprises one or more chemical modifications.

150. The system or method of any of the preceding embodiments, wherein the heterologous object sequence is disposed between the promoter and the sequence that binds the polypeptide.

10 151. The system or method of any of the preceding embodiments, wherein the promoter is disposed between the heterologous object sequence and the sequence that binds the polypeptide.

152. The system or method of any of the preceding embodiments, wherein the heterologous object sequence comprises an open reading frame (or the reverse complement thereof) in a 5' to
15 3' orientation on the template RNA.

153. The system or method of any of the preceding embodiments, wherein the heterologous object sequence comprises an open reading frame (or the reverse complement thereof) in a 3' to
20 5' orientation on the template RNA.

154. The system or method of any of the preceding embodiments, wherein the polypeptide comprises (a) a reverse transcriptase domain and (b) an endonuclease domain, wherein at least one of (a) or (b) is heterologous.

25 155. The system or method of any of the preceding embodiments, wherein the polypeptide comprises (a) a target DNA binding domain, (b) a reverse transcriptase domain and (c) an endonuclease domain, wherein at least one of (a), (b) or (c) is heterologous.

156. A substantially pure polypeptide comprising (a) a reverse transcriptase domain and (b) a
30 heterologous endonuclease domain.

157. A substantially pure polypeptide comprising (a) a target DNA binding domain, (b) a reverse transcriptase domain and (c) an endonuclease domain, wherein at least one of (a), (b) or (c) is heterologous.

5 158. A substantially pure polypeptide comprising (a) a reverse transcriptase domain, (b) an endonuclease domain, and (c) a heterologous target DNA binding domain.

159. A polypeptide or a nucleic acid encoding the polypeptide, wherein the polypeptide comprises (a) a reverse transcriptase domain and (b) an endonuclease domain, wherein at least
10 one of (a) or (b) is heterologous to the other.

160. A polypeptide or a nucleic acid encoding the polypeptide, wherein the polypeptide comprises (a) a target DNA binding domain, (b) a reverse transcriptase domain and (c) an endonuclease domain, wherein at least one of (a), (b) or (c) is heterologous to the other.

15

161. Any polypeptide of numbered embodiments 156-160, wherein the reverse transcriptase domain has at least 80% identity (e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identity) to a reverse transcriptase domain of an APE-type or RLE-type non-LTR retrotransposon listed in any of Tables 1-3.

20

162. Any polypeptide of numbered embodiments 156-161, wherein the endonuclease domain has at least 80% identity e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identity, to a endonuclease domain of an APE-type or RLE-type non-LTR retrotransposon listed in any of Tables 1-3.

25

163. Any polypeptide of numbered embodiments 156-162 or any preceding numbered method, wherein the DNA binding domain has at least 80% identity e.g., at least 85%, 90%, 95%, 97%, 98%, 99%, 100% identity, to a DNA binding domain of a sequence listed in Table 1, 2, or 3.

30 164. A nucleic acid encoding the polypeptide of any preceding numbered embodiment.

165. A vector comprising the nucleic acid of numbered embodiment 164.

166. A host cell comprising the nucleic acid of numbered embodiment 164.

5 167. A host cell comprising the polypeptide of any preceding numbered embodiment.

168. A host cell comprising the vector of numbered embodiment 165.

10 169. A host cell (e.g., a human cell) comprising: (i) a heterologous object sequence (e.g., a sequence encoding a therapeutic polypeptide) at a target site in a chromosome, and (ii) one or both of an untranslated region (e.g., a retrotransposon untranslated sequence, e.g., a sequence of column 6 of Table 3) on one side (e.g., upstream) of the heterologous object sequence, and an untranslated region (e.g., a retrotransposon untranslated sequence, e.g., a sequence of column 7 of Table 3) on the other side (e.g., downstream) of the heterologous object sequence.

15

170. A host cell (e.g., a human cell) comprising: (i) a heterologous object sequence (e.g., a sequence encoding a therapeutic polypeptide) at a target site in a chromosome, wherein the target locus is a Natural HarborTM site, e.g., a site of Table 4 herein.

20 171. The host cell of embodiment 170, which further comprises (ii) one or both of an untranslated region 5' of the heterologous object sequence, and an untranslated region 3' of the heterologous object sequence.

25 172. The host cell of embodiment 170, which further comprises (ii) one or both of an untranslated region (e.g., a retrotransposon untranslated sequence, e.g., a sequence of column 6 of Table 3) on one side (e.g., upstream) of the heterologous object sequence, and an untranslated region (e.g., a retrotransposon untranslated sequence, e.g., a sequence of column 7 of Table 3) on the other side (e.g., downstream) of the heterologous object sequence.

30 173. The host cell of any of embodiments 169-173, which comprises heterologous object sequence at only the target site.

174. A pharmaceutical composition, comprising any preceding numbered system, nucleic acid, polypeptide, or vector; and a pharmaceutically acceptable excipient or carrier.

5 175. The pharmaceutical composition of embodiment 174, wherein the pharmaceutically acceptable excipient or carrier is selected from a vector (e.g., a viral or plasmid vector), a vesicle (e.g., a liposome, an exosome, a natural or synthetic lipid bilayer), a lipid nanoparticle.

176. A polypeptide of any of the preceding embodiments, wherein the polypeptide further
10 comprises a nuclear localization sequence.

177. A method of modifying a target DNA strand in a cell, tissue or subject, comprising administering any preceding numbered system to the cell, tissue or subject, thereby modifying the target DNA strand.

15 178. Any preceding numbered embodiment, wherein the polypeptide comprises an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to an amino acid sequence listed in Table 5 (e.g., any one of SEQ ID NOs: 1017-1022), or a functional fragment thereof.

20 179. Any preceding numbered embodiment, wherein the reverse transcriptase domain comprises an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the reverse transcriptase domain of an amino acid sequence listed in Table 5 (e.g., any one of SEQ ID NOs: 1017-1022), or a functional fragment
25 thereof.

180. Any preceding numbered embodiment, wherein the retrotransposase comprises an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to an amino acid sequence listed in Table 5 (e.g., any one of SEQ ID NOs:
30 1017-1022), or a functional fragment thereof.

181. Any preceding numbered embodiment, wherein the polypeptide comprises an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 1023) or GGGS (SEQ ID NO: 1024).

5

182. Any preceding numbered embodiment, wherein the reverse transcriptase domain comprises an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 1023) or GGGS (SEQ ID NO: 1024).

10

183. Any preceding numbered embodiment, wherein the retrotransposase comprises an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 1023) or GGGS (SEQ ID NO: 1024).

15

184. Any preceding numbered embodiment, wherein the polypeptide, reverse transcriptase domain, or retrotransposase comprises a linker comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 1023) or GGGS (SEQ ID NO: 1024).

20

185. Any preceding numbered embodiment, wherein the polypeptide comprises a DNA binding domain covalently attached to the remainder of the polypeptide by a linker, e.g., a linker comprising at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 125, 150, 200, 300, 400, or 500 amino acids.

25

186. Numbered embodiment 185, wherein the linker is attached to the remainder of the polypeptide at a position in the DNA binding domain, RNA binding domain, reverse transcriptase domain, or endonuclease domain (e.g., as shown in any of Figures 17A-17F).

30

187. Numbered embodiment 185 or 186, wherein the linker is attached to the remainder of the polypeptide at a position in the N-terminal side of an alpha helical region of the polypeptide, e.g., at a position corresponding to version v1 as described in Example 26.

5 188. Numbered embodiment 185 or 186, wherein the linker is attached to the remainder of the polypeptide at a position in the C-terminal side of an alpha helical region of the polypeptide, e.g., preceding an RNA binding motif (e.g., a -1 RNA binding motif), e.g., at a position corresponding to version v2 as described in Example 26.

10 189. Numbered embodiment 185 or 186, wherein the linker is attached to the remainder of the polypeptide at a position in the C-terminal side of a random coil region of the polypeptide, e.g., N-terminal relative to a DNA binding motif (e.g., a c-myb DNA binding motif), e.g., at a position corresponding to version v3 as described in Example 26.

15 190. Any one of numbered embodiments 185-189, wherein the linker comprises an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 1023) or GGGS (SEQ ID NO: 1024).

20 191. Any preceding numbered embodiment, wherein a polynucleotide sequence comprising at least about 500, 1000, 2000, 3000, 3500, 3600, 3700, 3800, 3900, or 4000 contiguous nucleotides from the 5' end of the template RNA sequence are integrated into a target cell genome.

25 192. Any preceding numbered embodiment, wherein a polynucleotide sequence comprising at least about 500, 1000, 2000, 2500, 2600, 2700, 2800, 2900, or 3000 contiguous nucleotides from the 3' end of the template RNA sequence are integrated into a target cell genome.

30 193. Any preceding numbered embodiment, wherein the nucleic acid sequence of the template RNA, or a portion thereof (e.g., a portion comprising at least about 100, 200, 300, 400, 500, 1000, 2000, 2500, 3000, 3500, or 4000 nucleotides) integrates into the genomes of a population

of target cells at a copy number of at least about 0.21, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, or 1.0 integrants/genome.

194. Any preceding numbered embodiment, wherein the nucleic acid sequence of the template
 5 RNA, or a portion thereof (e.g., a portion comprising at least about 100, 200, 300, 400, 500, 1000, 2000, 2500, 3000, 3500, or 4000 nucleotides) integrates into the genomes of a population of target cells at a copy number of at least about 0.085, 0.09, 0.1, 0.15, or 0.2 integrants/genome.

195. Any preceding numbered embodiment, wherein the nucleic acid sequence of the template
 10 RNA, or a portion thereof (e.g., a portion comprising at least about 100, 200, 300, 400, 500, 1000, 2000, 2500, 3000, 3500, or 4000 nucleotides) integrates into the genomes of a population of target cells at a copy number of at least about 0.036, 0.04, 0.05, 0.06, 0.07, or 0.08 integrants/genome.

15 196. Any preceding numbered embodiment, wherein the polypeptide comprises a functional endonuclease domain (e.g., wherein the endonuclease domain does not comprise a mutation that abolishes endonuclease activity, e.g., as described herein).

197. Any preceding numbered embodiment, wherein the polypeptide comprises an amino acid
 20 sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the R2 polypeptide from a medium ground finch, e.g., *Geospiza fortis* (e.g., as described herein, e.g., R2-1_GFo), or a functional fragment thereof.

198. Any preceding numbered embodiment, wherein the reverse transcriptase domain
 25 comprises an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the R2 polypeptide from a medium ground finch, e.g., *Geospiza fortis* (e.g., as described herein, e.g., R2-1_GFo), or a functional fragment thereof.

199. Any preceding numbered embodiment, wherein the retrotransposase comprises an amino
 30 acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%

sequence identity to the R2 polypeptide from a medium ground finch, e.g., *Geospiza fortis* (e.g., as described herein, e.g., R2-1_GFo), or a functional fragment thereof.

200. Any one of numbered embodiments 197-199, wherein the nucleic acid sequence of the template RNA, or a portion thereof (e.g., a portion comprising at least about 100, 200, 300, 400, 500, 1000, 2000, 2500, 3000, 3500, or 4000 nucleotides) integrates into the genomes of a population of target cells at a copy number of at least about 0.21 integrants/genome.

201. Any preceding numbered embodiment, wherein the polypeptide comprises an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the R4 polypeptide from a large roundworm, e.g., *Ascaris lumbricoides* (e.g., as described herein, e.g., R4_AL), or a functional fragment thereof.

202. Any preceding numbered embodiment, wherein the reverse transcriptase domain comprises an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the R4 polypeptide from a large roundworm, e.g., *Ascaris lumbricoides* (e.g., as described herein, e.g., R4_AL), or a functional fragment thereof.

203. Any preceding numbered embodiment, wherein the retrotransposase comprises an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the R4 polypeptide from a large roundworm, e.g., *Ascaris lumbricoides* (e.g., as described herein, e.g., R4_AL), or a functional fragment thereof.

204. Any one of numbered embodiments 201-203, wherein the nucleic acid sequence of the template RNA, or a portion thereof (e.g., a portion comprising at least about 100, 200, 300, 400, 500, 1000, 2000, 2500, 3000, 3500, or 4000 nucleotides) integrates into the genomes of a population of target cells at a copy number of at least about 0.085 integrants/genome.

205. Any preceding numbered embodiment, wherein introduction of the system into a target cell does not result in alteration (e.g., upregulation) of p53 and/or p21 protein levels, H2AX

phosphorylation (e.g., gamma H2AX), ATM phosphorylation, ATR phosphorylation, Chk1 phosphorylation, Chk2 phosphorylation, and/or p53 phosphorylation.

206. Any preceding numbered embodiment, wherein introduction of the system into a target cell results in upregulation of p53 protein level in the target cell to a level that is less than about 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 30%, 40%, 45%, 50%, 55%, 60%, 70%, 80%, or 90% of the p53 protein level induced by introducing a site-specific nuclease, e.g., Cas9, that targets the same genomic site as said system.

207. Numbered embodiment 205 or 206, wherein the p53 protein level is determined according to the method described in Example 30.

208. Any preceding numbered embodiment, wherein introduction of the system into a target cell results in upregulation of p53 phosphorylation level in the target cell to a level that is less than about 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 30%, 40%, 45%, 50%, 55%, 60%, 70%, 80%, or 90% of the p53 phosphorylation level induced by introducing a site-specific nuclease, e.g., Cas9, that targets the same genomic site as said system.

209. Any preceding numbered embodiment, wherein introduction of the system into a target cell results in upregulation of p21 protein level in the target cell to a level that is less than about 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 30%, 40%, 45%, 50%, 55%, 60%, 70%, 80%, or 90% of the p53 protein level induced by introducing a site-specific nuclease, e.g., Cas9, that targets the same genomic site as said system.

210. Numbered embodiment 205 or 209, wherein the p21 protein level is determined according to the method described in Example 30.

211. Any preceding numbered embodiment, wherein introduction of the system into a target cell results in upregulation of H2AX phosphorylation level in the target cell to a level that is less than about 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 30%, 40%, 45%, 50%, 55%, 60%, 70%, 80%, or 90% of the H2AX

5 phosphorylation level induced by introducing a site-specific nuclease, e.g., Cas9, that targets the same genomic site as said system.

212. Any preceding numbered embodiment, wherein introduction of the system into a target cell results in upregulation of ATM phosphorylation level in the target cell to a level that is less than about 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 30%, 40%, 45%, 50%, 55%, 60%, 70%, 80%, or 90% of the ATM

10 phosphorylation level induced by introducing a site-specific nuclease, e.g., Cas9, that targets the same genomic site as said system.

213. Any preceding numbered embodiment, wherein introduction of the system into a target cell results in upregulation of ATR phosphorylation level in the target cell to a level that is less than about 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 30%, 40%, 45%, 50%, 55%, 60%, 70%, 80%, or 90% of the ATR

15 phosphorylation level induced by introducing a site-specific nuclease, e.g., Cas9, that targets the same genomic site as said system.

214. Any preceding numbered embodiment, wherein introduction of the system into a target cell results in upregulation of Chk1 phosphorylation level in the target cell to a level that is less than about 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%,

25 10%, 15%, 20%, 30%, 40%, 45%, 50%, 55%, 60%, 70%, 80%, or 90% of the Chk1 phosphorylation level induced by introducing a site-specific nuclease, e.g., Cas9, that targets the same genomic site as said system.

215. Any preceding numbered embodiment, wherein introduction of the system into a target cell results in upregulation of Chk2 phosphorylation level in the target cell to a level that is less than about 0.001%, 0.005%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%,

30

10%, 15%, 20%, 30%, 40%, 45%, 50%, 55%, 60%, 70%, 80%, or 90% of the Chk2 phosphorylation level induced by introducing a site-specific nuclease, e.g., Cas9, that targets the same genomic site as said system.

5 *Definitions*

Domain: The term “domain” as used herein refers to a structure of a biomolecule that contributes to a specified function of the biomolecule. A domain may comprise a contiguous region (e.g., a contiguous sequence) or distinct, non-contiguous regions (e.g., non-contiguous sequences) of a biomolecule. Examples of protein domains include, but are not limited to, an
10 endonuclease domain, a DNA binding domain, a reverse transcription domain; an example of a domain of a nucleic acid is a regulatory domain, such as a transcription factor binding domain.

Exogenous: As used herein, the term exogenous, when used with reference to a biomolecule (such as a nucleic acid sequence or polypeptide) means that the biomolecule was introduced into a host genome, cell or organism by the hand of man. For example, a nucleic acid
15 that is as added into an existing genome, cell, tissue or subject using recombinant DNA techniques or other methods is exogenous to the existing nucleic acid sequence, cell, tissue or subject.

Genomic safe harbor site (GSH site): A genomic safe harbor site is a site in a host genome that is able to accommodate the integration of new genetic material, e.g., such that the
20 inserted genetic element does not cause significant alterations of the host genome posing a risk to the host cell or organism. A GSH site generally meets 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the following criteria: (i) is located >300kb from a cancer-related gene; (ii) is >300kb from a miRNA/other functional small RNA; (iii) is >50kb from a 5' gene end; (iv) is >50kb from a replication origin; (v) is >50kb away from any ultraconserved element; (vi) has low transcriptional activity (i.e.

no mRNA +/- 25 kb); (vii) is not in copy number variable region; (viii) is in open chromatin; and/or (ix) is unique, with 1 copy in the human genome. Examples of GSH sites in the human genome that meet some or all of these criteria include (i) the adeno-associated virus site 1 (AAVS1), a naturally occurring site of integration of AAV virus on chromosome 19; (ii) the chemokine (C-C motif) receptor 5 (CCR5) gene, a chemokine receptor gene known as an HIV-1 coreceptor; (iii) the human ortholog of the mouse Rosa26 locus; (iv) the rDNA locus. Additional GSH sites are known and described, e.g., in Pellenz et al. epub August 20, 2018 (<https://doi.org/10.1101/396390>).

Heterologous: The term heterologous, when used to describe a first element in reference to a second element means that the first element and second element do not exist in nature disposed as described. For example, a heterologous polypeptide, nucleic acid molecule, construct or sequence refers to (a) a polypeptide, nucleic acid molecule or portion of a polypeptide or nucleic acid molecule sequence that is not native to a cell in which it is expressed, (b) a polypeptide or nucleic acid molecule or portion of a polypeptide or nucleic acid molecule that has been altered or mutated relative to its native state, or (c) a polypeptide or nucleic acid molecule with an altered expression as compared to the native expression levels under similar conditions. For example, a heterologous regulatory sequence (e.g., promoter, enhancer) may be used to regulate expression of a gene or a nucleic acid molecule in a way that is different than the gene or a nucleic acid molecule is normally expressed in nature. In another example, a heterologous domain of a polypeptide or nucleic acid sequence (e.g., a DNA binding domain of a polypeptide or nucleic acid encoding a DNA binding domain of a polypeptide) may be disposed relative to other domains or may be a different sequence or from a different source, relative to other domains or portions of a polypeptide or its encoding nucleic acid. In certain embodiments,

a heterologous nucleic acid molecule may exist in a native host cell genome, but may have an altered expression level or have a different sequence or both. In other embodiments, heterologous nucleic acid molecules may not be endogenous to a host cell or host genome but instead may have been introduced into a host cell by transformation (e.g., transfection, electroporation), wherein the added molecule may integrate into the host genome or can exist as extra-chromosomal genetic material either transiently (e.g., mRNA) or semi-stably for more than one generation (e.g., episomal viral vector, plasmid or other self-replicating vector).

Mutation or Mutated: The term “mutated” when applied to nucleic acid sequences means that nucleotides in a nucleic acid sequence may be inserted, deleted or changed compared to a reference (e.g., native) nucleic acid sequence. A single alteration may be made at a locus (a point mutation) or multiple nucleotides may be inserted, deleted or changed at a single locus. In addition, one or more alterations may be made at any number of loci within a nucleic acid sequence. A nucleic acid sequence may be mutated by any method known in the art.

Nucleic acid molecule: Nucleic acid molecule refers to both RNA and DNA molecules including, without limitation, cDNA, genomic DNA and mRNA, and also includes synthetic nucleic acid molecules, such as those that are chemically synthesized or recombinantly produced, such as RNA templates, as described herein. The nucleic acid molecule can be double-stranded or single-stranded, circular or linear. If single-stranded, the nucleic acid molecule can be the sense strand or the antisense strand. Unless otherwise indicated, and as an example for all sequences described herein under the general format “SEQ. ID NO:,” “nucleic acid comprising SEQ. ID NO:1” refers to a nucleic acid, at least a portion which has either (i) the sequence of SEQ. ID NO:1, or (ii) a sequence complimentary to SEQ. ID NO:1. The choice between the two is dictated by the context in which SEQ. ID NO:1 is used. For instance, if the nucleic acid is used

as a probe, the choice between the two is dictated by the requirement that the probe be complimentary to the desired target. Nucleic acid sequences of the present disclosure may be modified chemically or biochemically or may contain non-natural or derivatized nucleotide bases, as will be readily appreciated by those of skill in the art. Such modifications include, for example, labels, methylation, substitution of one or more naturally occurring nucleotides with an analog, inter-nucleotide modifications such as uncharged linkages (for example, methyl phosphonates, phosphotriesters, phosphoramidates, carbamates, etc.), charged linkages (for example, phosphorothioates, phosphorodithioates, etc.), pendant moieties, (for example, polypeptides), intercalators (for example, acridine, psoralen, etc.), chelators, alkylators, and modified linkages (for example, alpha anomeric nucleic acids, etc.). Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions. Such molecules are known in the art and include, for example, those in which peptide linkages substitute for phosphate linkages in the backbone of a molecule. Other modifications can include, for example, analogs in which the ribose ring contains a bridging moiety or other structure such as modifications found in “locked” nucleic acids.

Gene expression unit: a gene expression unit is a nucleic acid sequence comprising at least one regulatory nucleic acid sequence operably linked to at least one effector sequence. A first nucleic acid sequence is operably linked with a second nucleic acid sequence when the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter or enhancer is operably linked to a coding sequence if the promoter or enhancer affects the transcription or expression of the coding sequence. Operably

linked DNA sequences may be contiguous or non-contiguous. Where necessary to join two protein-coding regions, operably linked sequences may be in the same reading frame.

Host: The terms host genome or host cell, as used herein, refer to a cell and/or its genome into which protein and/or genetic material has been introduced. It should be understood that such terms are intended to refer not only to the particular subject cell and/or genome, but to the progeny of such a cell and/or the genome of the progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term “host cell” as used herein. A host genome or host cell may be an isolated cell or cell line grown in culture, or genomic material isolated from such a cell or cell line, or may be a host cell or host genome which composing living tissue or an organism. In some instances, a host cell may be an animal cell or a plant cell, e.g., as described herein. In certain instances, a host cell may be a bovine cell, horse cell, pig cell, goat cell, sheep cell, chicken cell, or turkey cell. In certain instances, a host cell may be a corn cell, soy cell, wheat cell, or rice cell.

Pseudoknot: A “pseudoknot sequence” sequence, as used herein, refers to a nucleic acid (e.g., RNA) having a sequence with suitable self-complementarity to form a pseudoknot structure, e.g., having: a first segment, a second segment between the first segment and a third segment, wherein the third segment is complementary to the first segment, and a fourth segment, wherein the fourth segment is complementary to the second segment. The pseudoknot may optionally have additional secondary structure, e.g., a stem loop disposed in the second segment, a stem-loop disposed between the second segment and third segment, sequence before the first segment, or sequence after the fourth segment. The pseudoknot may have additional sequence

between the first and second segments, between the second and third segments, or between the third and fourth segments. In some embodiments, the segments are arranged, from 5' to 3': first, second, third, and fourth. In some embodiments, the first and third segments comprise five base pairs of perfect complementarity. In some embodiments, the second and fourth segments
5 comprise 10 base pairs, optionally with one or more (e.g., two) bulges. In some embodiments, the second segment comprises one or more unpaired nucleotides, e.g., forming a loop. In some embodiments, the third segment comprises one or more unpaired nucleotides, e.g., forming a loop.

Stem-loop sequence: As used herein, a “stem-loop sequence” refers to a nucleic acid
10 sequence (e.g., RNA sequence) with sufficient self-complementarity to form a stem-loop, e.g., having a stem comprising at least two (e.g., 3, 4, 5, 6, 7, 8, 9, or 10) base pairs, and a loop with at least three (e.g., four) base pairs. The stem may comprise mismatches or bulges.

15 BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic of the Gene Writing genome editing system.

Figure 2 is a schematic of the structure of the Gene Writer genome editor polypeptide.

Figure 3 is a schematic of a Gene Writer genome editor polypeptide comprising a heterologous DNA binding domain designed to target different sites of the genome.

20 Figure 4 is a schematic of the structure of Gene Writer genome editor template RNA.

Figure 5 is a schematic showing the Gene Writing genome editing system to add a gene expression unit into a safe harbor site in the genome.

Figure 6 is a schematic showing Gene Writing genome editing to add a new exon into an specific intron in the genome and replace downstream exons.

Figure 7 illustrates a schematic of Miseq library construction. Nested PCR was performed across the R2Tg-rDNA junction using (1) outer forward primer and tailed inner reverse primer followed by (2) tailed inner forward primer and tail reverse primer. The inner reverse primer contains a 1-4 base stagger, an 8-nucleotide randomized UMI, and a multiplexing barcode. The UMI allows for counting of individual amplification events to eliminate PCR bias.

Figures 8A-8B: Results of Miseq and Matlab analysis of DNA-mediated R2Tg integration into Hek293T cells. Each graph shows analysis of (Fig. 8A) the experimental R2Tg (Fig. 8B) and 1bp deletion negative control. The y-axis indicates aligned counts of unique sequences determined via unique UMIs found via Matlab. The X-axis indicates the sequence position of sequence coverage. The vertical gray line at the left of the graph indicates the position of the forward primer, while the vertical gray line at the right of the graph indicates the expected Tg-rDNA junction site. Bars at the right end of the graph indicate insertion without a truncation, and bars at the left end of the graph indicate truncation. Fig 8A shows that most sequences show high alignment to the expected integration product.

Figure 9 shows a ddPCR evaluation of copy number variation of the R2Tg-rDNA junction in human cells across transfection conditions. Forward primer and probe were expected to bind to the 3' UTR of the R2Tg, while reverse primer was targeted to the human rDNA. The resulting ddPCR signal was normalized to that of reference assay RPP30 to determine copy number. Significantly higher average copies per genome were found with the wildtype (WT, left set of bars) R2Tg as compared to genetic control altering translation with a 1-bp deletion (Frameshift mutant control, right set of bars).

Figure 10 illustrates the sequence alignment and coverage of TOPO cloning the nested PCR product from Example 7. The gray line at the right edge of the graph indicates the expected

transgene-rDNA junction. Most sequences showed high alignment to the expected integrated product.

Figure 11 is a schematic of an exemplary template RNA. It comprises a payload domain in the center (e.g., a heterologous object sequence, e.g., comprising a promoter and a protein-coding sequence). The payload domain is flanked by 5' and 3' protein interaction domains, e.g., sequences capable of binding the Gene Writer polypeptide, e.g., 5' and 3' UTR sequences shown in Table 3. Flanking the protein interaction domains are 5' and 3' homology domains, which have homology to the desired insertion region in the genome.

Figure 12 is a graph showing retrotransposition efficiency measured by ddPCR (digital droplet PCR) using different transfection conditions. Bars A-C represent samples that were transfected using 0.15µl Lipofectamine™ RNAiMAX with 100ng, 250ng, or 500ng respectively. Bars D-F represent samples that were transfected using 0.3µl Lipofectamine™ RNAiMAX with 100ng, 250ng, or 500ng respectively. Bars G-I represent samples that were transfected using 1µl TransIT®-mRNA transfection kit with 100ng, 250ng, or 500ng respectively.

Figure 13. Schematic of trans-transgene delivery machinery. This schematic illustrates a driver plasmid (left) with a pCEP4 backbone, which encodes the reverse transcriptase R2Tg, with a promoter and Kozak sequence upstream, and a polyadenylation signal downstream. The driver plasmid can drive expression of the GeneWriter protein. The transgene plasmid (right), with a pCDNA backbone, comprises (in order) a CMV promoter, an rDNA homology sequence, a 5' UTR, an antisense-orientation insert, a 3' UTR, a second rDNA homology sequence, a second polyadenylation signal, and a TK promoter driving a mKate2 marker. The antisense-orientation insert comprises an EF1α promoter, a coding region for EGFP that comprises an intron, and a polyadenylation signal. Use of the CMV promoter in the transgene plasmid drives

expression of a template RNA comprising the rDNA homology regions, the UTRs, and the antisense-orientation insert.

Figure 14 shows ddPCR evaluation of copy number variation of the transgene-rDNA junction in human cells across transfection conditions. Forward primer and probe were designed to bind to the 3' UTR of the R2Tg, while reverse primer was targeted to the human rDNA. The resulting ddPCR signal was normalized to that of reference assay RPP30 to determine copy number. Significantly higher average copies per genome were found with the wildtype (WT) R2Tg as compared to backbone construct with no R2Tg sequence involved. Condition 1 denotes a driver plasmid : transgene plasmid molar ratio of 9:1; condition 2 denotes the ratio is 4:1, condition 3 denotes the ratio is 1:1, condition 4 denotes the ratio is 1:4, and condition 5 denotes the ratio is 1:9.

Figures 15A and 15B. Fig. 15A: Hybrid capture of R2Tg identified on-target integrations in the human genome. The read coverage as aligned to the expected target integration in the R2 ribosomal site is indicated on the y-axis. The 5' junction between rDNA and R2Tg is indicated by the left vertical line, while the 3' junction is indicated by the right vertical line. Next-generation sequencing identifies reads spanning the expected junctions. Fig. 15B shows the number of reads from this experiment categorized as on-target integration or off-target integration at the 5' end and 3' end of the integrated sequence.

Figure 16. Sanger sequencing result of the 3' junction nested PCR. Lowercase nucleotides represent the designed SNP. Shaded uppercase nucleotides represent WT sequence. Figure 16 discloses SEQ ID NO: 1538.

Figures 17A-17F are schematic diagrams depicting various covalently dimerized Gene Writer protein configurations. The proteins depicted are: Fig. 17A: a wild-type full length

enzyme. Fig. 17B, two full-length enzymes (each comprising a DNA-binding domain, an RNA-binding domain, a reverse transcriptase domain, and an endonuclease domain) connected by a linker. Fig. 17C, a DNA binding domain and an RNA binding domain connected by a linker to a full-length enzyme. Fig. 17D, a DNA-binding domain and an RNA-binding domain connected
5 by a linker to an RNA-binding domain, a reverse transcriptase domain, and an endonuclease domain. Fig. 17E, a DNA-binding domain connected by a first linker to an RNA-binding domain, which is connected by a second linker to a second RNA-binding domain, a reverse transcriptase domain, and an endonuclease domain. Fig. 17F, a DNA-binding domain connected by a first linker to an RNA-binding domain, which is connected by a second linker to a plurality
10 of RNA-binding domains (in this figure, the molecule comprises three RNA-binding domains), which are connected by a linker to a reverse transcriptase domain and an endonuclease domain. In some embodiments, each R2 binds UTRs in the template RNA. In some embodiments, at least one module comprises a reverse transcriptase domain and an endonuclease domain. In some embodiments, the protein comprises a plurality of RNA-binding domains. In some
15 embodiments, the modular system is split and is only active when it binds on DNA where the system uses two different DNA binding modules, e.g., a first protein comprising a first DNA binding module that is fused to an RNA binding module that recruits the RNA template for target primed reverse transcription, and second protein that comprises a second DNA binding module that binds at the site of integration and is fused to the reverse transcription and endonuclease
20 modules. In some embodiments, the nucleic acid encoding the GeneWriter comprises an intein such that the GeneWriter protein is expressed from two separate genes and is fused by protein splicing after being translated. In some embodiments, the GeneWriter is derived from a non-LTR protein, e.g., an R2 protein.

Figures 18A-18F are a schematic diagram showing different modular components of a GeneWriter protein. The proteins depicted are: Fig. 18A: a wild-type full length enzyme. Fig. 18B: the DNA-binding domain of a GeneWriter may comprise zinc fingers, Cas9, or a transcription factor, or a fragment or variant of any of the foregoing. Fig. 18C: the reverse transcriptase domain and RNA-binding domain together may comprise a reverse transcriptase domain (e.g., from an R2 protein) that is heterologous to one or more other domains of the protein, and may optionally further comprise one or more additional RNA binding domains, or a fragment or variant of any of the foregoing. Fig. 18D: the RNA binding domain may comprise, e.g., a B-box protein, an MS2 coat protein, a dCas protein, or a UTR binding protein, or a fragment or variant of any of the foregoing. Fig. 18E: the reverse transcriptase domain may comprise, e.g., a truncated reverse transcriptase domain, e.g., from an R2 protein; a reverse transcriptase domain from a virus (e.g., HIV), or a reverse transcriptase domain from AMV (avian myeloblastosis virus), or a fragment or variant of any of the foregoing. Fig. 18F: the endonuclease domain can comprise, e.g., a Cas9 nickase, a Cas ortholog, Fok I, or a restriction enzyme, or a fragment or variant of any of the foregoing. In some embodiments, a separate DNA binding domain can be attached to a polypeptide described herein (e.g., a DNA binding domain having stronger affinity for the target DNA sequence than an existing or prior DNA binding domain of the polypeptide, or a DNA binding sequence that binds to a different target DNA sequence than the existing or prior DNA binding domain of the polypeptide). In some embodiments, DNA binding domain mutants can be generated, e.g., having increased affinity to the target DNA sequence. In embodiments, the DNA binding domain comprises a zinc finger. In embodiments, the DNA binding domain is attached to the polypeptide (e.g., at the N-terminal or C-terminal ends) via a linker, e.g., as described herein. In embodiments, a zinc finger is

attached to a DNA binding domain mutant (e.g., as described herein), such that the polypeptide exhibits increased binding to the target DNA sequence (e.g., as dictated by the zinc finger) without competition with the rDNA.

Figure 19 is a graph showing linker mutant integration into the genome of HEK293T
5 cells, assessed by a ddPCR assay evaluating copy number of R2Tg integration per genome. In v1 mutants, an insertion is located at the N-terminal side of an alpha helical region of R2Tg that preceded the predicted -1 RNA binding motif; in v2 mutants, an insertion is located at the C-terminal side of an alpha helical region of R2Tg that preceded the predicted -1 RNA binding motif; and in v3 mutants, an insertion is located C-terminal to a random coil region that came
10 after the predicted c-myb DNA binding motif of R2Tg.

Figures 20A-20B are a series of graphs showing long-read sequencing confirming fidelity of R2Tg cis integration. Unique sequence coverage, as determined by UMI, is graphed across the expected reference sequence. The left vertical bar indicates expected 5' junction of the rDNA and R2Tg, while the right vertical bar indicates the 3' junction. Two separate amplicons
15 spanning the 5' junction and 3' junction are shown.

Figures 21A-21B are a series of graphs showing long-read sequencing confirming fidelity of R2Tg cis integration. Unique sequence deletions (>3bp) as determined by UMI is graphed across the expected reference sequence. The left vertical bar indicates expected 5' junction of the rDNA and R2Tg, while the right vertical bar indicates the 3' junction. Two separate amplicons
20 spanning the 5' junction and 3' junction are shown.

Figure 22 is a diagram showing exemplary plasmid map PLV033 for cis integration of R2Gfo.

Figure 23 is a graph showing integration of R2Gfo, R4Al, and R2Tg in cis in HEK293T cells. The mean of four replicates is shown; error bars indicate standard deviation.

Figure 24 is a graph showing that R2Tg integrates into human fibroblasts in cis. Integration efficiency of the wild-type (WT) and endonuclease (EN) control R2Tg were plotted
5 over four replicate experiments as measured via ddPCR at the 3' junction of R2Tg and the rDNA target.

Figure 25 is a diagram showing Western Blot analysis for p53, p21, Actin, and Vinculin. U2OS cells were treated with the indicated compound or plasmid: GFP, R2Tg-WT (wild-type), or R2Tg-EN (endonuclease domain mutant). Plasmid transfections were performed with either
10 lipofectamine 3000 (Lipo) or Fugene HD (Fug). Cells were analyzed 24 hours after treatment or transfection.

DETAILED DESCRIPTION

This disclosure relates to compositions, systems and methods for targeting, editing,
15 modifying or manipulating a DNA sequence (e.g., inserting a heterologous object DNA sequence into a target site of a mammalian genome) at one or more locations in a DNA sequence in a cell, tissue or subject, e.g., *in vivo* or *in vitro*. The object DNA sequence may include, e.g., a coding sequence, a regulatory sequence, a gene expression unit.

More specifically, the disclosure provides retrotransposon-based systems for inserting a
20 sequence of interest into the genome. This disclosure is based, in part, on a bioinformatic analysis to identify retrotransposase sequences and the associated 5' UTR and 3' UTR from a variety of organisms (see Table 3). While not wishing to be bound by theory, in some embodiments, retrotransposases identified in homeothermic (warm blooded) species, like birds, may have improved thermostability relative to some other enzymes that evolved at lower

temperatures, and the thermostable retrotransposases may therefore be better suited for use in human cells. The disclosure also provides experimental evidence that several retrotransposases from different species, e.g., different species of animal and/or different species and clade of retrotransposon (e.g., as grouped by reverse transcriptase phylogeny, e.g., as described in Su et al. (2019) *RNA*; incorporated herein by reference in its entirety), can be used to catalyze DNA insertion into a target site in human cells (see Examples 7 and Example 28).

In some embodiments, systems described herein can have a number of advantages relative to various earlier systems. For instance, the disclosure describes retrotransposases capable of inserting long sequences (e.g., over 3000 nucleotides) of heterologous nucleic acid into a genome (see, e.g., Fig. 20A). In addition, retrotransposases described herein can insert heterologous nucleic acid in an endogenous site in the genome, such as the rDNA locus (see, e.g., Example 7). This is in contrast to Cre/loxP systems which require a first step of inserting an exogenous loxP site before a second step of inserting a sequence of interest into the loxP site.

Gene-writer™ genome editors

Non-long terminal repeat (LTR) retrotransposons are a type of mobile genetic elements that are widespread in eukaryotic genomes. They include two classes: the apurinic/aprimidinic endonuclease (APE)-type and the restriction enzyme-like endonuclease (RLE)-type. The APE class retrotransposons are comprised of two functional domains: an endonuclease/DNA binding domain, and a reverse transcriptase domain. The RLE class are comprised of three functional domains: a DNA binding domain, a reverse transcription domain, and an endonuclease domain. The reverse transcriptase domain of non-LTR retrotransposon functions by binding an RNA sequence template and reverse transcribing it into the host genome's target DNA. The RNA sequence template has a 3' untranslated region which is specifically bound to the transposase,

and a variable 5' region generally having Open Reading Frame(s) ("ORF") encoding transposase proteins. The RNA sequence template may also comprise a 5' untranslated region which specifically binds the retrotransposase.

The inventors have found that, surprisingly, the elements of such non-LTR

5 retrotransposons can be functionally modularized and/or modified to target, edit, modify or manipulate a target DNA sequence, e.g., to insert an object (e.g., heterologous) nucleic acid sequence into a target genome, e.g., a mammalian genome, by reverse transcription. Such modularized and modified nucleic acids, polypeptide compositions and systems are described herein and are referred to as Gene Writer™ gene editors. A Gene Writer™ gene editor system

10 comprises: (A) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain, and either (x) an endonuclease domain that contains DNA binding functionality or (y) an endonuclease domain and separate DNA binding domain; and (B) a template RNA comprising (i) a sequence that binds the polypeptide and (ii) a heterologous insert sequence. For example, the Gene Writer genome editor protein may

15 comprise a DNA-binding domain, a reverse transcriptase domain, and an endonuclease domain. In other embodiments, the Gene Writer genome editor protein may comprise a reverse transcriptase domain and an endonuclease domain. In certain embodiments, the elements of the Gene Writer™ gene editor polypeptide can be derived from sequences of non-LTR retrotransposons, e.g., APE-type or RLE-type retrotransposons or portions or domains thereof.

20 In some embodiments the RLE-type non-LTR retrotransposon is from the R2, NeSL, HERO, R4, or CRE clade. In some embodiments the Gene Writer genome editor is derived from R4 element X4_Line, which is found in the human genome. In some embodiments the APE-type non-LTR retrotransposon is from the R1, or Tx1 clade. In some embodiments the Gene Writer genome

editor is derived from Tx1 element Mare6, which is found in the human genome. The RNA template element of a Gene Writer™ gene editor system is typically heterologous to the polypeptide element and provides an object sequence to be inserted (reverse transcribed) into the host genome. In some embodiments the Gene Writer genome editor protein is capable of target primed reverse transcription.

In some embodiments the Gene Writer genome editor is combined with a second polypeptide. In some embodiments the second polypeptide is derived from an APE-type non-LTR retrotransposon. In some embodiments the second polypeptide has a zinc knuckle-like motif. In some embodiments the second polypeptide is a homolog of Gag proteins.

Polypeptide component of Gene Writer gene editor system

RT domain:

In certain aspects of the present invention, the reverse transcriptase domain of the Gene Writer system is based on a reverse transcriptase domain of an APE-type or RLE-type non-LTR retrotransposon. A wild-type reverse transcriptase domain of an APE-type or RLE-type non-LTR retrotransposon can be used in a Gene Writer system or can be modified (e.g., by insertion, deletion, or substitution of one or more residues) to alter the reverse transcriptase activity for target DNA sequences. In some embodiments the reverse transcriptase is altered from its natural sequence to have altered codon usage, e.g. improved for human cells. In some embodiments the reverse transcriptase domain is a heterologous reverse transcriptase from a different retrovirus, LTR-retrotransposon, or non-LTR retrotransposon. In certain embodiments, a Gene Writer system includes a polypeptide that comprises a reverse transcriptase domain of an RLE-type non-LTR retrotransposon from the R2, NeSL, HERO, R4, or CRE clade, or of an APE-type non-LTR retrotransposon from the R1, or Tx1 clade. In certain embodiments, a Gene Writer system

includes a polypeptide that comprises a reverse transcriptase domain of a retrotransposon listed in Table 1, Table 2, or Table 3. In embodiments, the amino acid sequence of the reverse transcriptase domain of a Gene Writer system is at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identical to the amino acid sequence of a reverse transcriptase domain of a retrotransposon whose DNA sequence is referenced in Table 1, Table 2, or Table 3. A person having ordinary skill in the art is capable of identifying reverse transcription domains based upon homology to other known reverse transcription domains using routine tools as Basic Local Alignment Search Tool (BLAST). In some embodiments, reverse transcriptase domains are modified, for example by site-specific mutation. In embodiments, the reverse transcriptase domain is engineered to bind a heterologous template RNA.

Endonuclease domain:

In certain embodiments, the endonuclease/DNA binding domain of an APE-type retrotransposon or the endonuclease domain of an RLE-type retrotransposon can be used or can be modified (e.g., by insertion, deletion, or substitution of one or more residues) in a Gene Writer system described herein. In some embodiments the endonuclease domain or endonuclease/DNA binding domain is altered from its natural sequence to have altered codon usage, e.g. improved for human cells. In some embodiments the endonuclease element is a heterologous endonuclease element, such as Fok1 nuclease, a type-II restriction l-like endonuclease (RLE-type nuclease), or another RLE-type endonuclease (also known as REL). In some embodiments the heterologous endonuclease activity has nickase activity and does not form double stranded breaks. The amino acid sequence of an endonuclease domain of a Gene

Writer system described herein may be at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% identical to the amino acid sequence of an endonuclease domain of a retrotransposon whose DNA sequence is referenced in Table 1, 2, or 3. A person having ordinary skill in the art is capable of identifying endonuclease domains based upon homology to other known endonuclease domains using tools as Basic Local Alignment Search Tool (BLAST). In certain embodiments, the heterologous endonuclease is Fok1 or a functional fragment thereof. In certain embodiments, the heterologous endonuclease is a Holliday junction resolvase or homolog thereof, such as the Holliday junction resolving enzyme from *Sulfolobus solfataricus*—SsoI Hje (Govindaraju et al., *Nucleic Acids Research* 44:7, 2016). In certain embodiments, the heterologous endonuclease is the endonuclease of the large fragment of a spliceosomal protein, such as Prp8 (Mahbub et al., *Mobile DNA* 8:16, 2017). For example, a Gene Writer polypeptide described herein may comprise a reverse transcriptase domain from an APE- or RLE-type retrotransposon and an endonuclease domain that comprises Fok1 or a functional fragment thereof. In still other embodiments, homologous endonuclease domains are modified, for example by site-specific mutation, to alter DNA endonuclease activity. In still other embodiments, endonuclease domains are modified to remove any latent DNA-sequence specificity.

DNA binding domain:

In certain aspects, the DNA-binding domain of a Gene Writer polypeptide described herein is selected, designed, or constructed for binding to a desired host DNA target sequence. In certain embodiments, the DNA-binding domain of the engineered RLE is a heterologous DNA-binding protein or domain relative to a native retrotransposon sequence. In some embodiments

the heterologous DNA binding element is a zinc-finger element or a TAL effector element, e.g., a zinc-finger or TAL polypeptide or functional fragment thereof. In some embodiments the heterologous DNA binding element is a sequence-guided DNA binding element, such as Cas9, Cpf1, or other CRISPR-related protein that has been altered to have no endonuclease activity. In some embodiments the heterologous DNA binding element retains endonuclease activity. In some embodiments the heterologous DNA binding element replaces the endonuclease element of the polypeptide. In specific embodiments, the heterologous DNA-binding domain can be any one or more of Cas9, TAL domain, ZF domain, Myb domain, combinations thereof, or multiples thereof. In certain embodiments, the heterologous DNA-binding domain is a DNA binding domain of a retrotransposon described in Table 1, Table 2, or Table 3. A person having ordinary skill in the art is capable of identifying DNA binding domains based upon homology to other known DNA binding domains using tools as Basic Local Alignment Search Tool (BLAST). In still other embodiments, DNA-binding domains are modified, for example by site-specific mutation, increasing or decreasing DNA-binding elements (for example, number and/or specificity of zinc fingers), etc., to alter DNA-binding specificity and affinity. In some embodiments the DNA binding domain is altered from its natural sequence to have altered codon usage, e.g. improved for human cells

In certain aspects of the present invention, the host DNA-binding site integrated into by the Gene Writer system can be in a gene, in an intron, in an exon, an ORF, outside of a coding region of any gene, in a regulatory region of a gene, or outside of a regulatory region of a gene. In other aspects, the engineered RLE may bind to one or more than one host DNA sequence.

In certain embodiments, a Gene Writer™ gene editor system RNA further comprises an intracellular localization sequence, e.g., a nuclear localization sequence. The nuclear localization

sequence may be an RNA sequence that promotes the import of the RNA into the nucleus. In certain embodiments the nuclear localization signal is located on the template RNA. In certain embodiments, the retrotransposase polypeptide is encoded on a first RNA, and the template RNA is a second, separate, RNA, and the nuclear localization signal is located on the template RNA and not on an RNA encoding the retrotransposase polypeptide. While not wishing to be bound by theory, in some embodiments, the RNA encoding the retrotransposase is targeted primarily to the cytoplasm to promote its translation, while the template RNA is targeted primarily to the nucleus to promote its retrotransposition into the genome. In some embodiments the nuclear localization signal is at the 3' end, 5' end, or in an internal region of the template RNA. In some embodiments the nuclear localization signal is 3' of the heterologous sequence (e.g., is directly 3' of the heterologous sequence) or is 5' of the heterologous sequence (e.g., is directly 5' of the heterologous sequence). In some embodiments the nuclear localization signal is placed outside of the 5' UTR or outside of the 3' UTR of the template RNA. In some embodiments the nuclear localization signal is placed between the 5' UTR and the 3' UTR, wherein optionally the nuclear localization signal is not transcribed with the transgene (e.g., the nuclear localization signal is an anti-sense orientation or is downstream of a transcriptional termination signal or polyadenylation signal). In some embodiments the nuclear localization sequence is situated inside of an intron. In some embodiments a plurality of the same or different nuclear localization signals are in the RNA, e.g., in the template RNA. In some embodiments the nuclear localization signal is less than 5, 10, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900 or 1000 bp in length. Various RNA nuclear localization sequences can be used. For example, Lubelsky and Ulitsky, *Nature* 555 (107-111), 2018 describe RNA sequences which drive RNA localization into the nucleus. In some embodiments, the nuclear localization signal is a SINE-derived nuclear

RNA localization (SIRLOIN) signal. In some embodiments the nuclear localization signal binds a nuclear-enriched protein. In some embodiments the nuclear localization signal binds the HNRNPK protein. In some embodiments the nuclear localization signal is rich in pyrimidines, e.g., is a C/T rich, C/U rich, C rich, T rich, or U rich region. In some embodiments the nuclear localization signal is derived from a long non-coding RNA. In some embodiments the nuclear localization signal is derived from MALAT1 long non-coding RNA or is the 600 nucleotide M region of MALAT1 (described in Miyagawa et al., *RNA* 18, (738-751), 2012). In some embodiments the nuclear localization signal is derived from BORG long non-coding RNA or is a AGCCC motif (described in Zhang et al., *Molecular and Cellular Biology* 34, 2318-2329 (2014)). In some embodiments the nuclear localization sequence is described in Shukla et al., *The EMBO Journal* e98452 (2018). In some embodiments the nuclear localization signal is derived from a non-LTR retrotransposon, an LTR retrotransposon, retrovirus, or an endogenous retrovirus.

In certain embodiments, a Gene Writer™ gene editor system polypeptide further comprises an intracellular localization sequence, e.g., a nuclear localization sequence and/or a nucleolar localization sequence. The nuclear localization sequence and/or nucleolar localization sequence may be amino acid sequences that promote the import of the protein into the nucleus and/or nucleolus, where it can promote integration of heterologous sequence into the genome. In certain embodiments, a Gene Writer gene editor system polypeptide (e.g., a retrotransposase, e.g., a polypeptide according to any of Tables 1, 2, or 3 herein) further comprises a nucleolar localization sequence. In certain embodiments, the retrotransposase polypeptide is encoded on a first RNA, and the template RNA is a second, separate, RNA, and the nucleolar localization signal is encoded on the RNA encoding the retrotransposase polypeptide and not on the template RNA. In some embodiments, the nucleolar localization signal is located at the N-terminus, C-

terminus, or in an internal region of the polypeptide. In some embodiments, a plurality of the same or different nucleolar localization signals are used. In some embodiments, the nuclear localization signal is less than 5, 10, 25, 50, 75, or 100 amino acids in length. Various polypeptide nucleolar localization signals can be used. For example, Yang et al., *Journal of Biomedical Science* 22, 33 (2015), describe a nuclear localization signal that also functions as a nucleolar localization signal. In some embodiments, the nucleolar localization signal may also be a nuclear localization signal. In some embodiments, the nucleolar localization signal may overlap with a nuclear localization signal. In some embodiments, the nucleolar localization signal may comprise a stretch of basic residues. In some embodiments, the nucleolar localization signal may be rich in arginine and lysine residues. In some embodiments, the nucleolar localization signal may be derived from a protein that is enriched in the nucleolus. In some embodiments, the nucleolar localization signal may be derived from a protein enriched at ribosomal RNA loci. In some embodiments, the nucleolar localization signal may be derived from a protein that binds rRNA. In some embodiments, the nucleolar localization signal may be derived from MSP58. In some embodiments, the nucleolar localization signal may be a monopartite motif. In some embodiments, the nucleolar localization signal may be a bipartite motif. In some embodiments, the nucleolar localization signal may consist of a multiple monopartite or bipartite motifs. In some embodiments, the nucleolar localization signal may consist of a mix of monopartite and bipartite motifs. In some embodiments, the nucleolar localization signal may be a dual bipartite motif. In some embodiments, the nucleolar localization motif may be a KRASSQALGTIPKRRSSSRFIKRKK (SEQ ID NO: 1530). In some embodiments, the nucleolar localization signal may be derived from nuclear factor- κ B-inducing kinase. In some

embodiments, the nucleolar localization signal may be an RKKRKKK motif (SEQ ID NO: 1531) (described in Birbach et al., Journal of Cell Science, 117 (3615-3624), 2004).

In some embodiments, a nucleic acid described herein (e.g., an RNA encoding a GeneWriter polypeptide, or a DNA encoding the RNA) comprises a microRNA binding site. In some embodiments, the microRNA binding site is used to increase the target-cell specificity of a GeneWriter system. For instance, the microRNA binding site can be chosen on the basis that is recognized by a miRNA that is present in a non-target cell type, but that is not present (or is present at a reduced level relative to the non-target cell) in a target cell type. Thus, when the RNA encoding the GeneWriter polypeptide is present in a non-target cell, it would be bound by the miRNA, and when the RNA encoding the GeneWriter polypeptide is present in a target cell, it would not be bound by the miRNA (or bound but at reduced levels relative to the non-target cell). While not wishing to be bound by theory, binding of the miRNA to the RNA encoding the GeneWriter polypeptide may reduce production of the GeneWriter polypeptide, e.g., by degrading the mRNA encoding the polypeptide or by interfering with translation. Accordingly, the heterologous object sequence would be inserted into the genome of target cells more efficiently than into the genome of non-target cells. A system having a microRNA binding site in the RNA encoding the GeneWriter polypeptide (or encoded in the DNA encoding the RNA) may also be used in combination with a template RNA that is regulated by a second microRNA binding site, e.g., as described herein in the section entitled “Template RNA component of Gene Writer™ gene editor system.”

TABLE 1: APE-type non-LTR retrotransposon elements

Table 1

Family	Sequence Accession	Mobile Element	Name	Organism
Dewa	AB097143	ORF2	Danio rerio retrotransposon DewaDr1 DNA, complete sequence	Danio rerio
HeT-A	KJ081250	non-LTR retrotransposon:HeT-A	Drosophila melanogaster non-LTR retrotransposon HeT-A, partial sequence	Drosophila melanogaster
Keno	AB111948	ORF2	Tetraodon nigroviridis retrotransposon KenoTn1 DNA, partial sequence	Tetraodon nigroviridis
KenoDr1;Keno	AB097144	ORF2	Danio rerio retrotransposon KenoDr1 DNA, complete sequence	Danio rerio
KenoFr1;Keno	AB111947	ORF2	Takifugu rubripes retrotransposon KenoFr1 DNA, complete sequence	Takifugu rubripes
Kibi	AB097139	ORF2	Danio rerio retrotransposon KibiDr2 DNA, complete sequence	Danio rerio
Kibi	AB097138	ORF2	Danio rerio retrotransposon KibiDr1 DNA, complete sequence	Danio rerio
Kibi	AB097137	ORF2	Tetraodon nigroviridis retrotransposon KibiTn1 DNA, complete sequence	Tetraodon nigroviridis
Kibi	AB097136	ORF2	Takifugu rubripes retrotransposon KibiFr1 DNA, complete sequence	Takifugu rubripes
KoshiTn1	AB097135	ORF2	Tetraodon nigroviridis retrotransposon KoshiTn1 DNA, complete sequence	Tetraodon nigroviridis
Mutsu	AB097142	ORF2	Danio rerio retrotransposon MutsuDr3 DNA, partial sequence	Danio rerio
Mutsu	AB097141	ORF2	Danio rerio retrotransposon MutsuDr2 DNA, partial sequence	Danio rerio
Mutsu	AB097140	ORF2	Danio rerio retrotransposon MutsuDr1 DNA, complete sequence	Danio rerio
R1	HQ284568	non-LTR retrotransposon:R1-like	Trilocha sp. GAS-2011 isolate TrilSp.6 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Trilocha sp. GAS-2011
R1	HQ284534	non-LTR retrotransposon:R1-like	Scopula ornata isolate ScoOrn.6 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Scopula ornata
R1	HQ284496	non-LTR retrotransposon:R1-like	Perigonia ilus isolate PerIllus.31 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Perigonia ilus
R1	HQ284489	non-LTR retrotransposon:R1-like	Oxytenis modestia isolate OxyMod.2_3_4_7_9 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Oxytenis modestia
R1	HQ284488	non-LTR retrotransposon:R1-like	Oxytenis modestia isolate OxyMod.1 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Oxytenis modestia

Table 1

Family	Sequence Accession	Mobile Element	Name	Organism
R1	HQ284476	non-LTR retrotransposon:R1-like	Oeneis magna dubia isolate OenMag.26 non-LTR retrotransposon R1-like reverse transcriptase-like gene, partial sequence	Oeneis magna dubia
R1	HQ284437	non-LTR retrotransposon:R1-like	Lymantria dispar isolate LymDis.2 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Lymantria dispar
R1	HQ284435	non-LTR retrotransposon:R1-like	Lymantria dispar isolate LymDis.1 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Lymantria dispar
R1	HQ284432	non-LTR retrotransposon:R1-like	Janiodes laverna isolate JanLav.911 non-LTR retrotransposon R1-like reverse transcriptase-like gene, partial sequence	Janiodes laverna
R1	HQ284431	non-LTR retrotransposon:R1-like	Janiodes laverna isolate JanLav.811 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Janiodes laverna
R1	HQ284430	non-LTR retrotransposon:R1-like	Janiodes laverna isolate JanLav.5 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Janiodes laverna
R1	HQ284428	non-LTR retrotransposon:R1-like	Janiodes laverna isolate JanLav.411 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Janiodes laverna
R1	HQ284426	non-LTR retrotransposon:R1-like	Janiodes laverna isolate JanLav.211 non-LTR retrotransposon R1-like reverse transcriptase-like gene, partial sequence	Janiodes laverna
R1	HQ284421	non-LTR retrotransposon:R1-like	Heteropterus morpheus isolate HetMor.3 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Heteropterus morpheus
R1	HQ284402	non-LTR retrotransposon:R1-like	Erinnyis ello isolate EriEllo.22 non-LTR retrotransposon R1-like reverse transcriptase-like gene, partial sequence	Erinnyis ello
R1	HQ284399	non-LTR retrotransposon:R1-like	Erebia theano isolate EreThe.29 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Erebia theano
R1	HQ284398	non-LTR retrotransposon:R1-like	Erebia theano isolate EreThe.28 non-LTR retrotransposon R1-like reverse transcriptase-like gene, partial sequence	Erebia theano

Table 1				
Family	Sequence Accession	Mobile Element	Name	Organism
R1	HQ284397	non-LTR retrotransposon:R1-like	Erebia theano isolate EreThe.27 non-LTR retrotransposon R1-like reverse transcriptase-like gene, partial sequence	Erebia theano
R1	HQ284391	non-LTR retrotransposon:R1-like	Emesis lucinda isolate EmeLuc.23 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Emesis lucinda
R1	HQ284390	non-LTR retrotransposon:R1-like	Emesis lucinda isolate EmeLuc.2 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Emesis lucinda
R1	HQ284364	non-LTR retrotransposon:R1-like	Coenonympha glycerion isolate CoeGly.9 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Coenonympha glycerion
R1	HQ284363	non-LTR retrotransposon:R1-like	Coenonympha glycerion isolate CoeGly.8 non-LTR retrotransposon R1-like reverse transcriptase-like gene, partial sequence	Coenonympha glycerion
R1	HQ284362	non-LTR retrotransposon:R1-like	Coenonympha glycerion isolate CoeGly.7 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Coenonympha glycerion
R1	HQ284361	non-LTR retrotransposon:R1-like	Coenonympha glycerion isolate CoeGly.5 non-LTR retrotransposon R1-like reverse transcriptase-like gene, partial sequence	Coenonympha glycerion
R1	HQ284357	non-LTR retrotransposon:R1-like	Coenonympha glycerion isolate CoeGly.13 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Coenonympha glycerion
R1	HQ284356	non-LTR retrotransposon:R1-like	Coenonympha glycerion isolate CoeGly.11 non-LTR retrotransposon R1-like reverse transcriptase-like gene, partial sequence	Coenonympha glycerion
R1	HQ284350	non-LTR retrotransposon:R1-like	Catocyclotis adelina isolate CatAde.18 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Catocyclotis adelina
R1	HQ284340	non-LTR retrotransposon:R1-like	Caria rhacotis isolate CarRha.11 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Caria rhacotis
R1	HQ284339	non-LTR retrotransposon:R1-like	Caria rhacotis isolate CarRha.1 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Caria rhacotis
R1	HQ284319	non-LTR retrotransposon:R1-like	Archiearis parthenias isolate BrePar.1 non-LTR retrotransposon R1-like	Archiearis parthenias

Table 1				
Family	Sequence Accession	Mobile Element	Name	Organism
			reverse transcriptase gene, partial cds	
R1	HQ284318	non-LTR retrotransposon:R1-like	Brangas neora isolate BraNeo.32 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Brangas neora
R1	HQ284292	non-LTR retrotransposon:R1-like	Araschnia levana isolate AraLev.31 non-LTR retrotransposon R1-like reverse transcriptase-like gene, partial sequence	Araschnia levana
R1	HQ284286	non-LTR retrotransposon:R1-like	Araschnia levana isolate AraLev.1 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Araschnia levana
R1	HQ284280	non-LTR retrotransposon:R1-like	Anteros formosus isolate AntForm.34 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Anteros formosus
R1	HQ284279	non-LTR retrotransposon:R1-like	Anteros formosus isolate AntForm.32 non-LTR retrotransposon R1-like reverse transcriptase-like gene, partial sequence	Anteros formosus
R1	HQ284278	non-LTR retrotransposon:R1-like	Anteros formosus isolate AntForm.31 non-LTR retrotransposon R1-like reverse transcriptase-like gene, partial sequence	Anteros formosus
R1	HQ284270	non-LTR retrotransposon:R1-like	Agrotis exclamationis isolate AgrExcl.27 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Agrotis exclamationis
R1	HQ284267	non-LTR retrotransposon:R1-like	Agrius cingulata isolate AgrCing.36_39 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Agrius cingulata
R1	HQ284266	non-LTR retrotransposon:R1-like	Agrius cingulata isolate AgrCing.3 non-LTR retrotransposon R1-like reverse transcriptase-like gene, partial sequence	Agrius cingulata
R1	HQ284263	non-LTR retrotransposon:R1-like	Aglia tau isolate AglTau.8 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Aglia tau
R1	HQ284262	non-LTR retrotransposon:R1-like	Aglia tau isolate AglTau.7 non-LTR retrotransposon R1-like reverse transcriptase gene, partial cds	Aglia tau
R1	DQ836362	MalR1	Maculinea alcon R1-like non-LTR retrotransposon R1 reverse transcriptase (RT) pseudogene, partial sequence	Phengaris alcon

Table 1

Family	Sequence Accession	Mobile Element	Name	Organism
R1	DQ836391	MnaR1	Maculinea nausithous R1-like non-LTR retrotransposon R1 reverse transcriptase (RT) gene, partial cds	Phengaris nausithous
R1	KU543683	non-LTR retrotransposon and non-LTR retrovirus reverse transcriptase; Region: RT_nLTR	Bactrocera tryoni clone Btry_5404 non-LTR retrotransposon R1, complete sequence	Bactrocera tryoni
R1	KU543682	non-LTR retrotransposon and non-LTR retrovirus reverse transcriptase; Region: RT_nLTR	Bactrocera tryoni clone Btry_5167 non-LTR retrotransposon R1, complete sequence	Bactrocera tryoni
R1	KU543679	non-LTR retrotransposon and non-LTR retrovirus reverse transcriptase; Region: RT_nLTR	Bactrocera tryoni clone Btry_4956 non-LTR retrotransposon R1, complete sequence	Bactrocera tryoni
R1	KU543678	non-LTR retrotransposon and non-LTR retrovirus reverse transcriptase; Region: RT_nLTR	Bactrocera tryoni clone Btry_5979 non-LTR retrotransposon R1, complete sequence	Bactrocera tryoni
R1	AB078933	ORF1	Papilio xuthus non-LTR retrotransposon gene for gag-like protein, partial cds, clone: SARTPx2-2	Papilio xuthus
R1	AB078932	ORF1	Papilio xuthus non-LTR retrotransposon gene for gag-like protein, partial cds, clone: SARTPx2-1	Papilio xuthus
R1	AB078936	ORF2	Papilio xuthus non-LTR retrotransposon genes for gag-like protein, reverse transcriptase, partial cds, clone: SARTPx4-N18	Papilio xuthus
R1	AB078935	ORF2	Papilio xuthus non-LTR retrotransposon genes for gag-like protein, reverse transcriptase, partial and complete cds, clone: SARTPx3-N7	Papilio xuthus
R1	AB078934	ORF2	Papilio xuthus non-LTR retrotransposon genes for gag-like protein, reverse transcriptase, partial cds, clone: SARTPx3-N3	Papilio xuthus
R1	AB078931	ORF2	Papilio xuthus non-LTR retrotransposon genes for gag-like protein, reverse transcriptase, partial and complete cds, clone: SARTPx1-N14	Papilio xuthus
R1	AB078930	ORF2	Papilio xuthus non-LTR retrotransposon genes for gag-like	Papilio xuthus

Table 1

Family	Sequence Accession	Mobile Element	Name	Organism
			protein, reverse transcriptase, partial and complete cds, clone: SARTPx1-N5	
R1	AB078929	ORF2	Papilio xuthus non-LTR retrotransposon genes for gag-like protein, reverse transcriptase, partial and complete cds, clone: SARTPx1-N4	Papilio xuthus
R1	AB078928	ORF2	Papilio xuthus non-LTR retrotransposon gene for gag-like protein, reverse transcriptase, complete and partial cds, clone: SARTPx1-3	Papilio xuthus
R1	KP771712	ORF2; contains endonuclease, reverse transcriptase and RNaseH	Blattella germanica non-LTR retrotransposon TRAS-like 2, complete sequence	Blattella germanica
R1	KP771711	ORF2; contains endonuclease, reverse transcriptase and RNaseH	Blattella germanica non-LTR retrotransposon TRAS-like 1, complete sequence	Blattella germanica
R1	AF015813	R1 ORF	Dugesia sp. retrotransposon R1 reverse transcriptase gene, partial cds	Aphonopelma sp. WDB-1998
R1	AF015489	R1 ORF	Dugesia sp. retrotransposon R1 reverse transcriptase gene, partial cds	Aphonopelma sp. WDB-1998
R1Bm	AB182560	non-LTR retrotransposon R1Bmks ORF2	Bombyx mori genes for non-LTR retrotransposon R1Bmks ORF1 protein, non-LTR retrotransposon R1Bmks ORF2 protein, complete cds	Bombyx mori
R6	AB090819	ORF2	Anopheles gambiae retrotransposon R6Ag3 DNA, complete sequence	Anopheles gambiae
R6	AB090818	ORF2	Anopheles gambiae retrotransposon R6Ag2 DNA, complete sequence	Anopheles gambiae
R6	AB090817	ORF2	Anopheles gambiae retrotransposon R6Ag1 DNA, complete sequence	Anopheles gambiae
R6	KJ958615	R2	Bacillus rossius non-LTR retrotransposon reVIR6, partial sequence	Bacillus rossius
R6	KJ958596	R2	Bacillus rossius non-LTR retrotransposon reBER6, partial sequence	Bacillus rossius
R6	AF352480	transposon:NLRCth1-like non-LTR retrotransposon	Chironomus circumdatus clone cir6 transposon NLRCth1-like non-LTR retrotransposon reverse transcriptase gene, partial cds	Chironomus circumdatus
R6	AF373367	transposon:non-LTR retrotransposon LINE2	Clelia rustica clone CR6 non-LTR retrotransposon LINE2 reverse	Paraphimophis rusticus

Table 1				
Family	Sequence Accession	Mobile Element	Name	Organism
			transcriptase pseudogene, partial sequence	
R7	AB090820	ORF2	Anopheles gambiae retrotransposon R7Ag1 DNA, complete sequence	Anopheles gambiae
R7	AB090821	ORF2	Anopheles gambiae retrotransposon R7Ag2 DNA, complete sequence	Anopheles gambiae
R7	KJ958622	R2	Bacillus rossius non-LTR retrotransposon trKOR7, partial sequence	Bacillus rossius
R7	KJ958616	R2	Bacillus rossius non-LTR retrotransposon reVIR7, partial sequence	Bacillus rossius
R7	KJ958597	R2	Bacillus rossius non-LTR retrotransposon reBER7, partial sequence	Bacillus rossius
R7	AF352514	transposon:NLRCh1-like non-LTR retrotransposon	Chironomus circumdatus clone cir7 transposon NLRCh1-like non-LTR retrotransposon reverse transcriptase pseudogene, partial sequence	Chironomus circumdatus
Rt2	AY379084	truncated; similar to reverse transcriptase	Leptocheirus plumulosus retrotransposon LpRt2, partial sequence	Leptocheirus plumulosus
Rt2	MSQRT2RET		Anopheles gambiae retrotransposon RT2, complete sequence	Anopheles gambiae
RTAg4	AB090813	ORF2	Anopheles gambiae retrotransposon RTAg4 DNA, complete sequence	Anopheles gambiae
TRAS1	BMOTRAS1	DNA binding domain at AA1103-1120.	Bombyx mori gene, complete sequence of retrotransposon TRAS1	Bombyx mori
TRAS3	JX875955	similar to reverse transcriptases	Acyrtosiphon pisum clone LSR1 non-LTR retrotransposon TRAS3, complete sequence	Acyrtosiphon pisum
Tx1	AJ621359	transposon:non-LTR retrotransposon TX1-1_Tet	Tetraodon nigroviridis non-LTR retrotransposon TX1-1_Tet, complete sequence	Tetraodon nigroviridis
Tx1	AJ621360	transposon:non-LTR retrotransposon TX1-2_Tet	Tetraodon nigroviridis partial non-LTR retrotransposon TX1-2_Tet	Tetraodon nigroviridis
Tx1	AJ621361	transposon:non-LTR retrotransposon TX1-3_Tet	Tetraodon nigroviridis partial non-LTR retrotransposon TX1-3_Tet	Tetraodon nigroviridis
Tx1	AJ621362	transposon:non-LTR retrotransposon TX1-4_Tet	Tetraodon nigroviridis partial non-LTR retrotransposon TX1-4_Tet	Tetraodon nigroviridis
Tx1	DQ118004	transposon:Tx1-like retrotransposon Tx1Aru	Acipenser ruthenus clone dg194 transposon Tx1-like retrotransposon Tx1Aru reverse transcriptase-like gene, partial sequence	Acipenser ruthenus

Table 1

Family	Sequence Accession	Mobile Element	Name	Organism
Tx1	AB097134	ORF2	Takifugu rubripes retrotransposon KoshiFr1 DNA, complete sequence	Takifugu rubripes
Tx1	AB090816	ORF2	Anopheles gambiae retrotransposon MinoAg1 DNA, complete sequence	Anopheles gambiae
Tx1	AB090812	ORF2	Anopheles gambiae retrotransposon RTAg3 DNA, complete sequence	Anopheles gambiae
Waldo	AH009917	non-LTR retrotransposon:Waldo-A	Drosophila melanogaster Waldo-A non-LTR retrotransposon, 5' sequence	Drosophila melanogaster
Waldo	AH009916	non-LTR retrotransposon:Waldo-A	Drosophila melanogaster clone CBE9 Waldo-A non-LTR retrotransposon, 5' sequence	Drosophila melanogaster
Waldo	AH009915	non-LTR retrotransposon:Waldo-A	Drosophila melanogaster Waldo-A non-LTR retrotransposon, 5' sequence	Drosophila melanogaster
Waldo	AH009914	non-LTR retrotransposon:Waldo-A	Drosophila melanogaster Waldo-A non-LTR retrotransposon	Drosophila melanogaster
Waldo	AH009920	non-LTR retrotransposon:Waldo-B	Drosophila melanogaster Waldo-B non-LTR retrotransposon, 5' sequence	Drosophila melanogaster
Waldo	AH009919	non-LTR retrotransposon:Waldo-B		Drosophila melanogaster
Waldo	AH009918	non-LTR retrotransposon:Waldo-B		Drosophila melanogaster
Waldo	AB090815	ORF2	Anopheles gambiae retrotransposon WaldoAg2 DNA, complete sequence	Anopheles gambiae
Waldo	AB090814	ORF2	Anopheles gambiae retrotransposon WaldoAg1 DNA, complete sequence	Anopheles gambiae
Waldo	AB078939	ORF2	Forficula scudderii non-LTR retrotransposon pseudogene for reverse transcriptase, clone: WaldoFs1-26	Forficula scudderii
Waldo	AB078938	ORF2	Forficula scudderii non-LTR retrotransposon pseudogene for reverse transcriptase, clone: WaldoFs1-2	Forficula scudderii
Waldo	AB078937	ORF2	Forficula scudderii non-LTR retrotransposon pseudogene for reverse transcriptase, clone: WaldoFs1-1	Forficula scudderii

TABLE 2: RLE-type non-LTR retrotransposon elements

Table 2				
Family	Accession	Mobile Element	Name/Description	Organism
CRE	EF067892		Colletotrichum cereale clone 9F8-1558 Cret3 non-LTR retrotransposon, partial sequence	Colletotrichum cereale
CRE	EF067894		Colletotrichum cereale clone 9F8-2137 Cret3 non-LTR retrotransposon, partial sequence	Colletotrichum cereale
CRE	MG028000	non-LTR retrotransposon:Rex3	Characidium gomesi voucher MNRJ20998 non-LTR retrotransposon Rex3, partial sequence	Characidium gomesi
CRE	KY566213	non-LTR retrotransposon:Rex3	Characidium gomesi non-LTR retrotransposon Rex3, partial sequence	Characidium gomesi
CRE	GU949558		Kalotermes flavicollis isolate Crete non-LTR retrotransposon R2, complete sequence; and R2 protein gene, complete cds	Kalotermes flavicollis
CRE;CRE2	CFU19151	poly dA tracts in 5' and 3' UTRs	Crithidia fasciculata retrotransposon CRE2 in mini-exon gene, putative reverse transcriptase gene, complete cds	Crithidia fasciculata
CZAR	BR000987	pol	TPA_inf: Capsaspora owczarzaki DNA, non-LTR retrotransposon CoL4, complete sequence, strain: ATCC 30864	Capsaspora owczarzaki
CZAR	BR000986	pol	TPA_inf: Capsaspora owczarzaki DNA, non-LTR retrotransposon CoL3, complete sequence, strain: ATCC 30864	Capsaspora owczarzaki
CZAR	BR000985	pol	TPA_inf: Capsaspora owczarzaki DNA, non-LTR retrotransposon CoL2, complete sequence, strain: ATCC 30864	Capsaspora owczarzaki
CZAR	BR000984	pol	TPA_inf: Capsaspora owczarzaki DNA, non-LTR retrotransposon CoL1, complete sequence, strain: ATCC 30864	Capsaspora owczarzaki
DongAg;Dong	AB097127	rt	Anopheles gambiae retrotransposon DongAg DNA, partial sequence	Anopheles gambiae
EhRLE2	AB097128	rt	Entamoeba histolytica retrotransposon EhRLE2 DNA, complete sequence	Entamoeba histolytica
EhRLE3	AB097129	rt	Entamoeba histolytica retrotransposon EhRLE3 DNA, complete sequence	Entamoeba histolytica
Genie	AF440196	endonuclease	Giardia intestinalis non-LTR retrotransposon GENIE 1 pol polyprotein gene, complete cds	Giardia intestinalis
Genie	BK000097	endonuclease	TPA_exp: Giardia intestinalis non-LTR retrotransposon Genie 1A gene, partial sequence	Giardia intestinalis
Genie	BK000095	endonuclease	TPA_exp: Giardia intestinalis non-LTR retrotransposon Genie 1 gene, partial sequence	Giardia intestinalis
Genie	BK000096	insertion site for non-LTR retrotransposon Genie 1	TPA_exp: Giardia intestinalis non-LTR retrotransposon Genie 1 target site sequence	Giardia intestinalis

Table 2				
Genie	AY216701	non-experimental evidence, no additional details recorded	Girardia tigrina GENIE retrotransposon, complete sequence	Girardia tigrina
Genie	BK000098	similar to endonuclease	TPA_exp: Giardia intestinalis non-LTR retrotransposon Genie 2 gene, complete sequence	Giardia intestinalis
GILD	AF433877	(tca)n (SEQ ID NO: 1532) or (tga)n (SEQ ID NO: 1533), n=2-4	Giardia intestinalis inactive non-LTR retrotransposon GILD, consensus sequence	Giardia intestinalis
GILM	AF433875	poly(dA) tract	Giardia intestinalis non-LTR LINE-like retrotransposon GILM, complete sequence	Giardia intestinalis
Hero	AB097132	rt	Danio rerio retrotransposon HERODr DNA, complete sequence	Danio rerio
Hero	AB097130	rt	Takifugu rubripes retrotransposon HEROFr DNA, complete sequence	Takifugu rubripes
HERO Tn	AB097131	rt	Tetraodon nigroviridis retrotransposon HEROTn DNA, complete sequence	Tetraodon nigroviridis
NeSL_3_135_68117	FJ905846	non-LTR retrotransposon:NeSL_3_135_68117	Daphnia pulex non-LTR retrotransposon NeSL_3_135_68117, complete sequence	Daphnia pulex
NeSL; NeSL-1	DQ099731	target site duplication	Caenorhabditis briggsae transposon NeSL-1-like non-LTR retrotransposon NeSL-1Cb reverse transcriptase (pol) gene, complete cds	Caenorhabditis briggsae
PERERE-9	BN000800		TPA_exp: Schistosoma mansoni Perere-9 non-LTR retrotransposon	Schistosoma mansoni
R2	AF015814	R2	Limulus polyphemus retrotransposon R2, complete sequence	Limulus polyphemus
R2	AF090145	R2	Nasonia vitripennis R2 non-LTR retrotransposable element reverse transcriptase gene, partial cds	Nasonia vitripennis
R2	AF015818	R2	Porcellio scaber retrotransposon R2, complete sequence	Porcellio scaber
R2	AF015815	R2	Anurida maritima retrotransposon R2, complete sequence	Anurida maritima
R2	M16558	R2	Bombyx mori rDNA insertion element R2 (typell), complete cds	Bombyx mori
R2	AF015819	R2	Forficula auricularia retrotransposon R2, complete sequence.	Forficula auricularia
R2	EU854578	R2	Triops cancriformis non-LTR retrotransposon R2 reverse transcriptase gene, complete cds	Triops cancriformis
R2	GU949555	R2	Reticulitermes lucifugus non-LTR retrotransposon R2, complete sequence; and R2 protein gene, complete cds	Reticulitermes lucifugus
R2	AB097123	rt	Ciona intestinalis retrotransposon R2Ci-C DNA, partial sequence	Ciona intestinalis
R2	AB097124	rt	Ciona intestinalis retrotransposon R2Ci-D DNA, partial sequence	Ciona intestinalis

Table 2				
R2	FJ461304	R2	Rhynchosciara americana non-LTR retrotransposon RaR2 reverse transcriptase gene, complete cds	Rhynchosciara americana
R2	AB097121	rt	Ciona intestinalis retrotransposon R2Ci-A DNA, complete sequence	Ciona intestinalis
R2	KP657892	R2	Bacillus rossius isolate roCAP(full).9 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP657890	R2	Bacillus rossius isolate roCAP(full).7 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP657888	R2	Bacillus rossius isolate roCAP(full).5 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP657870	R2	Bacillus rossius isolate roCAP(full).1 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP657833	R2	Bacillus rossius isolate roANZ(full).13 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP657832	R2	Bacillus rossius isolate roANZ(full).12 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP657830	R2	Bacillus rossius isolate roANZ(full).10 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP657807	R2	Bacillus rossius isolate roANZ(-101).8 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP657806	R2	Bacillus rossius isolate roANZ(-101).7 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP657805	R2	Bacillus rossius isolate roANZ(-101).6 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP657802	R2	Bacillus rossius isolate roANZ(-101).3 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP657799	R2	Bacillus rossius isolate roANZ(-101).1 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	FJ461304	R2	Rhynchosciara americana non-LTR retrotransposon RaR2 reverse transcriptase gene, complete cds	Rhynchosciara americana
R2	JQ082370	polyA_signal_sequence	Eyprepocnemis plorans non-LTR retrotransposon R2 R2 protein gene, complete cds	Eyprepocnemis plorans
R2	KJ958672	R2	Bacillus rossius non-LTR retrotransposon reCUR4_deg, partial sequence	Bacillus rossius
R2	KJ958671	R2	Bacillus rossius non-LTR retrotransposon reCUR3_deg, partial sequence	Bacillus rossius
R2	KJ958670	R2	Bacillus rossius non-LTR retrotransposon reCUR2_deg, partial sequence	Bacillus rossius
R2	KJ958669	R2	Bacillus rossius non-LTR retrotransposon reCUR1_deg, partial sequence	Bacillus rossius
R2	KJ958668	R2	Bacillus rossius non-LTR retrotransposon reCDF6_deg, partial sequence	Bacillus rossius
R2	KJ958667	R2	Bacillus rossius non-LTR retrotransposon reCDF5_deg, partial sequence	Bacillus rossius
R2	KJ958666	R2	Bacillus rossius non-LTR retrotransposon reCDF4_deg, partial sequence	Bacillus rossius

Table 2				
R2	KJ958 665	R2	Bacillus rossius non-LTR retrotransposon reCDF3_deg, partial sequence	Bacillus rossius
R2	KJ958 664	R2	Bacillus rossius non-LTR retrotransposon reCDF2_deg, partial sequence	Bacillus rossius
R2	KJ958 663	R2	Bacillus rossius non-LTR retrotransposon reCDF1_deg, partial sequence	Bacillus rossius
R2	KJ958 662	R2	Bacillus rossius non-LTR retrotransposon reCOM10_deg, partial sequence	Bacillus rossius
R2	KJ958 661	R2	Bacillus rossius non-LTR retrotransposon reCOM5_deg, partial sequence	Bacillus rossius
R2	KJ958 660	R2	Bacillus rossius non-LTR retrotransposon reCOM7_deg, partial sequence	Bacillus rossius
R2	KJ958 659	R2	Bacillus rossius non-LTR retrotransposon reCOM3_deg, partial sequence	Bacillus rossius
R2	KJ958 658	R2	Bacillus rossius non-LTR retrotransposon reCOM2_deg, partial sequence	Bacillus rossius
R2	KJ958 657	R2	Bacillus rossius non-LTR retrotransposon reMSN6_deg, partial sequence	Bacillus rossius
R2	KJ958 656	R2	Bacillus rossius non-LTR retrotransposon reMSN5_deg, partial sequence	Bacillus rossius
R2	KJ958 655	R2	Bacillus rossius non-LTR retrotransposon reMSN4_deg, partial sequence	Bacillus rossius
R2	KJ958 654	R2	Bacillus rossius non-LTR retrotransposon reMSN3_deg, partial sequence	Bacillus rossius
R2	KJ958 653	R2	Bacillus rossius non-LTR retrotransposon reMSN2_deg, partial sequence	Bacillus rossius
R2	KJ958 652	R2	Bacillus rossius non-LTR retrotransposon reMSN1_deg, partial sequence	Bacillus rossius
R2	KJ958 651	R2	Bacillus rossius non-LTR retrotransposon rePAT8_deg, partial sequence	Bacillus rossius
R2	KJ958 650	R2	Bacillus rossius non-LTR retrotransposon rePAT7_deg, partial sequence	Bacillus rossius
R2	KJ958 649	R2	Bacillus rossius non-LTR retrotransposon rePAT6_deg, partial sequence	Bacillus rossius
R2	KJ958 648	R2	Bacillus rossius non-LTR retrotransposon rePAT5_deg, partial sequence	Bacillus rossius
R2	KJ958 647	R2	Bacillus rossius non-LTR retrotransposon rePAT4_deg, partial sequence	Bacillus rossius
R2	KJ958 646	R2	Bacillus rossius non-LTR retrotransposon rePAT3_deg, partial sequence	Bacillus rossius
R2	KJ958 645	R2	Bacillus rossius non-LTR retrotransposon rePAT2_deg, partial sequence	Bacillus rossius
R2	KJ958 644	R2	Bacillus rossius non-LTR retrotransposon rePAT1_deg, partial sequence	Bacillus rossius
R2	KJ958 643	R2	Bacillus rossius non-LTR retrotransposon rePAT9_deg, partial sequence	Bacillus rossius
R2	KJ958 642	R2	Bacillus rossius non-LTR retrotransposon trKOR2_deg, partial sequence	Bacillus rossius

Table 2				
R2	KJ958 641	R2	Bacillus rossius non-LTR retrotransposon reGAB9_deg, partial sequence	Bacillus rossius
R2	KJ958 640	R2	Bacillus rossius non-LTR retrotransposon reGAB8_deg, partial sequence	Bacillus rossius
R2	KJ958 639	R2	Bacillus rossius non-LTR retrotransposon reGAB7_deg, partial sequence	Bacillus rossius
R2	KJ958 638	R2	Bacillus rossius non-LTR retrotransposon reGAB6_deg, partial sequence	Bacillus rossius
R2	KJ958 637	R2	Bacillus rossius non-LTR retrotransposon reGAB5_deg, partial sequence	Bacillus rossius
R2	KJ958 636	R2	Bacillus rossius non-LTR retrotransposon reGAB4_deg, partial sequence	Bacillus rossius
R2	KJ958 635	R2	Bacillus rossius non-LTR retrotransposon reGAB3_deg, partial sequence	Bacillus rossius
R2	KJ958 634	R2	Bacillus rossius non-LTR retrotransposon reGAB2_deg, partial sequence	Bacillus rossius
R2	KJ958 633	R2	Bacillus rossius non-LTR retrotransposon reGAB1_deg, partial sequence	Bacillus rossius
R2	KJ958 632	R2	Bacillus rossius non-LTR retrotransposon reTDS1_deg, partial sequence	Bacillus rossius
R2	KJ958 631	R2	Bacillus rossius non-LTR retrotransposon reTDS11_deg, partial sequence	Bacillus rossius
R2	KJ958 629	R2	Bacillus rossius non-LTR retrotransposon reCOM8, partial sequence	Bacillus rossius
R2	KJ958 628	R2	Bacillus rossius non-LTR retrotransposon reCOM6, partial sequence	Bacillus rossius
R2	KJ958 627	R2	Bacillus rossius non-LTR retrotransposon reCOM4, partial sequence	Bacillus rossius
R2	KJ958 626	R2	Bacillus rossius non-LTR retrotransposon reCOM1, partial sequence	Bacillus rossius
R2	KJ958 624	R2	Bacillus rossius non-LTR retrotransposon trKOR10, partial sequence	Bacillus rossius
R2	KJ958 623	R2	Bacillus rossius non-LTR retrotransposon reGAB10, partial sequence	Bacillus rossius
R2	KJ958 619	R2	Bacillus rossius non-LTR retrotransposon trKOR4, partial sequence	Bacillus rossius
R2	KJ958 618	R2	Bacillus rossius non-LTR retrotransposon trKOR3, partial sequence	Bacillus rossius
R2	KJ958 617	R2	Bacillus rossius non-LTR retrotransposon trKOR1, partial sequence	Bacillus rossius
R2	KJ958 613	R2	Bacillus rossius non-LTR retrotransposon reVIR4, partial sequence	Bacillus rossius
R2	KJ958 612	R2	Bacillus rossius non-LTR retrotransposon reVIR3, partial sequence	Bacillus rossius
R2	KJ958 611	R2	Bacillus rossius non-LTR retrotransposon reVIR2, partial sequence	Bacillus rossius
R2	KJ958 610	R2	Bacillus rossius non-LTR retrotransposon reVIR1, partial sequence	Bacillus rossius

Table 2				
R2	KJ958609	R2	Bacillus rossius non-LTR retrotransposon reTDS17, partial sequence	Bacillus rossius
R2	KJ958608	R2	Bacillus rossius non-LTR retrotransposon reTDS16, partial sequence	Bacillus rossius
R2	KJ958607	R2	Bacillus rossius non-LTR retrotransposon reTDS15, partial sequence	Bacillus rossius
R2	KJ958606	R2	Bacillus rossius non-LTR retrotransposon reTDS14, partial sequence	Bacillus rossius
R2	KJ958605	R2	Bacillus rossius non-LTR retrotransposon reTDS13, partial sequence	Bacillus rossius
R2	KJ958604	R2	Bacillus rossius non-LTR retrotransposon reTDS12, partial sequence	Bacillus rossius
R2	KJ958603	R2	Bacillus rossius non-LTR retrotransposon reTDS7, partial sequence	Bacillus rossius
R2	KJ958602	R2	Bacillus rossius non-LTR retrotransposon reTDS6, partial sequence	Bacillus rossius
R2	KJ958601	R2	Bacillus rossius non-LTR retrotransposon reTDS5, partial sequence	Bacillus rossius
R2	KJ958600	R2	Bacillus rossius non-LTR retrotransposon reTDS4, partial sequence	Bacillus rossius
R2	KJ958599	R2	Bacillus rossius non-LTR retrotransposon reTDS3, partial sequence	Bacillus rossius
R2	KJ958598	R2	Bacillus rossius non-LTR retrotransposon reTDS2, partial sequence	Bacillus rossius
R2	KJ958594	R2	Bacillus rossius non-LTR retrotransposon reBER4, partial sequence	Bacillus rossius
R2	KJ958593	R2	Bacillus rossius non-LTR retrotransposon reBER2, partial sequence	Bacillus rossius
R2	KJ958592	R2	Bacillus rossius non-LTR retrotransposon reBER1, partial sequence	Bacillus rossius
R2	KJ958591	R2	Bacillus rossius non-LTR retrotransposon roFOL7, partial sequence	Bacillus rossius
R2	KJ958590	R2	Bacillus rossius non-LTR retrotransposon roFOL6, partial sequence	Bacillus rossius
R2	KJ958589	R2	Bacillus rossius non-LTR retrotransposon roFOL5, partial sequence	Bacillus rossius
R2	KJ958588	R2	Bacillus rossius non-LTR retrotransposon roFOL4, partial sequence	Bacillus rossius
R2	KJ958587	R2	Bacillus rossius non-LTR retrotransposon roFOL3, partial sequence	Bacillus rossius
R2	KJ958586	R2	Bacillus rossius non-LTR retrotransposon roFOL2, partial sequence	Bacillus rossius
R2	KJ958585	R2	Bacillus rossius non-LTR retrotransposon roFOL1, partial sequence	Bacillus rossius
R2	KJ958584	R2	Bacillus rossius non-LTR retrotransposon roANZ10, partial sequence	Bacillus rossius
R2	KJ958583	R2	Bacillus rossius non-LTR retrotransposon roANZ9, partial sequence	Bacillus rossius

Table 2				
R2	KJ958582	R2	Bacillus rossius non-LTR retrotransposon roANZ8, partial sequence	Bacillus rossius
R2	KJ958581	R2	Bacillus rossius non-LTR retrotransposon roANZ7, partial sequence	Bacillus rossius
R2	KJ958580	R2	Bacillus rossius non-LTR retrotransposon roANZ6, partial sequence	Bacillus rossius
R2	KJ958579	R2	Bacillus rossius non-LTR retrotransposon roANZ5, partial sequence	Bacillus rossius
R2	KJ958578	R2	Bacillus rossius non-LTR retrotransposon roANZ4, partial sequence	Bacillus rossius
R2	KJ958577	R2	Bacillus rossius non-LTR retrotransposon roANZ3, partial sequence	Bacillus rossius
R2	KJ958576	R2	Bacillus rossius non-LTR retrotransposon roANZ2, partial sequence	Bacillus rossius
R2	KJ958575	R2	Bacillus rossius non-LTR retrotransposon roANZ1, partial sequence	Bacillus rossius
R2	KJ958574	R2	Bacillus rossius non-LTR retrotransposon roCAP9, partial sequence	Bacillus rossius
R2	KJ958573	R2	Bacillus rossius non-LTR retrotransposon roCAP8, partial sequence	Bacillus rossius
R2	KJ958572	R2	Bacillus rossius non-LTR retrotransposon roCAP7, partial sequence	Bacillus rossius
R2	KJ958571	R2	Bacillus rossius non-LTR retrotransposon roCAP6, partial sequence	Bacillus rossius
R2	KJ958570	R2	Bacillus rossius non-LTR retrotransposon roCAP5, partial sequence	Bacillus rossius
R2	KJ958569	R2	Bacillus rossius non-LTR retrotransposon roCAP4, partial sequence	Bacillus rossius
R2	KJ958568	R2	Bacillus rossius non-LTR retrotransposon roCAP3, partial sequence	Bacillus rossius
R2	KJ958567	R2	Bacillus rossius non-LTR retrotransposon roCAP2, partial sequence	Bacillus rossius
R2	KJ958566	R2	Bacillus rossius non-LTR retrotransposon roCAP1, partial sequence	Bacillus rossius
R2	KJ958565	R2	Bacillus rossius non-LTR retrotransposon roCAP10, partial sequence	Bacillus rossius
R2	JN937654	R2	Lepidurus apus lubbocki isolate lu8a 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon psR2LI, complete sequence	Lepidurus apus lubbocki
R2	JN937653	R2	Lepidurus apus lubbocki isolate lu7a 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon psR2LI, complete sequence	Lepidurus apus lubbocki
R2	JN937652	R2	Lepidurus apus lubbocki isolate lu2a 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon psR2LI, complete sequence	Lepidurus apus lubbocki
R2	JN937651	R2	Lepidurus apus lubbocki isolate b7c7 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon psR2LI, complete sequence	Lepidurus apus lubbocki

Table 2				
R2	JN93 7650	R2	Lepidurus apus lubbocki isolate b6c4 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon psR2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7649	R2	Lepidurus apus lubbocki isolate lu5a 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon psR2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7648	R2	Lepidurus apus lubbocki isolate lu1a 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon psR2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7647	R2	Lepidurus apus lubbocki isolate b6c5 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon psR2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7646	R2	Lepidurus apus lubbocki isolate b6c6 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon psR2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7645	R2	Lepidurus apus lubbocki isolate lu4a 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon psR2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7644	R2	Lepidurus apus lubbocki isolate lu3a 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon psR2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7643	R2	Lepidurus apus lubbocki isolate b6c3 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon psR2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7642	R2	Lepidurus apus lubbocki isolate lu6a 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon psR2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7641	R2	Lepidurus apus lubbocki isolate LM5h2 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7640	R2	Lepidurus apus lubbocki isolate LM2h5 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7639	R2	Lepidurus apus lubbocki isolate LM2h4 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7638	R2	Lepidurus apus lubbocki isolate LM5h5 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7637	R2	Lepidurus apus lubbocki isolate LM5h4 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7636	R2	Lepidurus apus lubbocki isolate LM5h3 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7635	R2	Lepidurus apus lubbocki isolate LM5h1 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LI, complete sequence	Lepidurus apus lubbocki

Table 2				
R2	JN93 7634	R2	Lepidurus apus lubbocki isolate LM2h3 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7633	R2	Lepidurus apus lubbocki isolate LM2h2 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7632	R2	Lepidurus apus lubbocki isolate LM2h1 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LI, complete sequence	Lepidurus apus lubbocki
R2	JN93 7631	R2	Lepidurus arcticus isolate T6 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2La, complete sequence	Lepidurus arcticus
R2	JN93 7630	R2	Lepidurus arcticus isolate T5 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2La, complete sequence	Lepidurus arcticus
R2	JN93 7629	R2	Lepidurus arcticus isolate T4 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2La, complete sequence	Lepidurus arcticus
R2	JN93 7628	R2	Lepidurus arcticus isolate T3 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2La, complete sequence	Lepidurus arcticus
R2	JN93 7627	R2	Lepidurus arcticus isolate T2 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2La, complete sequence	Lepidurus arcticus
R2	JN93 7626	R2	Lepidurus arcticus isolate T1 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2La, complete sequence	Lepidurus arcticus
R2	JN93 7625	R2	Lepidurus arcticus isolate V4 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2La, complete sequence	Lepidurus arcticus
R2	JN93 7624	R2	Lepidurus arcticus isolate V3 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2La, complete sequence	Lepidurus arcticus
R2	JN93 7623	R2	Lepidurus arcticus isolate V2 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2La, complete sequence	Lepidurus arcticus
R2	JN93 7622	R2	Lepidurus arcticus isolate V1 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2La, complete sequence	Lepidurus arcticus
R2	JN93 7615	R2	Lepidurus couesii isolate D3a7f 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcB, complete sequence	Lepidurus couesii
R2	JN93 7614	R2	Lepidurus couesii isolate D3a5f 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcB, complete sequence	Lepidurus couesii
R2	JN93 7613	R2	Lepidurus couesii isolate D3a4f 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcB, complete sequence	Lepidurus couesii

Table 2				
R2	JN93 7612	R2	Lepidurus couesii isolate D3a3f 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcB, complete sequence	Lepidurus couesii
R2	JN93 7611	R2	Lepidurus couesii isolate D3a2f 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcB, complete sequence	Lepidurus couesii
R2	JN93 7610	R2	Lepidurus couesii isolate D3_8 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcB, complete sequence	Lepidurus couesii
R2	JN93 7609	R2	Lepidurus couesii isolate D3_7 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcB, complete sequence	Lepidurus couesii
R2	JN93 7608	R2	Lepidurus couesii isolate D3_6 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcB, complete sequence	Lepidurus couesii
R2	JN93 7607	R2	Lepidurus couesii isolate D3_5 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcB, complete sequence	Lepidurus couesii
R2	JN93 7606	R2	Lepidurus couesii isolate D3_4 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcB, complete sequence	Lepidurus couesii
R2	JN93 7605	R2	Lepidurus couesii isolate D3_3 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcB, complete sequence	Lepidurus couesii
R2	JN93 7604	R2	Lepidurus couesii isolate D3_2 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcB, complete sequence	Lepidurus couesii
R2	JN93 7603	R2	Lepidurus couesii isolate D3_1 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcB, complete sequence	Lepidurus couesii
R2	JN93 7602	R2	Lepidurus couesii isolate C2_5 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcA, complete sequence	Lepidurus couesii
R2	JN93 7601	R2	Lepidurus couesii isolate LcoC2r1_5 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcA, complete sequence	Lepidurus couesii
R2	JN93 7600	R2	Lepidurus couesii isolate LcoC2r1_6 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcA, complete sequence	Lepidurus couesii
R2	JN93 7599	R2	Lepidurus couesii isolate C2_8 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcA, complete sequence	Lepidurus couesii
R2	JN93 7598	R2	Lepidurus couesii isolate C2_4 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcA, complete sequence	Lepidurus couesii
R2	JN93 7597	R2	Lepidurus couesii isolate C2_9 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcA, complete sequence	Lepidurus couesii

Table 2				
R2	JN93 7596	R2	Lepidurus couesii isolate C2_7 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcA, complete sequence	Lepidurus couesii
R2	JN93 7595	R2	Lepidurus couesii isolate C2_6 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcA, complete sequence	Lepidurus couesii
R2	JN93 7594	R2	Lepidurus couesii isolate C2_3 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcA, complete sequence	Lepidurus couesii
R2	JN93 7593	R2	Lepidurus couesii isolate C2_2 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcA, complete sequence	Lepidurus couesii
R2	JN93 7592	R2	Lepidurus couesii isolate C2_1 28S ribosomal RNA gene, partial sequence; and non-LTR retrotransposon R2LcA, complete sequence	Lepidurus couesii
R2	AF01 5822	R2 ORF	Tenebrio molitor retrotransposon R2 reverse transcriptase gene, partial cds	Tenebrio molitor
R2	AF01 5817	R2 ORF	Tenebrio molitor retrotransposon R2 reverse transcriptase gene, partial cds	Tenebrio molitor
R2	AF01 5816	R2 ORF	Hippodamia convergens retrotransposon R2 reverse transcriptase gene, partial cds	Hippodamia convergens
R2	KP65 7866	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-714).5 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7865	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-714).4 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7863	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-714).2 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7862	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-714).1 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7861	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1297).9 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7860	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1297).8 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7859	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1297).7 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7858	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1297).6 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7857	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1297).5 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7856	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1297).4 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7855	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1297).3 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7854	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1297).2 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7853	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1297).10 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius

Table 2

R2	KP65 7852	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1297).1 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7851	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1172).9 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7850	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1172).8 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7849	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1172).7 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7848	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1172).6 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7847	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1172).5 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7846	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1172).4 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7845	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1172).3 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7844	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1172).2 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7843	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1172).10 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7842	retrotransposon:R2 Br	Bacillus rossius isolate roCAP(-1172).1 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7824	retrotransposon:R2 Br	Bacillus rossius isolate roANZ(-1062).5 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7823	retrotransposon:R2 Br	Bacillus rossius isolate roANZ(-1062).4 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7822	retrotransposon:R2 Br	Bacillus rossius isolate roANZ(-1062).3 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7820	retrotransposon:R2 Br	Bacillus rossius isolate roANZ(-1062).2 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7816	retrotransposon:R2 Br	Bacillus rossius isolate roANZ(-1062).16 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7814	retrotransposon:R2 Br	Bacillus rossius isolate roANZ(-1062).14 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7810	retrotransposon:R2 Br	Bacillus rossius isolate roANZ(-1062).10 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7809	retrotransposon:R2 Br	Bacillus rossius isolate roANZ(-1062).1 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7759	retrotransposon:R2 Br	Bacillus rossius isolate rePAT(-1297).8 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7757	retrotransposon:R2 Br	Bacillus rossius isolate rePAT(-1297).6 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius
R2	KP65 7751	retrotransposon:R2 Br	Bacillus rossius isolate rePAT(-1297).1 retrotransposon R2Br reverse transcriptase gene, partial cds	Bacillus rossius

Table 2				
R2	AB097125	rt	Ciona savignyi retrotransposon R2Cs-D DNA, partial sequence	Ciona savignyi
R2	AB097124	rt	Ciona intestinalis retrotransposon R2Ci-D DNA, partial sequence	Ciona intestinalis
R2	AB097123	rt	Ciona intestinalis retrotransposon R2Ci-C DNA, partial sequence	Ciona intestinalis
R2	AB097121	rt	Ciona intestinalis retrotransposon R2Ci-A DNA, complete sequence	Ciona intestinalis
R2	AB201417	rt	Triops longicaudatus non-LTR retrotransposon R2TI gene for reverse transcriptase, partial cds	Triops longicaudatus
R2	AB201416	rt	Procambarus clarkii non-LTR retrotransposon R2Pc gene for reverse transcriptase, partial cds	Procambarus clarkii
R2	AB201415	rt	Hasarius adansoni non-LTR retrotransposon R2Ha gene for reverse transcriptase, partial cds	Hasarius adansoni
R2	AB201414	rt	Metacrinus rotundus non-LTR retrotransposon R2Mr gene for reverse transcriptase, partial cds	Metacrinus rotundus
R2	AB201413	rt	Mauremys reevesii non-LTR retrotransposon R2Cr-B2 gene for reverse transcriptase, partial cds	Mauremys reevesii
R2	AB201412	rt	Mauremys reevesii non-LTR retrotransposon R2Cr-B1 gene for reverse transcriptase, partial cds	Mauremys reevesii
R2	AB201411	rt	Mauremys reevesii non-LTR retrotransposon R2Cr-A gene for reverse transcriptase, partial cds	Mauremys reevesii
R2	AB201410	rt	Oryzias latipes non-LTR retrotransposon R2OI-A gene for reverse transcriptase, partial cds	Oryzias latipes
R2	AB201409	rt	Tanichthys albonubes non-LTR retrotransposon R2Ta gene for reverse transcriptase, partial cds	Tanichthys albonubes
R2	AB201408	rt	Eptatretus burgeri non-LTR retrotransposon R2Eb gene for reverse transcriptase, partial cds	Eptatretus burgeri
R2	DQ099732	transposon:R2-like non-LTR retrotransposon R2Ag	Aedes aegypti transposon R2-like non-LTR retrotransposon R2Ag reverse transcriptase (pol) gene, partial cds	Aedes aegypti
R2	DQ099728	transposon:R2-like non-LTR retrotransposon R2Ag_B	Aedes aegypti transposon R2-like non-LTR retrotransposon R2Ag_B reverse transcriptase (pol) gene, partial cds	Aedes aegypti
R2	GU949559		Kaloterme flavicollis isolate Livorno non-LTR retrotransposon R2, complete sequence; and R2 protein gene, complete cds	Kaloterme flavicollis
R2	GU949557		Reticulitermes balkanensis non-LTR retrotransposon R2, partial sequence; and R2 protein gene, partial cds	Reticulitermes balkanensis
R2	GU949556		Reticulitermes grassei non-LTR retrotransposon R2, partial sequence; and R2 protein gene, partial cds	Reticulitermes grassei
R2	GU949554		Reticulitermes urbis non-LTR retrotransposon R2, complete sequence; and R2 protein gene, complete cds	Reticulitermes urbis

Table 2				
R2	AF412214		Schistosoma japonicum clone S10A non-LTR retrotransposon SjR2-like, partial sequence	Schistosoma japonicum
R2	AF015685		Drosophila mercatorum R2 retrotransposon reverse transcriptase domain protein gene, complete cds	Drosophila mercatorum
R2	KJ958674		Bacillus rossius retrotransposon R2Br, complete sequence	Bacillus rossius
R2	AF015814	R2	Limulus polyphemus retrotransposon R2, complete sequence	Limulus polyphemus
R2	M16558	R2	Bombyx mori rDNA insertion element R2 (typell), complete cds.	Bombyx mori
R2	GQ398057	R9Av	Adineta vaga copy 1 non-LTR retrotransposon R9, complete sequence; and disrupted 28S ribosomal RNA gene, partial sequence	Adineta vaga
R2Bm	AB076841	R2	Bombyx mori non-LTR retrotransposon R2Bm gene for reverse transcriptase, complete cds and 28S rRNA	Bombyx mori
R2Ci-B	AB097122	rt	Ciona intestinalis retrotransposon R2Ci-B DNA, complete sequence	Ciona intestinalis
R2Dr	AB097126	rt	Danio rerio retrotransposon R2Dr DNA, complete sequence	Danio rerio
R4	AH003588		Parascaris equorum transposon non-LTR retrotransposable element R4 reverse transcriptase gene, partial cds	Parascaris equorum
R4	ALU29445	R4	Ascaris lumbricoides site-specific non-LTR retrotransposable element R4 in 26S rDNA, complete sequence	Ascaris lumbricoides
R4	L08889	R4 Dong	Bombyx mori reverse transcriptase gene, complete cds	Bombyx mori
R4	DQ836390	MalR4-5	Maculinea alcon R4-like non-LTR retrotransposon R4-5 reverse transcriptase (RT) pseudogene, partial sequence	Phengaris alcon
R4	DQ836385	MnaR4-3	Maculinea nausithous R4-like non-LTR retrotransposon R4-3 reverse transcriptase (RT) pseudogene, partial sequence	Phengaris nausithous
R4	DQ836386	MnaR4-4	Maculinea nausithous R4-like non-LTR retrotransposon R4-4 reverse transcriptase (RT) pseudogene, partial sequence	Phengaris nausithous
R4	DQ836387	MnaR4-7	Maculinea nausithous R4-like non-LTR retrotransposon R4-7 reverse transcriptase (RT) pseudogene, partial sequence	Phengaris nausithous
R4	DQ836388	MnaR4-8	Maculinea nausithous R4-like non-LTR retrotransposon R4-8 reverse transcriptase (RT) pseudogene, partial sequence	Phengaris nausithous
R4	DQ836389	MnaR4-9	Maculinea nausithous R4-like non-LTR retrotransposon R4-9 reverse transcriptase (RT) pseudogene, partial sequence	Phengaris nausithous

Table 2				
R4	DQ83 6379	MteR4-1	Maculinea teleius R4-like non-LTR retrotransposon R4-1 reverse transcriptase (RT) pseudogene, partial sequence	Phengaris teleius
R4	DQ83 6384	MteR4-10	Maculinea teleius R4-like non-LTR retrotransposon R4-10 reverse transcriptase (RT) pseudogene, partial sequence	Phengaris teleius
R4	DQ83 6367	MteR4-2	Maculinea teleius R4-like non-LTR retrotransposon R4-2 reverse transcriptase (RT) gene, partial cds	Phengaris teleius
R4	DQ83 6380	MteR4-3	Maculinea teleius R4-like non-LTR retrotransposon R4-3 reverse transcriptase (RT) pseudogene, partial sequence	Phengaris teleius
R4	DQ83 6381	MteR4-4	Maculinea teleius R4-like non-LTR retrotransposon R4-4 reverse transcriptase (RT) pseudogene, partial sequence	Phengaris teleius
R4	DQ83 6382	MteR4-6	Maculinea teleius R4-like non-LTR retrotransposon R4-6 reverse transcriptase (RT) pseudogene, partial sequence	Phengaris teleius
R4	DQ83 6383	MteR4-8	Maculinea teleius R4-like non-LTR retrotransposon R4-8 reverse transcriptase (RT) pseudogene, partial sequence	Phengaris teleius
R4	DQ83 6374	transposon:R4-like non-LTR retrotransposon R4-1	Maculineaalcon R4-like non-LTR retrotransposon R4-1 reverse transcriptase (RT) gene, partial cds	Phengarisalcon
R4	DQ83 6373	transposon:R4-like non-LTR retrotransposon R4-1	Maculinea nausithous R4-like non-LTR retrotransposon R4-1 reverse transcriptase (RT) gene, partial cds	Phengaris nausithous
R4	DQ83 6375	transposon:R4-like non-LTR retrotransposon R4-2	Maculineaalcon R4-like non-LTR retrotransposon R4-2 reverse transcriptase (RT) gene, partial cds	Phengarisalcon
R4	DQ83 6376	transposon:R4-like non-LTR retrotransposon R4-3	Maculineaalcon R4-like non-LTR retrotransposon R4-3 reverse transcriptase (RT) gene, partial cds	Phengarisalcon
R4	DQ83 6377	transposon:R4-like non-LTR retrotransposon R4-4	Maculineaalcon R4-like non-LTR retrotransposon R4-4 reverse transcriptase (RT) gene, partial cds	Phengarisalcon
R4	DQ83 6371	transposon:R4-like non-LTR retrotransposon R4-5	Maculinea nausithous R4-like non-LTR retrotransposon R4-5 reverse transcriptase (RT) gene, partial cds	Phengaris nausithous
R4	DQ83 6368	transposon:R4-like non-LTR retrotransposon R4-5	Maculinea teleius R4-like non-LTR retrotransposon R4-5 reverse transcriptase (RT) gene, partial cds	Phengaris teleius
R4	DQ83 6378	transposon:R4-like non-LTR	Maculineaalcon R4-like non-LTR retrotransposon R4-6 reverse transcriptase (RT) gene, partial cds	Phengarisalcon

Table 2				
		retrotransposon R4-6		
R4	DQ83 6372	transposon:R4-like non-LTR retrotransposon R4-6	Maculinea nausithous R4-like non-LTR retrotransposon R4-6 reverse transcriptase (RT) gene, partial cds	Phengaris nausithous
R4	DQ83 6369	transposon:R4-like non-LTR retrotransposon R4-7	Maculinea teleius R4-like non-LTR retrotransposon R4-7 reverse transcriptase (RT) gene, partial cds	Phengaris teleius
R4	DQ83 6370	transposon:R4-like non-LTR retrotransposon R4-9	Maculinea teleius R4-like non-LTR retrotransposon R4-9 reverse transcriptase (RT) gene, partial cds	Phengaris teleius
R4	AF28 6191	transposon:retrotransposon Rex6	Xiphophorus maculatus retrotransposon Rex6 reverse transcriptase pseudogene, partial sequence	Xiphophorus maculatus
R5	KJ958 673	R2	Bacillus rossius non-LTR retrotransposon reCUR5_deg, partial sequence	Bacillus rossius
R5	KJ958 620	R2	Bacillus rossius non-LTR retrotransposon trKOR5, partial sequence	Bacillus rossius
R5	KJ958 614	R2	Bacillus rossius non-LTR retrotransposon reVIR5, partial sequence	Bacillus rossius
R5	KJ958 595	R2	Bacillus rossius non-LTR retrotransposon reBER5, partial sequence	Bacillus rossius
R5	AJ006 560	transposon:Amer5 non-LTR retrotransposon	Anopheles merus Amer5 non-LTR retrotransposon encoding reverse transcriptase, partial	Anopheles merus
R5	AF35 2479	transposon:NLRCh1-like non-LTR retrotransposon	Chironomus circumdatus clone cir5 transposon NLRCh1-like non-LTR retrotransposon reverse transcriptase gene, partial cds	Chironomus circumdatus
R5	AF35 2454	transposon:NLRCh1-like non-LTR retrotransposon	Chironomus alpestris clone dor50 transposon NLRCh1-like non-LTR retrotransposon reverse transcriptase gene, partial cds	Chironomus alpestris
R5	AF35 2404	transposon:NLRCh1-like non-LTR retrotransposon	Chironomus luridus clone lur5 transposon NLRCh1-like non-LTR retrotransposon reverse transcriptase gene, partial cds	Chironomus luridus
R8	FR85 2798	poly(A) tail	Beta vulgaris subsp. vulgaris LINE-type retrotransposon Belline2_3	Beta vulgaris subsp. vulgaris
R8	KJ958 630	R2	Bacillus rossius non-LTR retrotransposon reVIR8_deg, partial sequence	Bacillus rossius
R8	KJ958 621	R2	Bacillus rossius non-LTR retrotransposon trKOR8, partial sequence	Bacillus rossius
R8	KP00 1560	rex3-RT pseudogene_Contig ILU_TR8	Iberochondrostoma lusitanicum clone tr8a non-LTR retrotransposon Rex3, complete sequence	Iberochondrostoma lusitanicum
R8	FR85 2885	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon BNR114 (Beline1_114)	Beta vulgaris

Table 2				
				subsp. vulgaris
R8	FR85 2856	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon BNR45 (Belline1_45)	Beta vulgaris subsp. vulgaris
R8	FR85 2844	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon BNR22 (Belline1_22)	Beta vulgaris subsp. vulgaris
R8	FR85 2836	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon Belline17_6	Beta vulgaris subsp. vulgaris
R8	FR85 2834	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon Belline17_4	Beta vulgaris subsp. vulgaris
R8	FR85 2831	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon Belline17_1	Beta vulgaris subsp. vulgaris
R8	FR85 2829	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon Belline16_2	Beta vulgaris subsp. vulgaris
R8	FR85 2827	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon Belline15_3	Beta vulgaris subsp. vulgaris
R8	FR85 2819	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon Belline12_2	Beta vulgaris subsp. vulgaris
R8	FR85 2813	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon Belline9_5	Beta vulgaris subsp. vulgaris
R8	FR85 2807	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon Belline8_1	Beta vulgaris subsp. vulgaris
R8	FR85 2806	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon Belline7_18	Beta vulgaris subsp. vulgaris
R8	FR85 2799	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon Belline2_4	Beta vulgaris subsp. vulgaris

Table 2				
R8	FR85 2795	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon BNR19 (Belline1_19)	Beta vulgaris subsp. vulgaris
R8	AF35 2481	transposon:NLRCh1 -like non-LTR retrotransposon	Chironomus circumdatus clone cir8 transposon NLRCh1-like non-LTR retrotransposon reverse transcriptase gene, partial cds	Chironomus circumdatu s
R8;R5	FR85 2861	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon BNR59 (Belline1_59)	Beta vulgaris subsp. vulgaris
R8;R5	FR85 2857	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon BNR51 (Belline1_51)	Beta vulgaris subsp. vulgaris
R8;R7	FR85 2866	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon BNR76 (Belline1_76)	Beta vulgaris subsp. vulgaris
R8;R7	FR85 2838	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon BNR7 (Belline1_7)	Beta vulgaris subsp. vulgaris
R8;R7	AF35 2455	transposon:NLRCh1 -like non-LTR retrotransposon	Chironomus alpestris clone dor70 note identical sequence found in dor80 transposon NLRCh1-like non- LTR retrotransposon reverse transcriptase gene, partial cds	Chironomus alpestris
R8;R9	FR85 2878	right terminal repeat	Beta vulgaris subsp. vulgaris LINE-type retrotransposon BNR96 (Belline1_96)	Beta vulgaris subsp. vulgaris
R9	KJ958 625	R2	Bacillus rossius non-LTR retrotransposon trKOR9, partial sequence	Bacillus rossius
Rex6	AJ293 547	en	Oreochromis niloticus Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Oni-3	Oreochromi s niloticus
Rex6	AJ293 546	en	Oreochromis niloticus Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Oni-2	Oreochromi s niloticus
Rex6	AJ293 545	en	Oreochromis niloticus Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Oni-1	Oreochromi s niloticus
Rex6	AJ293 517	en	Xiphophorus maculatus Rex6 retrotransposon partial en pseudogene for endonuclease, clone Rex6-Xma-6	Xiphophoru s maculatus
Rex6	AJ293 516	en	Xiphophorus maculatus Rex6 retrotransposon partial en pseudogene for endonuclease, clone Rex6-Xma-5	Xiphophoru s maculatus
Rex6	AJ293 515	en	Xiphophorus maculatus Rex6 retrotransposon partial en pseudogene for endonuclease, clone Rex6-Xma-4	Xiphophoru s maculatus
Rex6	AJ293 514	en	Xiphophorus maculatus Rex6 retrotransposon partial en pseudogene for endonuclease, clone Rex6-Xma-3	Xiphophoru s maculatus
Rex6	AJ293 513	en	Xiphophorus maculatus Rex6 retrotransposon partial en pseudogene for endonuclease, clone Rex6-Xma-2	Xiphophoru s maculatus

Table 2				
Rex6	AJ293 512	en	Xiphophorus maculatus Rex6 retrotransposon partial en pseudogene for endonuclease, clone Rex6-Xma-1	Xiphophorus maculatus
Rex6	AJ293 538	en	Poecilia formosa Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Pfo-6	Poecilia formosa
Rex6	AJ293 537	en	Poecilia formosa Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Pfo-5	Poecilia formosa
Rex6	AJ293 536	en	Poecilia formosa Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Pfo-4	Poecilia formosa
Rex6	AJ293 535	en	Poecilia formosa Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Pfo-3	Poecilia formosa
Rex6	AJ293 534	en	Poecilia formosa Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Pfo-2	Poecilia formosa
Rex6	AJ293 533	en	Poecilia formosa Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Pfo-1	Poecilia formosa
Rex6	AJ293 526	en	Poeciliopsis gracilis Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Pgr-4	Poeciliopsis gracilis
Rex6	AJ293 525	en	Poeciliopsis gracilis Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Pgr-3	Poeciliopsis gracilis
Rex6	AJ293 524	en	Poeciliopsis gracilis Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Pgr-2	Poeciliopsis gracilis
Rex6	AJ293 523	en	Poeciliopsis gracilis Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Pgr-1	Poeciliopsis gracilis
Rex6	AJ293 522	en	Oryzias latipes Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Ola-5	Oryzias latipes
Rex6	AJ293 521	en	Oryzias latipes Rex6 retrotransposon partial en pseudogene for endonuclease, clone Rex6-Ola-4	Oryzias latipes
Rex6	AJ293 520	en	Oryzias latipes Rex6 retrotransposon partial en pseudogene for endonuclease, clone Rex6-Ola-3	Oryzias latipes
Rex6	AJ293 519	en	Oryzias latipes Rex6 retrotransposon partial en pseudogene for endonuclease, clone Rex6-Ola-2	Oryzias latipes
Rex6	AJ293 518	en	Oryzias latipes Rex6 retrotransposon partial en pseudogene for endonuclease, clone Rex6-Ola-1	Oryzias latipes
Rex6	AJ293 549	en	Cichlasoma labridens Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Cla-2	Herichthys labridens
Rex6	AJ293 548	en	Cichlasoma labridens Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Cla-1	Herichthys labridens
Rex6	AJ293 544	en	Heterandria bimaculata Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Hbi-6	Pseudoxiphophorus bimaculatus
Rex6	AJ293 543	en	Heterandria bimaculata Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Hbi-5	Pseudoxiphophorus bimaculatus
Rex6	AJ293 542	en	Heterandria bimaculata Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Hbi-4	Pseudoxiphophorus bimaculatus
Rex6	AJ293 541	en	Heterandria bimaculata Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Hbi-3	Pseudoxiphophorus bimaculatus

Table 2				
Rex6	AJ293540	en	Heterandria bimaculata Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Hbi-2	Pseudoxiphophorus bimaculatus
Rex6	AJ293539	en	Heterandria bimaculata Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Hbi-1	Pseudoxiphophorus bimaculatus
Rex6	AJ293532	en	Gambusia affinis Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Gaf-5	Gambusia affinis
Rex6	AJ293531	en	Gambusia affinis Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Gaf-5	Gambusia affinis
Rex6	AJ293530	en	Gambusia affinis Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Gaf-4	Gambusia affinis
Rex6	AJ293529	en	Gambusia affinis Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Gaf-3	Gambusia affinis
Rex6	AJ293528	en	Gambusia affinis Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Gaf-2	Gambusia affinis
Rex6	AJ293527	en	Gambusia affinis Rex6 retrotransposon partial en pseudogene for endonuclease, clone rex6-Gaf-1	Gambusia affinis
Rex6	JX576459	non-LTR retrotransposon:Rex6	Symphysodon discus isolate i non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576458	non-LTR retrotransposon:Rex6	Symphysodon discus isolate h non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576457	non-LTR retrotransposon:Rex6	Symphysodon discus isolate g non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576456	non-LTR retrotransposon:Rex6	Symphysodon discus isolate f non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576455	non-LTR retrotransposon:Rex6	Symphysodon discus isolate e non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576454	non-LTR retrotransposon:Rex6	Symphysodon discus isolate d non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576453	non-LTR retrotransposon:Rex6	Symphysodon discus isolate c non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576452	non-LTR retrotransposon:Rex6	Symphysodon discus isolate b non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576451	non-LTR retrotransposon:Rex6	Symphysodon discus isolate a non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576450	non-LTR retrotransposon:Rex6	Symphysodon discus isolate z8 non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus

Table 2				
Rex6	JX576 449	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate z7 non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 448	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate z6 non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 447	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate z5 non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 446	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate z4 non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 445	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate z3 non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 444	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate z2 non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 443	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate z1 non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 442	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate z non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 441	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate x non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 440	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate v non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 439	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate u non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 438	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate t non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 437	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate s non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 436	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate r non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 435	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate q non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus
Rex6	JX576 434	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate p non-LTR retrotransposon Rex6, partial sequence	Symphysodon discus

Table 2				
Rex6	JX576 433	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate o non-LTR retrotransposon Rex6, partial sequence	Symphysodon on discus
Rex6	JX576 432	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate n non-LTR retrotransposon Rex6, partial sequence	Symphysodon on discus
Rex6	JX576 431	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate m non-LTR retrotransposon Rex6, partial sequence	Symphysodon on discus
Rex6	JX576 430	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate l non-LTR retrotransposon Rex6, partial sequence	Symphysodon on discus
Rex6	JX576 429	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate k non-LTR retrotransposon Rex6, partial sequence	Symphysodon on discus
Rex6	JX576 428	non-LTR retrotransposon:Rex 6	Symphysodon discus isolate j non-LTR retrotransposon Rex6, partial sequence	Symphysodon on discus
Rex6	JX576 427	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone e non-LTR retrotransposon Rex6, partial sequence	Pterophyllum m scalare
Rex6	JX576 426	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone d non-LTR retrotransposon Rex6, partial sequence	Pterophyllum m scalare
Rex6	JX576 425	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone c non-LTR retrotransposon Rex6, partial sequence	Pterophyllum m scalare
Rex6	JX576 424	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone b non-LTR retrotransposon Rex6, partial sequence	Pterophyllum m scalare
Rex6	JX576 423	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone a non-LTR retrotransposon Rex6, partial sequence	Pterophyllum m scalare
Rex6	JX576 422	non-LTR retrotransposon:Rex 6	Geophagus proximus clone g non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	JX576 421	non-LTR retrotransposon:Rex 6	Geophagus proximus clone f non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	JX576 420	non-LTR retrotransposon:Rex 6	Geophagus proximus clone e non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	JX576 419	non-LTR retrotransposon:Rex 6	Geophagus proximus clone d non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	JX576 418	non-LTR retrotransposon:Rex 6	Geophagus proximus clone c non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus

Table 2				
Rex6	JX576 417	non-LTR retrotransposon:Rex 6	Geophagus proximus clone b non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	JX576 416	non-LTR retrotransposon:Rex 6	Geophagus proximus clone a non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	JX576 415	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone g non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	JX576 414	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone f non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	JX576 413	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone e non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	JX576 412	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone d non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	JX576 411	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone c non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	JX576 410	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone b non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	JX576 409	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone a non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	JX576 408	non-LTR retrotransposon:Rex 6	Cichla monoculus clone h non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus
Rex6	JX576 407	non-LTR retrotransposon:Rex 6	Cichla monoculus clone g non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus
Rex6	JX576 406	non-LTR retrotransposon:Rex 6	Cichla monoculus clone f non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus
Rex6	JX576 405	non-LTR retrotransposon:Rex 6	Cichla monoculus clone e non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus
Rex6	JX576 404	non-LTR retrotransposon:Rex 6	Cichla monoculus clone d non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus
Rex6	JX576 403	non-LTR retrotransposon:Rex 6	Cichla monoculus clone c non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus
Rex6	JX576 402	non-LTR retrotransposon:Rex 6	Cichla monoculus clone b non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus

Table 2				
Rex6	JX576 401	non-LTR retrotransposon:Rex 6	Cichla monoculus clone a non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus
Rex6	KF13 1853	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone z7 non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1852	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone z6 non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1851	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone z5 non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1850	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone z4 non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1849	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone z3 non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1848	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone z2 non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1847	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone z1 non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1846	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone z non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1845	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone x non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1844	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone v non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1843	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone u non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1842	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone t non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1841	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone s non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1840	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone r non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1839	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone q non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare

Table 2				
Rex6	KF13 1838	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone p non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1837	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone n non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1836	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone m non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1835	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone l non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1834	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone k non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1833	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone j non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1832	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone i non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1831	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone h non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1830	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone g non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1829	non-LTR retrotransposon:Rex 6	Pterophyllum scalare clone f non-LTR retrotransposon Rex6, partial sequence	Pterophyllu m scalare
Rex6	KF13 1828	non-LTR retrotransposon:Rex 6	Geophagus proximus clone z6 non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1827	non-LTR retrotransposon:Rex 6	Geophagus proximus clone z5 non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1826	non-LTR retrotransposon:Rex 6	Geophagus proximus clone z4 non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1825	non-LTR retrotransposon:Rex 6	Geophagus proximus clone z3 non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1824	non-LTR retrotransposon:Rex 6	Geophagus proximus clone z2 non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1823	non-LTR retrotransposon:Rex 6	Geophagus proximus clone z1 non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus

Table 2				
Rex6	KF13 1822	non-LTR retrotransposon:Rex 6	Geophagus proximus clone z non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1821	non-LTR retrotransposon:Rex 6	Geophagus proximus clone x non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1820	non-LTR retrotransposon:Rex 6	Geophagus proximus clone v non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1819	non-LTR retrotransposon:Rex 6	Geophagus proximus clone u non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1818	non-LTR retrotransposon:Rex 6	Geophagus proximus clone t non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1817	non-LTR retrotransposon:Rex 6	Geophagus proximus clone s non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1816	non-LTR retrotransposon:Rex 6	Geophagus proximus clone r non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1815	non-LTR retrotransposon:Rex 6	Geophagus proximus clone q non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1814	non-LTR retrotransposon:Rex 6	Geophagus proximus clone p non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1813	non-LTR retrotransposon:Rex 6	Geophagus proximus clone o non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1812	non-LTR retrotransposon:Rex 6	Geophagus proximus clone n non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1811	non-LTR retrotransposon:Rex 6	Geophagus proximus clone m non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1810	non-LTR retrotransposon:Rex 6	Geophagus proximus clone l non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1809	non-LTR retrotransposon:Rex 6	Geophagus proximus clone k non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1808	non-LTR retrotransposon:Rex 6	Geophagus proximus clone j non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1807	non-LTR retrotransposon:Rex 6	Geophagus proximus clone i non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus

Table 2				
Rex6	KF13 1806	non-LTR retrotransposon:Rex 6	Geophagus proximus clone h non-LTR retrotransposon Rex6, partial sequence	Geophagus proximus
Rex6	KF13 1805	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone z10 non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1804	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone z9 non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1803	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone z8 non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1802	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone z7 non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1801	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone z6 non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1800	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone z5 non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1799	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone z4 non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1798	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone z3 non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1797	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone z2 non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1796	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone z1 non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1795	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone z non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1794	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone x non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1793	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone v non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1792	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone u non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1791	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone t non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus

Table 2				
Rex6	KF13 1790	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone s non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1789	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone r non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1788	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone q non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1787	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone p non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1786	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone o non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1785	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone n non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1784	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone m non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1783	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone l non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1782	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone k non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1781	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone j non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1780	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone i non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1779	non-LTR retrotransposon:Rex 6	Astronotus ocellatus clone h non-LTR retrotransposon Rex6, partial sequence	Astronotus ocellatus
Rex6	KF13 1778	non-LTR retrotransposon:Rex 6	Cichla monoculus clone z6 non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus
Rex6	KF13 1777	non-LTR retrotransposon:Rex 6	Cichla monoculus clone z5 non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus
Rex6	KF13 1776	non-LTR retrotransposon:Rex 6	Cichla monoculus clone z4 non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus
Rex6	KF13 1775	non-LTR retrotransposon:Rex 6	Cichla monoculus clone z3 non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus

Table 2				
Rex6	KF13 1774	non-LTR retrotransposon:Rex 6	Cichla monocolus clone z2 non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1773	non-LTR retrotransposon:Rex 6	Cichla monocolus clone z1 non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1772	non-LTR retrotransposon:Rex 6	Cichla monocolus clone z non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1771	non-LTR retrotransposon:Rex 6	Cichla monocolus clone x non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1770	non-LTR retrotransposon:Rex 6	Cichla monocolus clone v non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1769	non-LTR retrotransposon:Rex 6	Cichla monocolus clone u non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1768	non-LTR retrotransposon:Rex 6	Cichla monocolus clone t non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1767	non-LTR retrotransposon:Rex 6	Cichla monocolus clone s non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1766	non-LTR retrotransposon:Rex 6	Cichla monocolus clone r non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1765	non-LTR retrotransposon:Rex 6	Cichla monocolus clone q non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1764	non-LTR retrotransposon:Rex 6	Cichla monocolus clone p non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1763	non-LTR retrotransposon:Rex 6	Cichla monocolus clone o non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1762	non-LTR retrotransposon:Rex 6	Cichla monocolus clone n non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1761	non-LTR retrotransposon:Rex 6	Cichla monocolus clone m non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1760	non-LTR retrotransposon:Rex 6	Cichla monocolus clone l non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus
Rex6	KF13 1759	non-LTR retrotransposon:Rex 6	Cichla monocolus clone k non-LTR retrotransposon Rex6, partial sequence	Cichla monocolus

Table 2				
Rex6	KF13 1758	non-LTR retrotransposon:Rex 6	Cichla monoculus clone j non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus
Rex6	KF13 1757	non-LTR retrotransposon:Rex 6	Cichla monoculus clone i non-LTR retrotransposon Rex6, partial sequence	Cichla monoculus
SLACS	JN60 8782	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-46 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8781	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-45 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8780	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-41 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8779	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-40 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8778	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-38 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8777	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-37 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8776	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-36 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8775	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-35 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8774	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-34 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8773	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-33 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8772	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-32 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8771	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-30 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8770	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-29 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8769	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-28 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8768	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-27 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8767	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-26 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8766	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-25 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8765	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-24 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8764	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-23 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8763	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-22 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8762	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-20 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8761	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-19 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8760	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-18 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8759	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-17 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8758	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-16 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8757	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-13 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8756	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-12 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8755	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-11 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8754	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-10 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8753	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-08 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8752	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-07 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8751	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-06 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8750	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-05 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8749	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-04 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8748	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-03 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8747	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate Y-01 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8746	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-83 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8745	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-81 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8744	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-80 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8743	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-79 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8742	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-78 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8741	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-76 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8740	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-75 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8739	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-74 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8738	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-66 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8737	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-65 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8736	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-62 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8735	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-61 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8734	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-60 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8733	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-59 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8732	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-57 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8731	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-49 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8730	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-41 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8729	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-28 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8728	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-27 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8727	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-23 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8726	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-21 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8725	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-14 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8724	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-11 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8723	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-07 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8722	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-06 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8721	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-05 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8720	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-03 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8719	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate X-01 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8718	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG40 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8717	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG39 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8716	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG38 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8715	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG37 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8714	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG36 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8713	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG35 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8712	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG34 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8711	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG33 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8710	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG32 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8709	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG31 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8708	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG30 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8707	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG29 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8706	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG28 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8705	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG27 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8704	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG26 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8703	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG25 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8702	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG24 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8701	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG23 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8700	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG22 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8699	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG21 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8698	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG20 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8697	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG19 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8696	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG18 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8695	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG17 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8694	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG16 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8693	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG14 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8692	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG13 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8691	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG12 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8690	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG11 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8689	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG10 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8688	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG09 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8687	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG08 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8686	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG07 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8685	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG06 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8684	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG05 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8683	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG04 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8682	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG03 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8681	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG02 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8680	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mG01 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8679	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG40 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8678	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG39 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8677	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG38 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8676	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG37 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8675	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG36 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8674	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG35 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8673	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG34 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8672	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG33 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8671	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG32 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8670	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG31 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8669	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG30 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8668	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG29 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8667	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG28 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8666	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG27 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8665	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG26 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8664	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG25 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8663	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG24 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8662	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG23 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8661	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG22 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8660	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG21 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8659	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG20 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8658	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG19 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8657	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG18 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8656	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG17 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8655	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG16 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8654	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG15 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8653	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG14 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8652	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG13 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8651	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG11 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8650	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG10 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8649	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG09 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8648	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG08 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8647	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG07 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8646	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG06 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8645	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG05 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8644	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG03 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8643	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG02 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8642	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fG01 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8641	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-47 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8640	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-46 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8639	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-45 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8638	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-44 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8637	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-42 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8636	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-41 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8635	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-40 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8634	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-37 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8633	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-36 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8632	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-35 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8631	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-34 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8630	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-33 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8629	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-32 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8628	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-31 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8627	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-26 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8626	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-24 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8625	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-23 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8624	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-22 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8623	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-21 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8622	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-20 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8621	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-19 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8620	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-18 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8619	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-17 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8618	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-16 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8617	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-15 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8616	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-14 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8615	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-13 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8614	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-12 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8613	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-11 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8612	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-10 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8611	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-09 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8610	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-08 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8609	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-06 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8608	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-05 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8607	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-03 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8606	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate A-01 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8236	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR40 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8235	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR39 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8234	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR38 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8233	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR35 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8232	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR32 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8231	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR29 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8230	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR26 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8229	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR25 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8228	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR18 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8227	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR17 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8226	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR14 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8225	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR13 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8224	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR12 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8223	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR11 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8222	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR09 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8221	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR08 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8220	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR07 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8219	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR06 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8218	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR05 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8217	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR04 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8216	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR03 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8215	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR02 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8214	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mR01 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8213	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL40 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8212	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL39 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8211	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL37 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8210	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL36 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8209	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL35 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8208	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL34 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8207	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL33 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8206	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL32 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8205	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL31 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8204	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL28 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8203	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL27 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8202	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL25 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8201	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL24 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8200	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL23 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8199	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL19 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8198	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL16 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8197	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL13 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8196	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL12 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8195	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL11 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8194	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL10 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8193	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL07 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8192	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL04 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8191	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL03 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8190	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL02 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8189	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mL01 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8188	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF40 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8187	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF39 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8186	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF38 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8185	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF37 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8184	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF36 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8183	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF35 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8182	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF33 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8181	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF32 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8180	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF31 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8179	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF30 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8178	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF29 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8177	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF28 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8176	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF27 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8175	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF26 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8174	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF25 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8173	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF23 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8172	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF22 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8171	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF21 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8170	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF20 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8169	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF19 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8168	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF18 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8167	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF17 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8166	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF16 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8165	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF15 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8164	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF14 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8163	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF13 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8162	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF11 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8161	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF10 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8160	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF09 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8159	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF08 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8158	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF07 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8157	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF06 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8156	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF05 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8155	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF04 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8154	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF03 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8153	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF02 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8152	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate mF01 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8151	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR40 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8150	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR39 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8149	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR38 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8148	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR37 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8147	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR36 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8146	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR35 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8145	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR34 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8144	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR33 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8143	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR32 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8142	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR31 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8141	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR30 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8140	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR29 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8139	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR28 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8138	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR27 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8137	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR26 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8136	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR25 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8135	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR24 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8134	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR23 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8133	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR22 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8132	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR21 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8131	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR20 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8130	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR19 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8129	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR18 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8128	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR17 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8127	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR16 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8126	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR15 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8125	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR14 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8124	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR13 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8123	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR12 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8122	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR11 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8121	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR09 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8120	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR08 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8119	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR07 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8118	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR05 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8117	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR04 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8116	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR03 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8115	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR02 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8114	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fR01 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8113	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL40 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8112	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL39 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2

SLACS	JN60 8111	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL38 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8110	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL37 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8109	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL36 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8108	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL35 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8107	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL34 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8106	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL33 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8105	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL32 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8104	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL31 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8103	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL30 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8102	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL29 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8101	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL28 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8100	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL27 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8099	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL26 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8098	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL25 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8097	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL24 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8096	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL23 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN608095	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL22 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608094	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL20 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608093	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL19 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608092	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL18 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608091	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL17 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608090	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL16 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608089	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL15 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608088	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL14 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608087	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL13 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608086	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL12 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608085	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL11 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608084	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL10 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608083	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL09 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608082	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL07 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608081	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL06 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608080	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate fL05 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8079	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL04 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8078	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL03 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8077	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fL01 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8076	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fF40 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8075	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fF39 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8074	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fF38 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8073	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fF37 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8072	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fF36 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8071	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fF35 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8070	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fF34 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8069	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fF33 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8068	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fF32 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8067	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fF31 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8066	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fF30 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8065	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fF29 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8064	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate fF28 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN60 8063	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff27 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8062	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff26 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8061	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff24 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8060	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff23 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8059	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff21 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8058	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff20 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8057	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff18 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8056	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff17 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8055	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff16 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8054	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff15 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8053	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff14 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8052	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff13 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8051	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff12 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8050	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff11 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8049	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff10 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN60 8048	non-LTR retrotransposon:SLA CS-like	Silene latifolia isolate ff09 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia

Table 2				
SLACS	JN608047	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate ff08 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608046	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate ff06 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608045	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate ff05 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608044	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate ff04 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608043	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate ff03 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
SLACS	JN608042	non-LTR retrotransposon:SLACS-like	Silene latifolia isolate ff02 non-LTR retrotransposon SLACS-like, partial sequence	Silene latifolia
YURECi	AB097133	rt	Ciona intestinalis retrotransposon YURECi DNA, complete sequence	Ciona intestinalis
CRE	.	Cnl1	C. neoformans non-LTR retrotransposon - consensus.	Cryptococcus neoformans
CRE	.	CRE-1_ACas	CRE non-LTR retrotransposon: consensus.	Acanthamoeba castellanii
CRE	.	Cre-1_BM	Cre-1_BM non-LTR retrotransposon - consensus.	Bombyx mori
CRE	.	CRE-1_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	Cre-1_FCy	Cre-1_FCy non-LTR retrotransposon - conceptual consensus.	Fragilariopsis cylindrus
CRE	.	Cre-1_HM	Cre-1_HM non-LTR retrotransposon - consensus.	Hydra vulgaris
CRE	.	CRE-1_HRo	Cre-like non-LTR retrotransposon: consensus sequence.	Helobdella robusta
CRE	.	CRE-1_LSa	CRE non-LTR retrotransposon: consensus.	Lactuca sativa
CRE	.	Cre-1_MB	Cre-1_MB non-LTR retrotransposon - consensus.	Monosiga brevicollis
CRE	.	Cre-1_NV	Cre-1_NV non-LTR retrotransposon - consensus.	Nematostella vectensis
CRE	.	CRE-1_PXu	Non-LTR retrotransposon from Papilio xuthus: consensus.	Papilio xuthus
CRE	.	CRE-10_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	CRE-11_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus

Table 2				
CRE	.	CRE-12_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	CRE-13_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	CRE-14_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	CRE-15_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	CRE-16_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	CRE-17_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	Cre-2_BM	Cre-2_BM non-LTR retrotransposon - consensus.	Bombyx mori
CRE	.	CRE-2_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	CRE-2_HMa	CRE non-LTR retrotransposon: consensus.	Hydra vulgaris
CRE	.	CRE-2_HRo	Cre-like non-LTR retrotransposon: consensus sequence.	Helobdella robusta
CRE	.	CRE-2_NV	CRE non-LTR retrotransposon: consensus.	Nematostella vectensis
CRE	.	CRE-2_PXu	Non-LTR retrotransposon from Papilio xuthus: consensus.	Papilio xuthus
CRE	.	CRE-3_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	CRE-3_HRo	CRE-like non-LTR retrotransposon: consensus sequence.	Helobdella robusta
CRE	.	CRE-3_NV	CRE non-LTR retrotransposon: consensus.	Nematostella vectensis
CRE	.	CRE-4_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	CRE-4_HRo	CRE-like non-LTR retrotransposon: consensus sequence.	Helobdella robusta
CRE	.	CRE-5_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	CRE-5_HRo	CRE-like non-LTR retrotransposon: consensus sequence.	Helobdella robusta
CRE	.	CRE-6_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	CRE-6_HRo	CRE-like non-LTR retrotransposon: consensus sequence.	Helobdella robusta
CRE	.	CRE-7_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	CRE-8_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus
CRE	.	CRE-9_CCri	Non-LTR retrotransposon from the red seaweed: consensus.	Chondrus crispus

Table 2				
CRE	M33009	CRE1	C.fasciculata retrotransposable element (CRE1).	Crithidia fasciculata
CRE	U19151	CRE2	C.fasciculata retrotransposable element (CRE2).	Crithidia fasciculata
CRE	M62862	CZAR	T. cruzi SL-RNA-associated non-LTR retrotransposon.	Trypanosoma cruzi
R4	.	Dong	Bombyx mori non-LTR retrotransposable element.	Bombyx mori
R4	.	DONG_FR2	Non-LTR retrotransposon; site-specific LINE; R4/Dong superfamily; DONG_FR2.	Takifugu rubripes
R4	.	Dong-1_AFC	Dong/R4-type non-LTR retrotransposon - consensus.	Cichlidae
R4	.	Dong-1_HMM	Non-LTR retrotransposon family from Heliconius melpomene melpomene.	Heliconius melpomene melpomene
R4	.	Dong-1_NVe	A Dong non-LTR retrotransposon family from Nematostella vectensis.	Nematostella vectensis
R4	.	Dong-1_PPo	Non-LTR retrotransposon from Papilio polytes: consensus.	Papilio polytes
R4	.	Dong-1_PXu	Non-LTR retrotransposon from Papilio xuthus: consensus.	Papilio xuthus
R4	.	Dong-2_BM	Non-LTR retrotransposon - a consensus.	Bombyx mori
R4	.	Dong-2_HMM	Non-LTR retrotransposon family from Heliconius melpomene melpomene.	Heliconius melpomene melpomene
R4	.	Dong-2_Lch	Dong-like non-LTR retrotransposon - consensus.	Latimeria chalumnae
R4	.	Dong-2_PPo	Non-LTR retrotransposon from Papilio polytes: consensus.	Papilio polytes
R4	.	DongAa	A Dong non-LTR retrotransposon family from Aedes aegypti.	Aedes aegypti
R4	AB097127	DongAG	Anopheles gambiae non-LTR retrotransposon DongAg - a partial sequence.	Anopheles gambiae
R4	AB097128	EhRLE2	Entamoeba histolytica retrotransposon EhRLE2, complete sequence.	Entamoeba histolytica
R4	AB097129	EhRLE3	Entamoeba histolytica retrotransposon EhRLE3, complete sequence.	Entamoeba histolytica
HERO	.	HERO-1_AFC	Hero-type non-LTR retrotransposon - consensus.	Cichlidae
HERO	.	HERO-1_BF	Amphioxus HERO-1_BF autonomous non-LTR Retrotransposon - consensus.	Branchiostoma floridae
HERO	.	HERO-1_HR	A family of HERO non-LTR retrotransposons - a consensus sequence.	Helobdella robusta
HERO	.	HERO-1_PP	A family of HERO non-LTR retrotransposons - a consensus sequence.	Physarum polycephalum
HERO	AAGJ02121261	HERO-1_SP	Sea urchin HERO-1_SP autonomous non-LTR Retrotransposon - consensus.	Strongylocentrotus purpuratus

Table 2				
HERO	.	HERO-2_BF	Amphioxus HERO-2_BF autonomous non-LTR Retrotransposon - consensus.	Branchiostoma floridae
HERO	.	HERO-2_DR	HERO-2_DR is a family of HERO non-LTR retrotransposons - a consensus.	Danio rerio
HERO	.	HERO-2_HR	Non-LTR retrotransposon: consensus sequence.	Helobdella robusta
HERO	048B05	Hero-2_SPur	HERO-type non-ltr retrotransposon from sea urchin.	Strongylocentrotus purpuratus
HERO	.	HERO-3_BF	HERO-3_BF is a family of HERO non-LTR retrotransposons - a consensus.	Branchiostoma floridae
HERO	.	HERO-3_DR	HERO-3_DR is a family of HERO non-LTR retrotransposons - a consensus.	Danio rerio
HERO	.	HERO-3_HR	Non-LTR retrotransposon: consensus sequence.	Helobdella robusta
HERO	.	Hero-3_SPur	HERO-type non-LTR retrotransposon from sea urchin.	Strongylocentrotus purpuratus
HERO	.	HERO-4_DR	HERO-4_DR is a family of HERO non-LTR retrotransposons - a consensus.	Danio rerio
HERO	.	HERO-4_HR	Non-LTR retrotransposon: consensus sequence.	Helobdella robusta
HERO	.	HERO-5_HR	Non-LTR retrotransposon: consensus sequence.	Helobdella robusta
HERO	.	HERO-6_HR	Non-LTR retrotransposon: consensus sequence.	Helobdella robusta
HERO	.	HERO-7_HR	Non-LTR retrotransposon: consensus sequence.	Helobdella robusta
HERO	.	HERO-8_HR	Non-LTR retrotransposon: consensus sequence.	Helobdella robusta
HERO	.	HERO-9_HR	Non-LTR retrotransposon: consensus sequence.	Helobdella robusta
HERO	.	HERODr	HERODr is a family of HERO non-LTR retrotransposons - a consensus.	Danio rerio
HERO	.	HEROFr	A HERO clade non-LTR Retrotransposon family - consensus.	Takifugu rubripes
HERO	.	HEROTn	HEROTn or Zebulon non-LTR retrotransposon - a consensus sequence.	Tetraodon nigroviridis
NeSL	.	LIN10B_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	LIN11_SM	Non-LTR retrotransposon: consensus.	Schmidtea mediterranea
NeSL	.	LIN13_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea

Table 2				
NeSL	.	LIN14_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	LIN15_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	LIN2_SM	Non-LTR retrotransposon (consensus).	Schmidtea mediterranea
NeSL	.	LIN21_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	LIN23_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	LIN24_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	LIN24B_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	LIN25_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	LIN26_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	LIN3_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	LIN4_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	LIN4b_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	LIN5_SM	Non-LTR retrotransposon from Schmidtea mediterranea: consensus.	Schmidtea mediterranea
NeSL	.	LIN6_SM	Non-LTR retrotransposon from Schmidtea mediterranea: consensus.	Schmidtea mediterranea
NeSL	.	LIN7_SM	Non-LTR retrotransposon from Schmidtea mediterranea: consensus.	Schmidtea mediterranea
NeSL	.	LIN7B_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea

Table 2				
NeSL	.	LIN9_SM	Non-LTR retrotransposon: consensus.	Schmidtea mediterranea
CRE	JQ747487	MoTeR1	Telomere-specific non-LTR retrotransposon MoTeR1 from Magnaporthe oryzae.	Magnaporthe oryzae
CRE	JQ747488	MoTeR2	Telomere-specific non-LTR retrotransposon MoTeR2 from Magnaporthe oryzae.	Magnaporthe oryzae
NeSL	Z82058	NeSL-1	NeSL-1 is a non-LTR retrotransposon, complete sequence.	Caenorhabditis elegans
NeSL	.	NeSL-1_C11	A family of NeSL non-LTR retrotransposons.	Caenorhabditis tropicalis
NeSL	.	NeSL-1_CA	A family of NeSL non-LTR retrotransposons.	Caenorhabditis angaria
NeSL	.	NeSL-1_CBre	A family of NeSL non-LTR retrotransposons - consensus.	Caenorhabditis brenneri
NeSL	.	NeSL-1_CBri	A family of NeSL non-LTR retrotransposons.	Caenorhabditis briggsae
NeSL	.	NeSL-1_CJap	A family of NeSL non-LTR retrotransposons - consensus.	Caenorhabditis japonica
NeSL	.	NeSL-1_CRem	A family of NeSL non-LTR retrotransposons - consensus.	Caenorhabditis remanei
NeSL	.	NeSL-1_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	NeSL-1_TV	A family of NeSL non-LTR retrotransposons - consensus.	Trichomonas vaginalis
NeSL	.	NeSL-2_CBre	A family of NeSL non-LTR retrotransposons - consensus.	Caenorhabditis brenneri
NeSL	.	NeSL-2_CRem	A family of NeSL non-LTR retrotransposons - consensus.	Caenorhabditis remanei
NeSL	.	NeSL-2_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
NeSL	.	NeSL-3_CBre	A family of NeSL non-LTR retrotransposons - consensus.	Caenorhabditis brenneri
NeSL	.	NeSL-3_CRem	A family of NeSL non-LTR retrotransposons - consensus.	Caenorhabditis remanei
NeSL	chrUn	NeSL-4_CRem	A family of NeSL non-LTR retrotransposons.	Caenorhabditis remanei
NeSL	.	NeSL-4_SM	Non-LTR retrotransposon; consensus.	Schmidtea mediterranea
R2	BN000800	PERERE-9	Schistosoma mansoni Perere-9 non-LTR retrotransposon (EST).	Schistosoma mansoni
R4	.	Plat_R4	R4 Non-LTR Retrotransposon from Ornithorhynchus.	Ornithorhynchus
R2	AF015815	R2_AM	Anurida maritima retrotransposon R2, complete sequence.	Anurida maritima

Table 2				
R2	M16558	R2_BM	Bombyx mori rDNA insertion element R2 (type II), complete cds.	Bombyx mori
R2	AB097121	R2_CI	R2-type LINE.	Ciona intestinalis
R2	.	R2_CPB	Non-LTR retrotransposon: consensus.	Chrysemys picta bellii
R2	.	R2_DAn	28S rDNA-specific non-LTR retrotransposon R2 in Drosophila ananassae.	Drosophila ananassae
R2	X51967	R2_DM	LINE-like retrotransposable element R2DM.	Drosophila melanogaster
R2	.	R2_DPe	28S rDNA-specific non-LTR retrotransposon R2 in Drosophila persimilis.	Drosophila persimilis
R2	.	R2_DPs	28S rDNA-specific non-LTR retrotransposon R2 in Drosophila pseudoobscura.	Drosophila pseudoobscura
R2	.	R2_DSe	28S rDNA-specific non-LTR retrotransposon R2 in Drosophila sechellia.	Drosophila sechellia
R2	.	R2_DSi	28S rDNA-specific non-LTR retrotransposon R2 in Drosophila simulans.	Drosophila simulans
R2	.	R2_DYa	28S rDNA-specific non-LTR retrotransposon R2 in Drosophila yakuba.	Drosophila yakuba
R2	AF015819	R2_FA	Forficula auricularia retrotransposon R2, complete sequence.	Forficula auricularia
R2	AF015816	R2_HC	Hippodamia convergens retrotransposon R2 reverse transcriptase gene, partial cds.	Hippodamia convergens
R2	GU949558	R2_KF	28S rDNA-specific non-LTR retrotransposon R2 from Kaloterme flavicollis.	Kaloterme flavicollis
R2	AF015814	R2_LP	Limulus polyphemus retrotransposon R2, complete sequence.	Limulus polyphemus
R2	AF015818	R2_PS	Porcellio scaber retrotransposon R2, complete sequence.	Porcellio scaber
R2	GU949555	R2_RL	28S rDNA-specific non-LTR retrotransposon R2 from Reticulitermes lucifugus.	Reticulitermes lucifugus
R2	GU949554	R2_RU	28S rDNA-specific non-LTR retrotransposon R2 from Reticulitermes urbis.	Reticulitermes urbis
R2	.	R2-1_AAm	R2 non-LTR retrotransposon from lone star tick.	Amblyomma americanum
R2	.	R2-1_ACC	R2 non-LTR retrotransposon from golden eagle.	Aquila chrysaetos canadensis
R2	.	R2-1_ACh	R2 non-LTR retrotransposon from rifleman.	Acanthisitta chloris
R2	.	R2-1_AFo	R2 non-LTR retrotransposon from emperor penguin.	Aptenodytes forsteri

Table 2				
R2	.	R2-1_AMi	R2-type non-LTR retrotransposon.	Alligator mississippiensis
R2	.	R2-1_AOM	R2 non-LTR retrotransposon from kiwi.	Apteryx spp.
R2	.	R2-1_ApA	R2 non-LTR retrotransposon from north island brown kiwi.	Apteryx australis mantelli
R2	.	R2-1_APi	R2 non-LTR retrotransposon from pea aphid.	Acyrtosiphon pisum
R2	.	R2-1_BRG	R2 non-LTR retrotransposon from East African grey crowned crane.	Balearica regulorum gibbericeps
R2	.	R2-1_BTe	R2 non-LTR retrotransposon from buff-tailed bumblebee.	Bombus terrestris
R2	.	R2-1_CAnn	R2 non-LTR retrotransposon from Anna's hummingbird.	Calypte anna
R2	.	R2-1_CAu	R2 non-LTR retrotransposon from turkey vulture.	Cathartes aura
R2	.	R2-1_CBr	R2 non-LTR retrotransposon from American crow.	Corvus brachyrhynchos
R2	.	R2-1_CCa	R2 non-LTR retrotransposon from chuck-will's-widow.	Antrostomus carolinensis
R2	.	R2-1_CCan	R2 non-LTR retrotransposon from common cuckoo.	Cuculus canorus
R2	.	R2-1_CPu	R2 non-LTR retrotransposon from ruff.	Calidris pugnax
R2	.	R2-1_Crp	Non-LTR retrotransposon.	Crocodylus porosus
R2	.	R2-1_CSt	R2 non-LTR retrotransposon from speckled mousebird.	Colius striatus
R2	.	R2-1_CU	R2 non-LTR retrotransposon from MacQueen's bustard.	Chlamydotis macqueenii
R2	.	R2-1_CVo	R2 non-LTR retrotransposon from killdeer.	Charadrius vociferus
R2	.	R2-1_DWi	28S rDNA-specific non-LTR retrotransposon R2 in Drosophila willistoni.	Drosophila willistoni
R2	.	R2-1_EGa	R2 non-LTR retrotransposon from little egret.	Egretta garzetta
R2	.	R2-1_FAI	R2 non-LTR retrotransposon from collared flycatcher.	Ficedula albicollis
R2	.	R2-1_FCh	R2 non-LTR retrotransposon from Saker falcon.	Falco cherrug
R2	.	R2-1_FPe	R2 non-LTR retrotransposon from peregrine falcon.	Falco peregrinus

Table 2				
R2	.	R2-1_GA	R2 non-LTR retrotransposon from three-spined stickleback.	Gasterosteus aculeatus
R2	.	R2-1_Gav	Non-LTR retrotransposon.	Gavialis gangeticus
R2	.	R2-1_GFo	R2 non-LTR retrotransposon from medium ground finch.	Geospiza fortis
R2	.	R2-1_GSt	R2 non-LTR retrotransposon from red-throated loon.	Gavia stellata
R2	.	R2-1_HAI	R2 non-LTR retrotransposon from white-tailed eagle.	Haliaeetus albicilla
R2	.	R2-1_IS	R2 non-LTR retrotransposon from deer tick.	Ixodes scapularis
R2	.	R2-1_LCh	R2-type non-LTR retrotransposon - consensus.	Latimeria chalumnae
R2	.	R2-1_LDi	R2 non-LTR retrotransposon from cuckoo roller.	Leptosomus discolor
R2	.	R2-1_LSal	non-LTR retrotransposon, consensus.	Lepeophtheirus salmonis
R2	.	R2-1_LV	R2 non-LTR retrotransposon from green sea urchin.	Lytechinus variegatus
R2	.	R2-1_MDe	R2 non-LTR retrotransposon from Hessian fly.	Mayetiola destructor
R2	.	R2-1_MLe	R2 non-LTR retrotransposon - consensus.	Mnemiopsis leidyi
R2	.	R2-1_MR	R2 non-LTR retrotransposon from alfalfa leafcutter bee.	Megachile rotundata
R2	.	R2-1_MUn	R2 non-LTR retrotransposon fragment from budgerigar.	Melopsittacus undulatus
R2	.	R2-1_MUni	R2 non-LTR retrotransposon from brown mesite.	Mesitornis unicolor
R2	.	R2-1_MVi	R2 non-LTR retrotransposon from golden-collared manakin.	Manacus vitellinus
R2	.	R2-1_NNi	R2 non-LTR retrotransposon from crested ibis.	Nipponia nippon
R2	.	R2-1_NV	Starlet sea anemone R2-1_NV autonomous Non-LTR Retrotransposon - consensus.	Nematostella vectensis
R2	.	R2-1_OHo	R2 non-LTR retrotransposon from hoatzin.	Opisthocomus hoazin
R2	.	R2-1_PAd	R2 non-LTR retrotransposon from Adelie penguin.	Pygoscelis adeliae
R2	.	R2-1_PBa	R2 non-LTR retrotransposon from red harvester ant.	Pogonomyrmex barbatus
R2	.	R2-1_PCar	R2 non-LTR retrotransposon from great cormorant.	Phalacrocorax carbo
R2	.	R2-1_PCau	R2 non-LTR retrotransposon sequence.	Priapulus caudatus

Table 2				
R2	.	R2-1_PCr	R2 non-LTR retrotransposon from great crested grebe.	<i>Podiceps cristatus</i>
R2	.	R2-1_PCri	R2 non-LTR retrotransposon from Dalmatian pelican.	<i>Pelecanus crispus</i>
R2	.	R2-1_PGu	R2 non-LTR retrotransposon from sandgrouse.	<i>Pterocles gutturalis</i>
R2	.	R2-1_PLe	R2 non-LTR retrotransposon from tropicbird.	<i>Phaethon lepturus</i>
R2	.	R2-1_PM	R2-1_PM is a family of R2 non-LTR retrotransposons - consensus.	<i>Petromyzon marinus</i>
R2	.	R2-1_PPap	R2 non-LTR retrotransposon from sand fly.	<i>Phlebotomus papatasi</i>
R2	.	R2-1_PPu	R2 non-LTR retrotransposon from downy woodpecker.	<i>Picoides pubescens</i>
R2	.	R2-1_PRR	R2 non-LTR retrotransposon from American flamingo.	<i>Phoenicopterus ruber ruber</i>
R2	.	R2-1_PSi	R2 non-LTR retrotransposon from Chinese soft-shelled turtle.	<i>Pelodiscus sinensis</i>
R2	.	R2-1_RMi	R2 non-LTR retrotransposon from brown tick.	<i>Rhipicephalus microplus</i>
R2	.	R2-1_RPr	R2 non-LTR retrotransposon sequence.	<i>Rhodnius prolixus</i>
R2	.	R2-1_RPu	R2 non-LTR retrotransposon cDNA sequence from brown tick.	<i>Rhipicephalus pulchellus</i>
R2	.	R2-1_SCa	R2 non-LTR retrotransposon from Atlantic canary.	<i>Serinus canaria</i>
R2	.	R2-1_SK	R2 non-LTR retrotransposon from acorn worm.	<i>Saccoglossus kowalevskii</i>
R2	.	R2-1_SM	R2-type retrotransposon from Schmidtea mediterranea: consensus.	<i>Schmidtea mediterranea</i>
R2	.	R2-1_SP	R2 non-LTR retrotransposon from purple sea urchin.	<i>Strongylocentrotus purpuratus</i>
R2	AGKD 0107 2455	R2-1_SSa	R2-type non-LTR retrotransposon.	<i>Salmo salar</i>
R2	.	R2-1_StC	R2 non-LTR retrotransposon from ostrich.	<i>Struthio camelus australis</i>
R2	.	R2-1_TAI	R2 non-LTR retrotransposon from barn owl.	<i>Tyto alba</i>
R2	.	R2-1_TCas	R2 non-LTR retrotransposon from red flour beetle - consensus -.	<i>Tribolium castaneum</i>
R2	.	R2-1_TG	A family of R2 non-LTR retrotransposons - consensus sequence.	<i>Taeniopygia guttata</i>

Table 2				
R2	.	R2-1_TGut	R2 non-LTR retrotransposon from white-throated tinamou.	<i>Tinamus guttatus</i>
R2	.	R2-1_TSP	A family of R2 non-LTR retrotransposons in the <i>Trichinella spiralis</i> genome - a consensus.	<i>Trichinella spiralis</i>
R2	scaffo ld_6	R2-1_TUr	R2 non-LTR retrotransposon from twospotted spider mite.	<i>Tetranychus urticae</i>
R2	.	R2-1_XM	R2 non-LTR retrotransposon fragment from Southern platyfish.	<i>Xiphophorus maculatus</i>
R2	.	R2-1_ZA	R2 non-LTR retrotransposon from white-throated sparrow.	<i>Zonotrichia albicollis</i>
R2	.	R2-1_ZLM	R2 non-LTR retrotransposon from silvereye.	<i>Zosterops lateralis melanops</i>
R2	.	R2-2_API	R2 non-LTR retrotransposon from pea aphid.	<i>Acyrtosiphon pisum</i>
R2	.	R2-2_CCan	R2 non-LTR retrotransposon from common cuckoo.	<i>Cuculus canorus</i>
R2	.	R2-2_CMa	R2 non-LTR retrotransposon from Macqueen's bustard.	<i>Chlamydotis macqueenii</i>
R2	.	R2-2_DWi	28S rDNA-specific non-LTR retrotransposon R2 in <i>Drosophila willistoni</i> .	<i>Drosophila willistoni</i>
R2	.	R2-2_HAI	R2 non-LTR retrotransposon from white-tailed eagle.	<i>Haliaeetus albicilla</i>
R2	.	R2-2_IS	R2 non-LTR retrotransposon from deer tick.	<i>Ixodes scapularis</i>
R2	.	R2-2_MR	R2 non-LTR retrotransposon from alfalfa leafcutter bee.	<i>Megachile rotundata</i>
R2	.	R2-2_MUn	R2 non-LTR retrotransposon fragment from budgerigar.	<i>Melopsittacus undulatus</i>
R2	.	R2-2_MUni	R2 non-LTR retrotransposon from brown mesite.	<i>Mesitornis unicolor</i>
R2	.	R2-2_NNi	R2 non-LTR retrotransposon from crested ibis.	<i>Nipponia nippon</i>
R2	.	R2-2_NV	Starlet sea anemone R2-2_NV autonomous Non-LTR Retrotransposon - consensus.	<i>Nematostella vectensis</i>
R2	.	R2-2_PBa	R2 non-LTR retrotransposon from red harvester ant.	<i>Pogonomyrmex barbatus</i>
R2	.	R2-2_PM	R2-2_PM is a family of R2 non-LTR retrotransposons - a consensus.	<i>Petromyzon marinus</i>
R2	.	R2-2_RPr	R2 non-LTR retrotransposon sequence.	<i>Rhodnius prolixus</i>
R2	.	R2-2_SMed	R2 non-LTR retrotransposon from <i>Schmidtea mediterranea</i> : consensus.	<i>Schmidtea mediterranea</i>
R2	.	R2-2_TCas	R2 non-LTR retrotransposon from red flour beetle.	<i>Tribolium castaneum</i>

Table 2				
R2	scaffo ld_37	R2-2_TUr	R2 non-LTR retrotransposon from twospotted spider mite.	Tetranychus urticae
R2	ABJB 0105 5516 9	R2-3_IS	R2 non-LTR retrotransposon from deer tick.	Ixodes scapularis
R2	.	R2-3_MR	R2 non-LTR retrotransposon from alfalfa leafcutter bee.	Megachile rotundata
R2	.	R2-4_MR	R2 non-LTR retrotransposon from alfalfa leafcutter bee.	Megachile rotundata
R2	.	R2-5_MR	R2 non-LTR retrotransposon from alfalfa leafcutter bee.	Megachile rotundata
R2	.	R2-6_MR	R2 non-LTR retrotransposon from alfalfa leafcutter bee.	Megachile rotundata
R2	.	R2-7_MR	R2 non-LTR retrotransposon from alfalfa leafcutter bee.	Megachile rotundata
R2	.	R2-8_MR	R2 non-LTR retrotransposon from alfalfa leafcutter bee.	Megachile rotundata
R2	.	R2-N1_Gav	Non-LTR retrotransposon.	Gavialis gangeticus
R2	.	R2-N2_Gav	Non-LTR retrotransposon.	Gavialis gangeticus
R2	.	R2-N2B_Gav	Non-LTR retrotransposon.	Gavialis gangeticus
R2	.	R2A_NVi	28S rDNA-specific non-LTR retrotransposon R2 in Nasonia vitripennis.	Nasonia vitripennis
R2	AF01 5817	R2A_TM	Tenebrio molitor retrotransposon R2 reverse transcriptase gene, partial cds.	Tenebrio molitor
R2	.	R2Amel	R2Amel - R2 non-LTR retrotransposon from the honeybee Apis mellifera.	Apis mellifera
R2	AF01 5685	R2B_DM	Drosophila mercatorum R2 retrotransposon reverse transcriptase domain protein gene, complete cds.	Drosophila mercatoru m
R2	.	R2B_NVi	28S rDNA-specific non-LTR retrotransposon R2 in Nasonia vitripennis.	Nasonia vitripennis
R2	AF01 5822	R2B_TM	Tenebrio molitor retrotransposon R2 reverse transcriptase gene, partial cds.	Tenebrio molitor
R2	.	R2C_NGi	28S rDNA-specific non-LTR retrotransposon R2 in Nasonia giraulti.	Nasonia giraulti
R2	AB09 7122	R2Ci-B	Ciona intestinalis retrotransposon R2Ci-B, complete sequence.	Ciona intestinalis
R2	.	R2Ci-D	Ciona intestinalis retrotransposon R2CiD, complete sequence.	Ciona intestinalis
R2	AB09 7121	R2CIA_CI	Ciona intestinalis retrotransposon R2Ci-A, complete sequence.	Ciona intestinalis
R2	AB09 7125	R2Cs-D	Ciona intestinalis retrotransposon R2CsD, partial sequence.	Ciona savignyi
R2	.	R2D_NGi	28S rDNA-specific non-LTR retrotransposon R2 in Nasonia giraulti.	Nasonia giraulti

Table 2				
R2	NM_001030097	R2Dr	R2 non-LTR retrotransposon in the <i>Danio rerio</i> genome - a single copy.	<i>Danio rerio</i>
R2	.	R2E_NLo	28S rDNA-specific non-LTR retrotransposon R2 in <i>Nasonia longicornis</i> .	<i>Nasonia longicornis</i>
R2	AB201408	R2Eb	R2 non-LTR retrotransposon from <i>Eptatretus burgeri</i> .	<i>Eptatretus burgeri</i>
R2	AB201415	R2Ha	R2 non-LTR retrotransposon from <i>Hasarius adansoni</i> .	<i>Hasarius adansoni</i>
R2	JN937617	R2La	R2-type non-LTR retrotransposon.	<i>Lepidurus arcticus</i>
R2	.	R2LcA	R2-type non-LTR retrotransposon.	<i>Lepidurus couesii</i>
R2	JN937619	R2LcB	R2-type non-LTR retrotransposon.	<i>Lepidurus couesii</i>
R2	.	R2LcC	R2-type non-LTR retrotransposon.	<i>Lepidurus couesii</i>
R2	JN937616	R2LI	R2-type non-LTR retrotransposon.	<i>Lepidurus apus lubbocki</i>
R2	AB201414	R2Mr	R2 non-LTR retrotransposon from <i>Metacrinus rotundus</i> .	<i>Metacrinus rotundus</i>
R2	.	R2NS-1_CGi	R2-type retrotransposon from <i>Crassostrea gigas</i> .	<i>Crassostrea gigas</i>
R2	.	R2NS-1_CSi	R2-type retrotransposon from <i>Clonorchis sinensis</i> : consensus.	<i>Clonorchis sinensis</i>
R2	.	R2NS-1_PMi	R2-like non-LTR retrotransposon from bat star.	<i>Patiria miniata</i>
R2	.	R2NS-1_SMed	R2-type retrotransposon from <i>Schmidtea mediterranea</i> : consensus.	<i>Schmidtea mediterranea</i>
R2	.	R2Nvec-A	R2Nvec-A - R2 non-LTR retrotransposon from the starlet sea anemone <i>Nematostella vectensis</i> .	<i>Nematostella vectensis</i>
R2	.	R2OI-A	R2 non-LTR retrotransposon from the medaka <i>Oryzias latipes</i> - consensus.	<i>Oryzias latipes</i>
R2	AB201416	R2Pc	R2 non-LTR retrotransposon from <i>Procambarus clarkii</i> .	<i>Procambarus clarkii</i>
R2	.	R2Sm-A	R2Sm-A - R2 non-LTR retrotransposon from the bloodfluke <i>Schistosoma mansoni</i> .	<i>Schistosoma mansoni</i>
R2	AB201409	R2Ta	R2 non-LTR retrotransposon from <i>Tanichthys albonubes</i> .	<i>Tanichthys albonubes</i>
R2	EU854578	R2Tc	R2-type non-LTR retrotransposon.	<i>Triops cancriformis</i>
R2	JN937621	R2Tc_it	R2-type non-LTR retrotransposon.	<i>Triops cancriformis</i>

Table 2				
R2	AB201417	R2TI	R2 non-LTR retrotransposon from <i>Triops longicaudatus</i> .	<i>Triops longicaudatus</i>
R4	U29445	R4_AL	<i>Ascaris lumbricoides</i> site-specific non-LTR retrotransposable element R4 in 26S rDNA, complete sequence.	<i>Ascaris lumbricoides</i>
R4	U29590	R4_HC	<i>Haemonchus contortus</i> non-LTR retrotransposon specific to the large subunit rRNA genes of nematodes.	<i>Haemonchus contortus</i>
R4	.	R4_Hmel	a R4 element from <i>Heliconius melpomene</i> .	<i>Heliconius melpomene</i>
R4	.	R4-1_AC	A family of R4 non-LTR retrotransposons - consensus sequence.	<i>Anolis carolinensis</i>
R4	.	R4-1_ADi	R4-type retrotransposon: consensus.	<i>Acropora digitifera</i>
R4	.	R4-1_BM	Non-LTR retrotransposon - a consensus.	<i>Bombyx mori</i>
R4	CADV01008175	R4-1_BX	An R4 non-LTR retrotransposon family from <i>Bursaphelenchus xylophilus</i> .	<i>Bursaphelenchus xylophilus</i>
R4	ABLE03011482	R4-1_CJap	An R4 non-LTR retrotransposon family from <i>Caenorhabditis japonica</i> .	<i>Caenorhabditis japonica</i>
R4	.	R4-1_CM	Non-LTR retrotransposon from the elephant shark - consensus.	<i>Callorhynchus milii</i>
R4	.	R4-1_CPB	Non-LTR retrotransposon: consensus.	<i>Chrysemys picta bellii</i>
R4	.	R4-1_ED	Autonomous non-LTR retrotransposon from the R4 clade - a consensus sequence.	<i>Entamoeba dispar</i>
R4	.	R4-1_HG	An R4 non-LTR retrotransposon family from <i>Heterodera glycines</i> .	<i>Heterodera glycines</i>
R4	.	R4-1_HMe	Non-LTR retrotransposon family from <i>Heliconius melpomene melpomene</i> .	<i>Heliconius melpomene melpomene</i>
R4	CABB01003843	R4-1_MI	An R4 non-LTR retrotransposon family from <i>Meloidogyne incognita</i> .	<i>Meloidogyne incognita</i>
R4	.	R4-1_PH	Non-LTR Retrotransposon, consensus.	<i>Parhyale hawaiiensis</i>
R4	CACX01002001	R4-1_SRa	An R4 non-LTR retrotransposon family from <i>Strongyloides ratti</i> .	<i>Strongyloides ratti</i>
R4	.	R4-1_TCa	R4-type retrotransposon: consensus.	<i>Tribolium castaneum</i>
R4	.	R4-1B_AC	Dong-type non-LTR retrotransposons - a consensus sequence.	<i>Anolis carolinensis</i>
R4	.	R4-2_AS	An R4 non-LTR retrotransposon family from <i>Ascaris suum</i> .	<i>Ascaris suum</i>

Table 2				
R4	CADV 0100 9048	R4-2_BX	An R4 non-LTR retrotransposon family from <i>Bursaphelenchus xylophilus</i> .	<i>Bursaphelenchus xylophilus</i>
R4	ABLA 0100 0389	R4-2_HG	An R4 non-LTR retrotransposon family from <i>Heterodera glycines</i> .	<i>Heterodera glycines</i>
R4	CACX 0100 2006	R4-2_SRa	An R4 non-LTR retrotransposon family from <i>Strongyloides ratti</i> .	<i>Strongyloides ratti</i>
R4	CADV 0100 8832	R4-3_BX	An R4 non-LTR retrotransposon family from <i>Bursaphelenchus xylophilus</i> .	<i>Bursaphelenchus xylophilus</i>
R4	.	R4-3_SRa	An R4 non-LTR retrotransposon family from <i>Strongyloides ratti</i> .	<i>Strongyloides ratti</i>
R4	.	R4-4_BX	An R4 non-LTR retrotransposon family from <i>Bursaphelenchus xylophilus</i> .	<i>Bursaphelenchus xylophilus</i>
R4	.	R4-4_SRa	An R4 non-LTR retrotransposon family from <i>Strongyloides ratti</i> .	<i>Strongyloides ratti</i>
R4	.	R4-5_BX	An R4 non-LTR retrotransposon family from <i>Bursaphelenchus xylophilus</i> .	<i>Bursaphelenchus xylophilus</i>
NeSL	AY21 6701	R5	<i>Girardia tigrina</i> R5 retrotransposon, complete sequence.	<i>Girardia tigrina</i>
NeSL	.	R5-1_SM	A family of planarian NeSL non-LTR retrotransposons - consensus.	<i>Schmidtea mediterranea</i>
NeSL	.	R5-2_SM	A family of planarian NeSL non-LTR retrotransposons - consensus.	<i>Schmidtea mediterranea</i>
R2	.	R8Hm-A	R8Hm-A - 18S rDNA-specific non-LTR retrotransposon from <i>Hydra magnipapillata</i> .	<i>Hydra vulgaris</i>
R2	.	R8Hm-B	R8Hm-B - 18S rDNA-specific non-LTR retrotransposon from <i>Hydra magnipapillata</i> .	<i>Hydra vulgaris</i>
R2	.	R9Av	R9Av, an rDNA-specific non-LTR retrotransposon family from rotifer.	<i>Adineta vaga</i>
R2	FJ461 304	RaR2	28S rDNA-specific non-LTR retrotransposon R2 from <i>Rhynchosciara americana</i> .	<i>Rhynchosciara americana</i>
R4	.	Rex6	Non-LTR retrotransposon; site-specific LINE; R4/Dong superfamily; REX6; DONG_FR.	<i>Takifugu rubripes</i>
R4	.	Rex6-1_OL	A Rex6 non-LTR retrotransposon family from <i>Olyzias latipes</i> .	<i>Olyzias latipes</i>
CRE	X170 78	SLACS	<i>Trypanosoma brucei</i> DNA for retrotransposable element SLACS.	<i>Trypanosoma brucei</i>
NeSL	.	Utopia-1_ACa	Utopia-1_ACa is a protozoan Utopia non-LTR retrotransposon - a complete sequence.	<i>Acanthamoeba castellanii</i>

Table 2				
NeSL	scaffo ld_47 4	Utopia-1_ACar	A family of NeSL non-LTR retrotransposons.	Anolis carolinensis
NeSL	.	Utopia-1_AEc	A family of Utopia non-LTR retrotransposons - consensus.	Acromyrme x echinator
NeSL	.	Utopia-1_AMi	A family of NeSL non-LTR retrotransposons - consensus.	Alligator mississippi ensis
NeSL	.	Utopia-1_APi	A family of Utopia non-LTR retrotransposons - consensus.	Acyrtosiph on pisum
NeSL	.	Utopia-1_API	A family of Utopia non-LTR retrotransposons.	Agrilus planipennis
NeSL	.	Utopia-1_CFI	A family of Utopia non-LTR retrotransposons - consensus.	Camponotu s floridanus
NeSL	.	Utopia-1_CMy	A family of Utopia non-LTR retrotransposons - consensus.	Chelonia mydas
NeSL	.	Utopia-1_CPB	A family of Utopia non-LTR retrotransposons - consensus.	Chrysemys picta bellii
NeSL	.	Utopia-1_Crp	Non-LTR retrotransposon.	Crocodylus porosus
NeSL	.	Utopia-1_DPo	A family of Utopia non-LTR retrotransposons.	Dendrocton us ponderosae
NeSL	.	Utopia-1_DPu	A family of Utopia non-LTR retrotransposons - consensus.	Daphnia pulex
NeSL	.	Utopia-1_DYak	A family of Utopia non-LTR retrotransposons - consensus.	Drosophila yakuba
NeSL	.	Utopia-1_EBr	A family of Utopia non-LTR retrotransposons - consensus.	Eimeria brunetti
NeSL	.	Utopia-1_EMi	A family of Utopia non-LTR retrotransposons - consensus.	Eimeria mitis
NeSL	.	Utopia-1_ENe	A family of Utopia non-LTR retrotransposons - consensus.	Eimeria necatrix
NeSL	.	Utopia-1_Gav	Non-LTR retrotransposon.	Gavialis gangeticus
NeSL	.	Utopia-1_GG1	A family of Utopia non-LTR retrotransposons.	Ganaspis
NeSL	.	Utopia-1_HAra	A family of Utopia non-LTR retrotransposons.	Hyaloperon ospora arabidopsid is
NeSL	.	Utopia-1_HG	A family of Utopia non-LTR retrotransposons.	Heterodera glycines
NeSL	.	Utopia-1_HMM	A family of Utopia non-LTR retrotransposons.	Heliconius melpomene melpomene
NeSL	.	Utopia-1_HSal	A family of Utopia non-LTR retrotransposons - consensus.	Harpegnath os saltator
NeSL	.	Utopia-1_IS	A family of Utopia non-LTR retrotransposons - consensus.	Ixodes scapularis

Table 2				
NeSL	.	Utopia-1_LAI	A family of Utopia non-LTR retrotransposons.	<i>Lasioglossum albipes</i>
NeSL	.	Utopia-1_LFu	A family of Utopia non-LTR retrotransposons.	<i>Ladona fulva</i>
NeSL	AGCV 0135 8106	Utopia-1_LV	A family of Utopia non-LTR retrotransposons.	<i>Lytechinus variegatus</i>
NeSL	.	Utopia-1_MRo	A family of Utopia non-LTR retrotransposons - consensus.	<i>Megachile rotundata</i>
NeSL	.	Utopia-1_NVit	A family of Utopia non-LTR retrotransposons - consensus.	<i>Nasonia vitripennis</i>
NeSL	.	Utopia-1_PAlni	NeSL non-LTR retrotransposon from <i>Phytophthora alni</i> .	<i>Phytophthora alni</i>
NeSL	.	Utopia-1_PArrh	NeSL non-LTR retrotransposon from <i>Pythium arrhenomanes</i> .	<i>Pythium arrhenomanes</i>
NeSL	.	Utopia-1_PBa	A family of Utopia non-LTR retrotransposons - consensus.	<i>Pogonomyrmex barbatus</i>
NeSL	.	Utopia-1_PCa	A family of Utopia non-LTR retrotransposons - consensus.	<i>Phytophthora capsici</i>
NeSL	.	Utopia-1_PCinn	NeSL non-LTR retrotransposon from <i>Phytophthora cinnamomi</i> .	<i>Phytophthora cinnamomi</i>
NeSL	AHJF 0100 4292	Utopia-1_PCu	A family of Utopia non-LTR retrotransposons - consensus.	<i>Pseudoperonospora cubensis</i>
NeSL	.	Utopia-1_Pi	A family of NeSL non-LTR retrotransposons - consensus.	<i>Phytophthora infestans</i>
NeSL	.	Utopia-1_PInsi	NeSL non-LTR retrotransposon from <i>Pythium insidiosum</i> .	<i>Pythium insidiosum</i>
NeSL	.	Utopia-1_PKern	NeSL non-LTR retrotransposon from <i>Phytophthora kernoviae</i> .	<i>Phytophthora kernoviae</i>
NeSL	.	Utopia-1_PLate	NeSL non-LTR retrotransposon from <i>Phytophthora lateralis</i> .	<i>Phytophthora lateralis</i>
NeSL	.	Utopia-1_PMi	A family of Utopia non-LTR retrotransposons.	<i>Patiria miniata</i>
NeSL	.	Utopia-1_PPac	A family of Utopia non-LTR retrotransposons.	<i>Pristionchus pacificus</i>
NeSL	.	Utopia-1_PPini	NeSL non-LTR retrotransposon from <i>Phytophthora pinifolia</i> .	<i>Phytophthora pinifolia</i>
NeSL	.	Utopia-1_PR	A family of Utopia non-LTR retrotransposons - consensus.	<i>Phytophthora ramorum</i>
NeSL	.	Utopia-1_PRe	A family of Utopia non-LTR retrotransposons - consensus.	<i>Panagrellus redivivus</i>
NeSL	.	Utopia-1_PS	A family of Utopia non-LTR retrotransposons - consensus.	<i>Phytophthora sojae</i>

Table 2				
NeSL	.	Utopia-1_PSi	A family of Utopia non-LTR retrotransposons - consensus.	Pelodiscus sinensis
NeSL	.	Utopia-1_PT	A family of Utopia non-LTR retrotransposons.	Parasteatoda tepidariorum
NeSL	ADOS 0100 1321	Utopia-1_PU	A family of Utopia non-LTR retrotransposons.	Pythium ultimum
NeSL	.	Utopia-1_PVexa	NeSL non-LTR retrotransposon from Phytophthora vexans.	Phytophthora aff. vexans
NeSL	.	Utopia-1_SaPa	A family of Utopia non-LTR retrotransposons.	Saprolegnia parasitica
NeSL	.	Utopia-1_SDicl	NeSL non-LTR retrotransposon from Saprolegnia diclina.	Saprolegnia diclina
NeSL	.	Utopia-1_SM	A family of Utopia non-LTR retrotransposons.	Strigamia maritima
NeSL	AAGJ 0214 0537	Utopia-1_SP	A family of Utopia non-LTR retrotransposons.	Strongylocentrotus purpuratus
NeSL	.	Utopia-1_TSP	A family of Utopia non-LTR retrotransposons.	Trichinella spiralis
NeSL	.	Utopia-1B_CPB	A family of Utopia non-LTR retrotransposons - consensus.	Chrysemys picta bellii
NeSL	.	Utopia-2_API	A family of Utopia non-LTR retrotransposons.	Acyrtosiphon pisum
NeSL	.	Utopia-2_CMy	A family of Utopia non-LTR retrotransposons - consensus.	Chelonia mydas
NeSL	.	Utopia-2_CPB	A family of Utopia non-LTR retrotransposons - consensus.	Chrysemys picta bellii
NeSL	.	Utopia-2_DPu	A family of Utopia non-LTR retrotransposons.	Daphnia pulex
NeSL	.	Utopia-2_LFu	A family of Utopia non-LTR retrotransposons.	Ladona fulva
NeSL	.	Utopia-2_PCa	A family of Utopia non-LTR retrotransposons - consensus.	Phytophthora capsici
NeSL	.	Utopia-2_PI	A family of NeSL non-LTR retrotransposons - consensus.	Phytophthora infestans
NeSL	.	Utopia-2_PR	A family of Utopia non-LTR retrotransposons - consensus.	Phytophthora ramorum
NeSL	.	Utopia-2_PS	A family of Utopia non-LTR retrotransposons - consensus.	Phytophthora sojae
NeSL	.	Utopia-2_PU	A family of Utopia non-LTR retrotransposons.	Pythium ultimum
NeSL	.	Utopia-3_CPB	A family of Utopia non-LTR retrotransposons - consensus.	Chrysemys picta bellii
NeSL	.	Utopia-3_DPu	A family of Utopia non-LTR retrotransposons.	Daphnia pulex

Table 2				
NeSL	.	Utopia-3_LFu	A family of Utopia non-LTR retrotransposons.	Ladona fulva
NeSL	.	Utopia-3_PCa	A family of Utopia non-LTR retrotransposons - consensus.	Phytophthora capsici
NeSL	.	Utopia-3_Pl	A family of NeSL non-LTR retrotransposons - consensus.	Phytophthora infestans
NeSL	.	Utopia-3_PR	A family of Utopia non-LTR retrotransposons - consensus.	Phytophthora ramorum
NeSL	.	Utopia-4_LFu	A family of Utopia non-LTR retrotransposons.	Ladona fulva
NeSL	AATU 0100 1281.1	Utopia-4_Pl	A family of NeSL non-LTR retrotransposons - a copy.	Phytophthora infestans
NeSL	.	Utopia-4_PR	A family of Utopia non-LTR retrotransposons - consensus.	Phytophthora ramorum
NeSL	.	Utopia-5_LFu	A family of Utopia non-LTR retrotransposons.	Ladona fulva
NeSL	.	Utopia-5_Pl	A family of Utopia non-LTR retrotransposons - consensus.	Phytophthora infestans
NeSL	.	Utopia-5_PR	A family of Utopia non-LTR retrotransposons - consensus.	Phytophthora ramorum
NeSL	.	Utopia-6_LFu	A family of Utopia non-LTR retrotransposons.	Ladona fulva
R4	.	X4_LINE	Conserved LINE element reconstructed from the human genome - consensus.	Vertebrata
NeSL	.	YURE_CSa	A NeSL non-LTR retrotransposon from Ciona savignyi.	Ciona savignyi
R2	.	YURE-2_Cis	YURE non-LTR retrotransposon from Ciona savignyi.	Ciona savignyi
NeSL	.	YURECi	Ciona intestinalis retrotransposon YURECi.	Ciona intestinalis

A skilled artisan can, based on the Accession numbers provided in Tables 1-3 determine the nucleic acid and corresponding polypeptide sequences of each retrotransposon and domains thereof, e.g., by using routine sequence analysis tools as Basic Local Alignment Search Tool (BLAST) or CD-Search for conserved domain analysis. Other sequence analysis tools are known and can be found, e.g., at <https://molbiol-tools.ca>, for example, at <https://molbiol-tools.ca/Motifs.htm>. SEQ ID NOs 1-112 align with each row in Table 1, and SEQ ID NOs 113-1015 align with the first 903 rows of Table 2.

Tables 1-3 herein provide the sequences of exemplary transposons, including the amino acid sequence of the retrotransposase, and sequences of 5' and 3' untranslated regions to allow the retrotransposase to bind the template RNA, and the full transposon nucleic acid sequence. In some embodiments, a 5' UTR of any of Tables 1-3 allows the retrotransposase to bind the

5 template RNA. In some embodiments, a 3' UTR of any of Tables 1-3 allows the retrotransposase to bind the template RNA. Thus, in some embodiments, a polypeptide for use in any of the systems described herein can be a polypeptide of any of Tables 1-3 herein, or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto. In some embodiments, the system further comprises one or both of a 5' or 3'

10 untranslated region of any of Tables 1-3 herein (or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto), e.g., from the same transposon as the polypeptide referred to in the preceding sentence, as indicated in the same row of the same table. In some embodiments, the system comprises one or both of a 5' or 3' untranslated region of any of Tables 1-3 herein, e.g., a segment of the full transposon sequence that encodes an RNA that is

15 capable of binding a retrotransposase, and/or the sub-sequence provided in the column entitled Predicted 5' UTR or Predicted 3' UTR.

In some embodiments, a polypeptide for use in any of the systems described herein can be a molecular reconstruction or ancestral reconstruction based upon the aligned polypeptide sequence of multiple retrotransposons. In some embodiments, a 5' or 3' untranslated region for

20 use in any of the systems described herein can be a molecular reconstruction based upon the aligned 5' or 3' untranslated region of multiple retrotransposons. A skilled artisan can, based on the Accession numbers provided herein, align polypeptides or nucleic acid sequences, e.g., by using routine sequence analysis tools as Basic Local Alignment Search Tool (BLAST) or CD-

Search for conserved domain analysis. Molecular reconstructions can be created based upon sequence consensus, e.g. using approaches described in *Ivics et al., Cell 1997, 501 – 510* ; *Wagstaff et al., Molecular Biology and Evolution 2013, 88-99*. In some embodiments, the retrotransposon from which the 5' or 3' untranslated region or polypeptide is derived is a young
 5 or a recently active mobile element, as assessed via phylogenetic methods such as those described in Boissinot et al., *Molecular Biology and Evolution* 2000, 915-928.

Table 3 (below) shows exemplary Gene Writer proteins and associated sequences from a variety of retrotransposases, identified using data mining. Column 1 indicates the family to
 10 which the retrotransposon belongs. Column 2 lists the element name. Column 3 indicates an accession number, if any. Column 4 lists an organism in which the retrotransposase is found. Column 5 lists the DNA sequence of the retrotransposon. Column 6 lists the predicted 5' untranslated region, and column 7 lists the predicted 3' untranslated region; both are segments of the sequence of column 5 that are predicted to allow the template RNA to bind the
 15 retrotransposase of column 8. (It is understood that columns 5-7 show the DNA sequence, and that an RNA sequence according to any of columns 5-7 would typically include uracil rather than thymidine.) Column 8 lists the predicted retrotransposase sequence encoded in the retrotransposon of column 5.

[illegible]

GAGATAATGGACCTGGGAGACGCCCTCTGCCAGGCGAGGCGAGCTGCCG GAGGAGAAACAATGGACGAGATAAGACGCCACCTGTATAAGGGCAAC GGACAGCAGAGACCCACCAAGCAAAATCAGAAAGAACAGCTGCAGGCT TACTATAAAAGACACTAGAGAACGACTTTCAGCTGGGCACTTAACA CCTTCCCCGAGCATTCAAGCAGGTAATGGAAGCCGGATATAAAGCT AGTAATCAATCAGACGCGCAGGACTCTCGGATGCTGGAATCCATA AGCCAAATAAGAACGGCAACCCGAGATAAAAGGACACGGTGACCCGG GAGAAACACCCAAAGAAACCTTTTTCAGAAAGTGATGAAGGACAGACA ATCAAAAAGGTAATTAATCTCGGTTCCAGCGTTATTTTATCTTGATAG AGGAAACTGGCTAAATCATTTTAGATGATATGAATGCTTGCTTGT GACATACCACTCAGTGAATTTATTCGGTTTTTAAACAAGATGGAAA CACTGGTAGCTTTAAAGCCTTGGGACTTTAAAACTTACGGGAAGC TGACAACACTGCCTTCAGAGAAATTAATACGGCTAAAGAAATTGAGAA AATGTGCAGGAAATGAGCAAGGCTCGGCTCCCGTCCAGACGGGATT ACTCTGGGACGTCGTAAGATGATCCCGAGTTTTCCCGGACCATGG AGATTTTCAATTTATGGTTAACAACTGGTAAATCCCGGACATGGTGAG GGGGTGCAAGCCGTTTTGATTCAAAATCATCAAGCCGGATCGTTTG AAAGACATTAATACTGGAGACTATCACGATCGGTCCATCTTGCTGA GACTGTTCTCAGGATTGTAACAGCTAGGCTGAGCAAGCGTGCCCT GAACCAAGGCAAGAGGCTTTATCAGAGCGCGGGATGCTCTGAAAA CTTAAACTCTGCAAACTATAATTTGTCGGCCAAAGAGAACACAGA CCACTGGGTGTTGATTCGTGGACATCGCAAGGCTTTTGACACCGTAA GCCACGACACATCATTCATGCTTTCAGCAAGAGAGGTGGATCCCA CATCGTCGTCTGGTGAGCAATATGTACGAGAACATCAGTACGTATATC ACCACAAAGAGGAACACACACAGACAAAATCCAGATCCGGTTGGA GTAAAGCAGGGTGACCCGATGTCGCCCTTTTATTAACCTGGCAATGG ACCTCTATTATGCAAGCTGGAAGAGAGTGGCAAGGATACACCCGAG GACAGAGCATCACAGCGATGGCATTTGCAGACGATCTGGTTTGCT GAGCGACTCTGGGAAATATGAATACAAATATTAGCATCTGGAGACC TTCTGCAATCTGACCGGTCTCAAAACAGGGGCAAAAGTCCACGGCT TTTACATCAAGCCGACAAAGGACTCTTACACCATCAATGACTGCGTGCC TGGAATATCAACGGCACACCCCTGAACATGATCGACCCCGCGAATCTG AGAAATACCTCGGCCCTGCAATTTTGCTTCAGCGGATCGATCAAGGTC CGGTCTCTCCAAAAAATAAGATTTTGGCTTCAGCGGATCGATCAAGCAC CACTTAAACCTCTGAGAAAATGATATTCTCAAAACATACACCATCCCT CGGCTGATCTACATAGCTGACCACTCAGAAAGTGAAGAACTGCACTACTCG AAACCTTGACCAAGAGATCCGGACAGCGGTCAAGGAATGGCTTACCT ACCTCCGTGACACCTGCGGATGCCATCCTGTACTCGAGCACGAGACGGC					ACAGAC TTTTACT ATTTTACT AACGCG TCAATTA CCACCT GATTTG GACCAA TTCACG GGATTT GTCCAA GGTGGA CGGGCC ACCTTTA CTTAACC CGGAAA AGGAAC ATATATA ATTTATG TGTGTC GATAAA (SEQ ID NO: 1263)	MDEIRRHDPKNGGQ QRPTKQKSEELQAY YKKTLEERLSAGALNT FPRAFQVMEGRDIK LVINQTAQDCFCGLE SISQIRTAIRDKKDTV TREKHPKKPFQKWM KDRAIKGNLYRFQR LFYLDRGKLAHILDDI ECLSCDIPLEIYSVFK TRWETTSFKSLGDF KTYGKADNTAFRELIT AKEIEKNVQEMSKGS APGPDGITLGDVVK MDPEFSRTMEIFNL WLTGKIPDMVRGC RTVLIPKSKPDRLKDI NNWRPITIGSILLRL SRIVTARLSKACPLNP RQGFIRAAAGCSENL KLLQTIWSAKREHRP LGVFVDIAKAFDTV SHQHIIHALQQRVED PHIVGLVSNMYENIS TYITTKRNTHTDKIQI RVGVKQGDPMSPLL FNLAMDPLLCLEES GKGYHRGQSSITAM AFADDLVLLSDSWEN MNTNISILETFCNLTG LKTQQQKCHGFYKIP TKDSYTIINDCAAWTI NGTPLNMIIDPGESEK YLGQFDPWIGIARS GLSTKLDLFWLQRIDQ APLPLQKTDILKTYTI PRLIYADHSEVKTALL ETLDQKIRTAKEWL

R2	R2-1_Gfo	.	Geospiza fortis	GGTTGGGCATCACCAAAATTGGCAGGACTGATCCCGAGCGTGCAGGCCGTAGACTGCGATCGGATCGCACAGTCACTGACGATACGATGAAATGCTTCATGAAAAAGAGAAAAATGGAACAGCTGCATAAGAAAATTGTGGATTCAAGCTGGAGGGACAGAGAACATACCTCGATTGGGAAGCACCACGTGAGTGAACCCCAACACGAGTGCACAAAATTCGGAATGGGAAGCACCAGCCAGAAAGATAAATTCCAAAGCCTTGCAATTGGAGGAAAAACGAATTCAAAAATGGACCAATGGCATCCCAAGCGCGGAAATTGTAAATTTGAAAGAGACAAAATTAGTAACCATTTGGATCCAATCTACAGACGCATACCTCACAGGAACTCTACTGCACATACTCACTCAGGCCAACGTTTACCCACGAGAGAAATTTAGCCAGGGGTAGACAAGACCAATACATCAAGCGTGTAGGCACTGCGATCGGACATTGAATCCTGCGCCACATCATCGCAACTGCCCACTGACACAGGACGCCCAATCAAGAGGACACAATTATCTGCGAACTGCTTCTCGAGGAGCGGAAGAAAGAAAGACTGGGTAGTGTTCAAGGAAACCGCACATAAGGGATTCCAACAAGGAACTGTACAAACCTGACCTGATATTTGTGAAGGATGCCCGTGCACTTGTGCTGGATGTGACAGTACGGTATGAAGCAGCAAAATCATCGCTGGAGGAAGCCGCTGCAGAGAAAGTGAGAAAAGTACAAACACCTGGAAACGGAAGTAAGACATCTCACGAAATGCAAAAGGACGTTACTTTGTGGCTTCCCTAGGAGCGCGGGGAAATGGCACCAAGATAACTTTAACTTTGACTGAGCTTGGCTCTCCAAATCGAGGCAAGTGAATAAGCAGAGACTTTCCACAGTAGCGCTCTTTCATCTGTGACATTGTACATAATGTTGCCAGTAGGGCCAGAAAAATCTATGTTATGTAATTCAGGTATTAGATGCTAGTTTGTACCTTCTGTTTGTTTAGGATTTGATAGTGTAGTATTTTATATTTGTACGATTGCATAATGTTCTTTTATACAGTCTGTTTAAATAAAATAGACGATAGCTAGAGACGTTAGGGCAGCCACAGCCAGTTAGGTAGCGGATAGTAGGTAGGACAGACTTTTACTATTTCAACGCGTCAATTACCACCTGATTGGACCAATTCACGGGATTTGTCCAAGGTGGACGGGCCACCTTACTTAACCCGGAAGAAGACATATATAATTTATGTGTTCGATAAA (SEQ ID NO: 1539)	AGACTTAAAGTGAGTTTGGTTACAACTGGGCATAGCTGCAGAGACCGCGCTCTCTCGGCGCCGCTGGTAAGCCCTTAAACAGGGTGACTAAGTCGGTCTCTGCCCCAGTCCGGAGTCTGATGGGACTACACGCCAACGATTCTCTCCAAAAATTCGGTGAACAAAAATTTCTCGGTGCAAGTCGCAAGGCTGTGACCCGAAACCTAGCCCCCGTCCGTCAGGGGCAACGGGTTCCGGAAGTGGGATGGCCACCCCGTTCCCGCAGACGAATCCGGCCATGAATCTGATCCATCTCTTAGGGAGGAGCTGCGGACAAACCCGACGCTTACTAAGGCAATCGGTTGGCACCAGACCTCCCGAGATGATATTTTACCATCTAAACCAACCAATTGACAGAGAAATGAATTGGACTTGTGCTGTGAACCTTCTTTAGAAATGTATAGGTCAGATCTGCAGGGATTGTGTCAGGAGGGGATTCT	AGACTTAA GTGAGTTT GGTTACAA CTGGGCAT AGCTGCAG AGACCGCG CCTCCTCG CGGCCCGG CTGGTAAG CCCTTAAC	GGTAGA TAATCTT TGTATA GTGGGG GGGGAT CTCATGT ACCGGG TTTCTTT TATTTGA TTTTCAA	HLPCTCDAILYSTR DGGLGITKLAGLPSV QARRLHRIAQSSDDT MKCFMEKEKMEQL HKKLWIAQAGGDRENI PSIWEAPPSSEPPNN VSTNSEWEAPTQKD KFKPCNWRKNEFKK WTKLASQGRGIVNFE RDKISNHWIQYRRI PHRLLTALQLRANV YPTREFLARGRDQY IKACRHCADIESCA HIIGNCPVTQDARIKR HNYICELLLEEAKKDD WVVFKEPHIRDSNKE LYKPDILFVKDARALV VDVTVRYEAAKSSLE EAAAEEKVRKYKHLET EVRHLTNAKDVTFVG FPLGARGKWHQDNF KLLTELGLSKSRQVK MAETFTVALFSSVDI VHMFASRARKSMV M (SEQ ID NO: 1016)

ATTTTCTGTGAATAGGGAGGTGTTAGAGGGGTTTCTGAGGTGATGA ACAACCTGCACCACCAACCGCAGTAGGGACGATTTAAACACCACTCTC CCACCGGACAATAATATATCGTACTTGAAGAGGGTAGCAGTGAAGCA GTGGAGGATGGCACACCGGAGGTAGCGCACCCCGTGCTGAAACCCAG GGCAAAGAGTCACCGAATAACATCGTGATGTAACCTTCCCAACAAAA ATCCACCATGCTCTGTAGGGTCAGACTGCATTCAGTACTGGCTCTG ATTGAACATCTTAAGGGTGCATGGGAAGAAGAGGCATGCTTTAGG TGTGCAAGTGTGGGAGGAGAACTTTAACTATCATAGTACTGTTTGTG ACATCGCAAAATGCAAGGGACCAAAAGTTGAGAAGGCCCAAGTGGGAG AGTGGATCTGTAGGTATGTGTAGGACTTTACAACCAAAATCGCCT GGGACAACATAAAAGATTGGCACATCCCTTGTTGTTAGAAACCAAGAAAG GATCGATGCTTCCCAACCGAAGGAGACATCAAAACAGAGGAGCCCAAA GAGATGTTGGACAAAAGAGGAGGAGGAGATGCTGATAAAGTTGGAGG TACAGTTCGAGGGACACAGAAACATCAATAAGCTTATCGCGAACACTT AACAACTAAACATCCAAACAGATTAGTGATAAAGGAGACTATTACCC AGAAAACAATTAACAGATCTAAGTAAGGGAGTGGCTGGACAGAAGTG CTGGACCCAGGACTGAGTCATCAACCCAGCTGGGGGTAGTTGACAAT GGACTTGGTGGGGTCACTGCCAGGGGGGCCAGCTGCTGAAGGAAG AACAAATAGAGCCATTAGGACACACCTTGATAAGGATAACGGTCACCGG GAAATCGCTGACGACACAGGCGAGGGAGGCTGCAGGCCCATTAACGA AAGAAGATAAGGAAGCGCTTTCAGAAAGGATGATTAGCAACTTCCCC GAAGTATTGAACAACACTGACTGCCAGGAAGCACAAACCATTGATCA ATCAAGCAGCGCAGGATGCTTTGGATGCTGGATTCAGCAAGCCAGAT AAGGAAGGCGCTCCGAAACAGAACACACAGAAAGACCAAGGGGATC AACCCAAAGACCACTCAGAGTGATGAAAAAAGAGCAGTTAAGA GGGTCACTTCTCCGTTTTCAGAAATTTTCATCTTGACAGGGGAA ATTGGCAAAGATTATTTGGACGACGTAGAGTGTTCCTGTGTATATA CCACCCAGTGAAATTTATCGGTATTCAAAGCCCGATGGGAAACACTG GACAGTTGCTGGCTTGGGATTCGAAATTAATAGGAAGCGAACA ATAAAGCCTTCAGGACTTAATTACGGCCAAAGAAATTCCTCAAAATGT GCGGAGATGACCAAGGGCTCGGCCCCAGGTCCAGATGGATCGCGCT TGGGGACATCAGGAAGATGGACCCCTGAGTACACCCGGACCGCCGAAC CTTCAACTTATGTTAACATCTGGTGAGATCCCGGACATGTTGAGGGG TGCAGAACTGTGTTAATCCCAAAATCGTCAAAACCGGAACGCCTGAAG ACATCAATAACTGGAGACCATCACGATTTGGATCCATCTTGTGAGACTT TTCTCCAGGATCAACAGCGAGGTTAACAAGGGCGTGCCTCCCAACC CTAGGCAAGAAAGCTTCATCAGTGGGAGGATGCTCCGAGAACTTGA AGCTCTGCAAAACCAATAATTCGGACTGCTAAAAATGAACACAGACCACT	AGGGTGA CTAA (SEQ ID NO: 1141)	TAAAC AGACGG TAGCTA GGTTCG CAAGGC AGCCAC AAGCCA AAGATA GGTAGG GTGCTC ATAGTG AGTAGG GACAGT GCCTTT GATTCA CAACGC GTCAAT ACCATCT GACACG GATACC CTTACCG GACTTG TCATGAT CTCCCA GACTTG TCCAAG GTGGAC GGCCA CCTTAC TTAACC GGAAAA GGAACA TATATTA ATTATAT GTGTC GAAAA (SEQ ID	PAPQPAVGGDDLNTSL PPDNNICVLEKGSSE AVEDGTPEVAHPVPE TQKESPNIVMVTL PNKNPPCCRVRLH SVLALIEHLKSHGKK RACFRVCVCGRENFN YHSTVCHIAKCKGPK VEKAPVGEWICEVCG RDFTTKIGLQHKRL AHPLVRNQRIDASQ PKETSNRGAHKRCW TKEEEMLIKLEVQFE GHRNINKLIAEHLTK TSKQISDKRRLPRKQ LTDLSKGVAGQKVL PGLSHQPLGWVDN GLGGHLPGGPAE GRTIEPLGHLDKDN GHREIADQHKAGRL QAHYRKKIRKRLSEG MISNFFEVFEQLDC QEAQPLINQAAQDC FGCLDSASQIRKALRK QNTQKQDQDQPKR PAQKWMKKRAVKR GHFLRFQKLFHLDRG KLAKIILDDVECLSDI PSEINVSFKARWETP QGFAGLGDFFINRKA NNKAFRDLITAKEILK NVREMTKGSAPGPD GIALGDIRKMDPEYT RTAELFNLWLTSGEIP DMVRGCRVTLIPKSS KPERLKDINNWRPITI GSILLRFSRIITARLTK ACPLNPRQSFISAA

GGGTGTGATTCTGGGACATCGCCAAAGGCTTTGACACCGTGAGCCAC CAACACATCATACATGTATTGCAAAGGAGGAGAGTGAGACCCACATCA TTGGATTGGTAAAAATATGTACAAAGACATCAGTACGGTTATCACCAC AAAGAAGACACATACACGGACAAAATCCAGATCCAGGTGGAGTAA GCAAGGTATCCGCTTTCGCCCTTCTATTCAACCTGGCGATGACCCCC TGTTGTCAAGCTGGAAGAACACGCAAGGATCCACGAGGACAGA GCAAGATAACAGCATGGCATTCGCTGATGACCTGGCTGTTGAGCGA TTCTGGGAAGACATGAATGGGAACATCAAGATACTGGAGACCTTCTGC GACCTACCGGTCTCAAAACACAGGGTCAAAAGTGCCACGGCTTCTACA TCAAGCTACAAAGGACTTTACACCGTCAACAACTGCGCTGCGTGGAC CATCAATGGCACACCCCTGAACATGATCAACCCCGGGGAATCAGAGAAA TACCTCGGCTGCAGTTGACCCCTGGGTGGGAATTGCAAAGACCAAGCC TCCCGAAAAAAGTGGACTTCTGGCTCGAACGCATTTGATCGAGCTCCACT CAAAACATTTCAAGAACTGGACATTTCTTAAGACATACACCATACCTCGAC TGACCTACGTAGCTGACCACTCAGAGATGAAAGCGGGGCCCTTGAAG CCCTTGACCGGACAATTGATCGGCGGTCAAGGACTGGCTGCACCTACC TTCGAGCACCTGTGATGCCATCTTTACACGAGCATGAAGGACGGTGGT TTGGAGTGACCAAAATGGTGGGACTGATTCGAGTGATCAAGCCCCGG AGGCTGCACAGGATTGCGCAGTACCGGAGGAGACGATGAAAGACTTC CTGAAAAAGGCCAGATGGAGAAGATGTACGAGAAATTTGGGTCCAA GCTGGAGGGAAAAAGAGAGGATGCCGTCAATTTGGGAAGCGCTCCC GGAGTTGTACCATCCATAGACACAGCCACAACCTTCGGAGTGGGAAGC ACCGAACCTAAAGTAAAGTACCTAGACCTTGTAAATGGCGCAGAAAA GAATTTAAAAAGTGGACTAAATTAATAGCCAGGGCTGGGAATTAGG TGTTTTAAGGGGACAAAATTAGTAACAATGGGATTCGACATTAGAT ACATACCTCAGGAACTTCTACTGCCATACAGCTCCGGGCCAGTGT GTACCCCAAGGGAAATTTCTCGCGGGGGAGGAGTCTGTGCCACATC TAACTTTGTAGGCACTGTAGGCGGCGAGAGAGTCTGTGCCACATC ATCGGATGTGTCCAGTGTGAGGGATGCCGAATCAAGAGGCACAAT CGCATTTGCGAGAGGCTGATGGAGGAGGGGGGAGAGGAGGACTGGAC GGTGTTCAGGAGCGCACATAAGGACGTACCAAGGAACTGTACAA ACCGACTTGATATTCTGTGAAAAGAGGCTTGCATTTGTGGATGTT ACAATACGGTTCGAGTCAACCAAGACAACTGGAGGAGGCTGCTGCA GAGAAGGTGAACAAGTACAAACATCTGGAGCCGGAAGTACGGAACCTC ACCAACGCTAAGGACGTTATCTTTATGGGGTTTCCCTTGGAGCGGG GACAATGGTACAATAAGAACTTTGAACCTTTGGACACTTTGGCCTCCC AGATCGAGGAGGACATTAATGCAAGACTTTATCCACGGACGCGCTCA TTTCATCTGTGGACATTATACATATGTTTGCCAGTAGAGGCAGAAAGACA					NO: 1264)	GCSENLKLLQTIIRTA KNEHRPLGVVFDIA KAFDTVSHQHIIHVL QRRVDPHIIGLVKN MYKJISTVITTKNTY TDKIQVGVKQGD LSPLLFNLAMDP LCK LEEKGFGHRGQSKI TAMAFADDLVLLSDS WEDMNANIKILETFC DLTGLKTQGGQKCHG FYIKPTKDSYTVNNCA AWTINGTPLNMINP GESEKYLGLQFDPWV GIAKTSLEKLDLFWLE RIDRAPLKPFEQLDIL KTYTIPRLTYWADHSE MKAGALEALDRIRS AVKDWLHLPSTCDA ILYTSMDKGGGLVTK LVGLIPSVQARRLHRI AQSPREETMKDFLEKA QMEKMYEKLWWQA GGKRKRMPISIWEAL PEVPSIDTATTSEW EAPNPKSKYPRPCN WRRKEFKKWTCLIA QGWGIRCFKGDKISN NWIRHYRYPHRLKLT AIQLRASVYPTREFLA RGREDNCVKSCRHCE AAEESCAHIIIGMCPV VRDARIKRHNRICERL MEEAGKRDWTVFQE PHIRDVTKEYKPDLI FVKEGLALVVDVIRF ESTKTTLEAAAAEKVN KYKHLETEVRNLTNA

R2	R2-1_ZA	.	Zonotrichia albicollis	GCATGCTAGGGTAGATAATCTTTGTATAGTGGGGGGGATCTCATGTA CCGGTTCTTTTATTTGATTTCAATAAAAAAGACGGTAGCTAGGTTCCG CAAGGCAGCCAAAGCAAGATAGGTAGGGTCTCATAGTGAGTAGG GACAGTGCCTTTTGAATCACAACGCGTCAATACCATCTGACACGGATACC CTTACCGGACTTGTATGATCTCCAGACTTGTCCAAGGTGGACGGGCC ACCTTTACTTAACCCGGAAAGGAACATATATTAATATATGTGTTCCGA AAA (SEQ ID NO: 1540)	CGACTTGA GAAGTCT GGTTACAA CTGGGCAT AGCTGCAG AGATCGCG CCTCCTCG TGGCCCG CTGGTAAG CCCTTAAC AGGGTGA CTAAGTCG ATCTCTGC CCCAGTCC AGGAGCC GCTGGGT TCACCAGC CCAGCGAT TCCTTCCA AATTCGGT GA (SEQ ID NO: 1142)	GATGTC ACATTG CACTTTC TGTAAC TGCACT GGGTGT GGGATG TGGGCC TGGGGT TGGGGT TATGGG GTATAT ATGTGG GATATTC TGGTGG GAATGT CCATTCA CTGTAT GCCTATC TTTTAA TAAAAA GACGGT AGCTAG GTTCCG GAAGCA GCCACA AGCCAA TAGCCA GTTAGG TAGCTC ATAGTG	KDVIFMGFPLGARG QWYNKNFELLDTLGL PRSRQDIIAKTLSTDA LISSVDIIHMFASRGR RQHA (SEQ ID NO: 1386)
R2	R2-1_ZA	.	Zonotrichia albicollis	CGACTTGAGAAGGTCTGGTTACAACCTGGCATAGCTGCAGAGATCGCG CCTCTCTGCGCCCGCTGGTAAGCCCTTAACAGGGTGACTAAGTCGAT CTCTGCCCAAGTCCAGGAGCCGCTGGTTTACCAGCCCAAGGATTCCTT CCAAATTCGGTGAACAAATTCCTCGGTAAAGCCGCTGGCTTATTGC CTGAAAACCTGGCCCGCTTTCAGACAGGGGCAAGAGTTCGGAAGT GGACTGACCAACCCAGCCGAAACCGAGAGCGAATCTGGTCATGACCCAA CTGTCCCAAACTCTGGTCCGTCTCTTGGAGCGGGGAAGGTGCACAGCC ACTACCTTACTCAGGGTATCGGTGGCACCCCAACCTGTGAAGAGGAC TTATAACATCTAGACCAACCAAAATACCCGGAATTGAATCAGAAATTAGG CCGCTGGTGAAGTTTCTTACAGGTTTACAGGTGAGATCTTAAGGGG GATGTCAATTTAGGGGATTCATTTCCAGATAATGGGGGTACTGG AGGGTTTCTGAGGTGTACGAACAACTGGCACCAACGCAACCAACGGGG GAGACGAGTTAAATCATAGTCTCCAGGGGACAGGGAGGGGATGTAC TTGAGAAGGATAGCAGCAAAAGGAGAGAGGGCTGCACCAAGAGGCA TTGCCCTCAGTGCAAAAGGCCCGCAGTGAACAGTTGCCAGATAACATCG TAAAGGTGACTGTTCCGACAAAATCCACCATGTCCCTGTGTGGTGT CCGCTTAACTCAGTTAGTCTGATTGAACATCTGAAGGGCTCACAC GGGAGGAGGGGTGTCTTTAGGTGTGCCAAATGTGGAGGGAGAA TTTTAACCATAGTACTGTTGTCTTACGCAAAAGTCAAAAGGTCCAC AGATTGAAAGGCCACAGTGGGAGAGTGGATCTGTGAGGTATGCGGA AGGACTTTCAGACCAAAATTTGGCTGGGACAAACACAAAGACATATG CATGCAATGGTGAGAAACAGGAAAGGATCGATGCTTCCCAACCGAAA GAGACATCAAAATCGAGGAGCCCAAGAGGTGTGTCGACGAAGGAGGA GGAAGAACTGCTCATGAAGTTGGAGGTACAGTTTGAGAAATCACAAAA CATCAATAAGCTTATCGCAGAGCAATTAACAACTAAAAACAGCTAAACAA ATTAGTGATAAAAGGAGAATGTGCTCAAAAAGGTAGGGGAGACAAT GGTAATTTGAAACAGAGCCCTGGGATGAGTCATCAATCGCAGGCAAAA GTTAAGGACAAATGGACTGGTGGGAGCCATCTCCGGGAGGACCAAGT GTCGATAAGGGAAACAAATAGGGAAGCCAGGACAACTCTTGACACAGAT AACAGCCATCAAAATACTGTGCAAGAAAGAAAGGGGAGGGCTGCA GGCTCGTTATAGAAGGAGAATAATGAACGATTAGCGGCCGGGACAAT	CGACTTGA GAAGTCT GGTTACAA CTGGGCAT AGCTGCAG AGATCGCG CCTCCTCG TGGCCCG CTGGTAAG CCCTTAAC AGGGTGA CTAAGTCG ATCTCTGC CCCAGTCC AGGAGCC GCTGGGT TCACCAGC CCAGCGAT TCCTTCCA AATTCGGT GA (SEQ ID NO: 1142)	GATGTC ACATTG CACTTTC TGTAAC TGCACT GGGTGT GGGATG TGGGCC TGGGGT TGGGGT TATGGG GTATAT ATGTGG GATATTC TGGTGG GAATGT CCATTCA CTGTAT GCCTATC TTTTAA TAAAAA GACGGT AGCTAG GTTCCG GAAGCA GCCACA AGCCAA TAGCCA GTTAGG TAGCTC ATAGTG	KDVIFMGFPLGARG QWYNKNFELLDTLGL PRSRQDIIAKTLSTDA LISSVDIIHMFASRGR RQHA (SEQ ID NO: 1386)

145

R2	R2Dr	AB097 126	Danio rerio	GGCCCTACCGATGACTGTACCACCCCACTAATACAGGTAATCTTTCCGAG TGGGAAGCACCCGAACCCCAAAAGTAAGTACCCAAACCTTGTGATTGGA GAAGGAAAGAGCTTAAAGAGTGACAAAATTTGAGTCCCAAGGTCGTG GAGTCAAAAATTTAGGAATGATACAAATAGTAACGATGGATCCAATA TTATAGACGCATACCTCAGGAACTCTCACTGCCATACAACTCAGG GCCAATGTATACCCACAAGGAATTTCTCGCGGGGGAGGGGTGAT AACTATGTTAAGTTTGTAGGCATGTGAAGCGGACCTTGAACCTGTG GCCATATCATCGGCTTTGCCAGTAACGAAGGACGCCCAATCAAGAG GCACAATCGCATATGCGACAGGCTTTCGAGGAGGCGAGCTAAGAGGGA ATGGGTGCTTCAAGGAGCGCACTTGAGGGATGCCACCAAGGAACT GTTAAACCGGATGTGATATTCGTGAAAGAGGACCGTGCACTGGTTGTG GATGTACAGTACGATATGAATCAGCCAA GACAACGCTGGAGGCGAGCT GCTATGGAGAAAGTGACAAGTACAAACATCTGGAGGCGAGAAGTGAA GGAACCTCACCAACGCAAGGACGTTGTTTTATGGGGTCCCCCTTGGA GCGGAGGGGAAATTTACAAAGGGAACCTTAACTTGTAGAGACTCTTG GCCTCCCAAAACGAGGCAATTTGAGTGTGCAAAAGACTATCCACGTA CGCGCTCATGTCTGTGACATTTGCAATGTCATGTTTCCAGTAGATCTA GGAACCAAAATGCTAGGTAGTGCATTTGCACTTCTGTAACCTGCACGT GGTGTGGGATGTGGCTGGGTGTTGTTATGGGGTATATATGTGG GATATTCTGGTGGAAATGCCATTCACTGTATGCTATCTTTTAAATAA AAGACGGTAGCTAGGTCGGAAGCAGCCACCAAGCAATAGCCAGTTA GGTAGCTCATAGTGGTAGGTGACAGGAACCTTGACTCAGAACCGGT CCATTAACTAGAACGGAACCAACTTCGGACATGCACCGATTAAACCG GATTTGCCAAGGTGGACGGGCCACCTTTACTTAACCCGGAAGGGAAC ATATAGTTATGTGTTCTGTAATA (SEQ ID NO: 1541)	MYEKLWVQAGGKK KGMPSEWEALPMTV PPTNTGNLSEWEAP NPKSKYPKCDWRR KELKWTKLESQGRG VKNFRNDTISNDWIQ YYRIPHRKLLTAIQL RANVYPTREFLARGR GDNVVKFCHCEADL ETCGHIIGFCPTKDA RIKRNHRCIDRLCEEA AKREWVVFKEPHLR DATTELFKPDVIFVKE DRALVVDVTVRYESA KTTLEAAAAMEKVDKY KHLEAEVKELTNAKD VVFMGFPLGARGKF YKGNFNLETLGLPKT RQLSVAKTSLTYALM SSVDIVHMFASRSRK PNV (SEQ ID NO: 1387)			AAATCC CAGCGG GATACA GCAAGA AGGTAT CGGATC TAATAA GGTTGA GCGAGG AGAGGG TGGAGA TCCTTTG GGGGGG	AATCCCC CTACCCAA TCCCCCG TCGTGACC TCCAGGCC AGGAATCA CGAGCGTA CGACAGTG GCCATCCG GCAATGAC AATAGCGT GACTAACG ACAATGAG	MESTAKGKSWMA RRPVEGATEGSLGRV PFVTRDPKRKPEAKR TLTHGLGLRECSVLT RLIEGRRGRDHTPSG WNAQRGMPNDESS VEEPNGPIPSNPIPTG TQALPEPMADGEQG EHPGVVVTLPRLDLN CPLCGGSASTAVKVQ RHIAFRHGTVPVRF5 CESCGKTSFGCHSVL CHIPKCRGPTGEPE

GATTGAGCTGTGAATCATGTGGAAGAACTTCTCCGGGTTGCCATTCCGTC	TCAGATCC	GTCTGT	CTCTGT	KVVKCEGCSRTFGTR
CTCTGTACATTCGGAATGTGCGGACCGACAGGCGAGCCGCTGAGA	ATGACCCT	AAGTGGTTAAGTCGAGGATGCAGTAGGACGTTTGGCACAAGGAGA	CTAAGTT	RACSIHEMHVHSEIR
AAGTGGTTAAGTCGAGGATGCAGTAGGACGTTTGGCACAAGGAGA	TGGAGTG	CGCTGTAGTATACATGAGATGCAGCTTCACTCAGAAATCCGCAATAGGA	CCCCTCT	NRKRIAQDRQEKSTS
AAAGAAATTGCTCAAGACAGGCAAGAAAGGAACTCGACAGATGGA	GGTTAACC	GAGGGAGAGTGGAGTCGAAAGGCTGACGCTGGGAAGGTCCCTC	CGGGTC	TDGEGRAGVERADA
GAGGGAGAGTGGAGTCGAAAGGCTGACGCTGGGAAGGTCCCTC	TCCGCCTC	TGGGAAGGGATCCCCCTAAACGTCCAGACGTCGAGAACGCCAG	CTCCAC	GEGPSGEGIPKRP
TGGGAAGGGATCCCCCTAAACGTCCAGACGTCGAGAACGCCAG	TTTAAAAA	AGAACCCTGAGCCCCCGCGAATCCGCCGATTCTCTCGCCACAACCC	GGTGAC	RARTPREPSEPPANP
AGAACCCTGAGCCCCCGCGAATCCGCCGATTCTCTCGCCACAACCC	C (SEQ ID	GATCTCCCCAGGAGGCTCCGGACCTACTCCGGGAGGTGGCCAGT	GCTCTAC	PILSPQDLPPGGLRD
GATCTCCCCAGGAGGCTCCGGACCTACTCCGGGAGGTGGCCAGT	NO: 1143)	GGTGGGTAAAGGCAGCGAGACGAGACGGGTACGGTGATTGACAGCGT	CCCTCCC	LLREVASGWVRAAR
GGTGGGTAAAGGCAGCGAGACGAGACGGGTACGGTGATTGACAGCGT		GCTCGAGCATGTTGGATGGCAACGATCGGCTCCCTGAGCTGGTTGAC	TCCTCGC	DGGTVIDSVLAAWL
GCGGCAGCAAGGACACTGCAGGGCTTACCTGCAGGGAGGTGGCC		CGAGACCCGCAACTTTTGTGCGCCTAACCGGAGGAGAGGACGGTGG	TCGTAG	DGNDRLPELVDAAT
CGAGACCCGCAACTTTTGTGCGCCTAACCGGAGGAGAGGACGGTGG		GGCGCCGGCTCAAACTGCTGCTAAGCGCCGCTACCCACGATTGCC	AACCCA	QRTLQGLPAGRLARR
GGCGCCGGCTCAAACTGCTGCTAAGCGCCGCTACCCACGATTGCC		AAATTGCTTCCGAAAGACCCAGCCGCTAGCCGCGAAATCCTAGA	ACGGTG	PATFVAPNRRRRGRW
AAATTGCTTCCGAAAGACCCAGCCGCTAGCCGCGAAATCCTAGA		CGGCAAAAGCGAAACAAAGTTGCCCAATCAATGAGCAAGCGATTATGA	AACACG	GRRLKLLAKRRAYHD
CGGCAAAAGCGAAACAAAGTTGCCCAATCAATGAGCAAGCGATTATGA		GCACCTTCGAAACAAATGGGCAAAATCCAAGTCCATTTGGTGGGTGGGA	GTTGCG	CQIRFRKDPARLAANI
GCACCTTCGAAACAAATGGGCAAAATCCAAGTCCATTTGGTGGGTGGGA		CGATTTGGGACGGAAACAGGGCCAAACACGCCACCTCTCGGGCCA	AGGATG	LDGKSETSCPINEQAI
CGATTTGGGACGGAAACAGGGCCAAACACGCCACCTCTCGGGCCA		ATCTCAAAAGCGAGGTCCAAACTAGCTCCGAAATGCATCGAACGCT	AAGTGA	HEHFRNKWANPSF
ATCTCAAAAGCGAGGTCCAAACTAGCTCCGAAATGCATCGAACGCT		CCACACGAGCCAGACGGCTTGGGAAAGGACATTTCCAATGGG	CGTGAG	GGGGRFGTENRANN
CCACACGAGCCAGACGGCTTGGGAAAGGACATTTCCAATGGG		ATCTGAGTGTGAGACCTCACTCAGCTGTTTAAACATGTGGTGTTCACA	GGGTAA	AHLLGPIKSEVQTS
ATCTGAGTGTGAGACCTCACTCAGCTGTTTAAACATGTGGTGTTCACA		GGTGTATCCCTCTCGCTTGAAGAAAGTCGTACGGTCTCTCGCCA	GACATG	RNASNASTPGPDGV
GGTGTATCCCTCTCGCTTGAAGAAAGTCGTACGGTCTCTCGCCA		AGTCTCAGACCCAGGAGGAGATGGAGATCGGCAACTGGAGACCAA	CGTACG	GKRDISNWDPECETL
AGTCTCAGACCCAGGAGGAGATGGAGATCGGCAACTGGAGACCAA		TCACCATCGGGTCGATGGTCTTGGGCTTTTCAAGGGTGATCAATAC	TGAGCG	TQLFNMMWWFTGVIP
TCACCATCGGGTCGATGGTCTTGGGCTTTTCAAGGGTGATCAATAC		GAGATTAACGGGAAGCCTGCTCGTTGCACCAAGACAGAGAGGGTTTCG	CGCATTT	SRLKSRVLLPKSSD
GAGATTAACGGGAAGCCTGCTCGTTGCACCAAGACAGAGAGGGTTTCG		ACGAAGCCCGGGTGTTCGGAGAACCTGGAACTCGAATGTCCTCTC	TTGCTGT	PGAEMEIGNWRPITI
ACGAAGCCCGGGTGTTCGGAGAACCTGGAACTCGAATGTCCTCTC		CGACACTCAAAGAAAGCGCAGCCAACTGGCAGTGGTATTCGTCGATT	TCTCTG	GSMVLRFLTRVINTR
CGACACTCAAAGAAAGCGCAGCCAACTGGCAGTGGTATTCGTCGATT		TTGCAAGCGTTTGACACCGTCTCTCATGAACACATGCTGTCAGTCCCT	GACTGG	LTEACPLHPRQRGFR
TTGCAAGCGTTTGACACCGTCTCTCATGAACACATGCTGTCAGTCCCT		GAGCAGATGAACGTGGATCCCCACATGGTAAATCTGATCCGGGAGATT	GTTTCGT	RSPGCSLENLECLL
GAGCAGATGAACGTGGATCCCCACATGGTAAATCTGATCCGGGAGATT		ACAAACAGCTGCACAAAGTTCGAGCTAGGCCGGAAGAGGGACCAAG	CCCCCTC	RHSKEKRSQAVVVFV
ACAAACAGCTGCACAAAGTTCGAGCTAGGCCGGAAGAGGGACCAAG		ACATCCCAAGTGAAGGTTGGTGTAAAGCAAGGGGATCTCTGTCCCGCT	ACAACC	DFAQAFDTVSHHIM
ACATCCCAAGTGAAGGTTGGTGTAAAGCAAGGGGATCTCTGTCCCGCT		GCTTTTCAACCTGGCTTTGGATCCTCTCATCCAAAGTCTCGAACGCACAG	ATCACTT	LSVLEQMNVDPHMV
GCTTTTCAACCTGGCTTTGGATCCTCTCATCCAAAGTCTCGAACGCACAG		GCAAAGGGTGTGAGGCCGAAGGTCAAAAGTGACAGCTTTAGCGTTGCG	ACACTAT	NLIREYNTSCTSVELG
GCAAAGGGTGTGAGGCCGAAGGTCAAAAGTGACAGCTTTAGCGTTGCG		CGGATGACCTGGCACTGTTGCGGGCTCGTGGGAGGGAATGGCACACA	AGGGGC	RKEGPDIPVRVGVKQ
CGGATGACCTGGCACTGTTGCGGGCTCGTGGGAGGGAATGGCACACA		ACCTTGGCTTGTAGACGAAATCTGCCTAACACCGCCCTCACAGTCCAA	ACAGCG	GDPLSPLLFNALDPL
ACCTTGGCTTGTAGACGAAATCTGCCTAACACCGCCCTCACAGTCCAA		CCCAAAAAGTCCACAGTTTATGTTAGGCTGAGGCTGAGAGGTGCTTTCA	GCTCCTA	IQSLERTGKGCEAG
CCCAAAAAGTCCACAGTTTATGTTAGGCTGAGGCTGAGAGGTGCTTTCA		CAGTGAACGACTGCCCCCATGGTTCTGGGGGGCAAGGCCCTGCAGC	CCTCCCT	HKVTALAFADDLALV
CAGTGAACGACTGCCCCCATGGTTCTGGGGGGCAAGGCCCTGCAGC		TAACAAACATCGAAAACTCCATCAAAATATCTGGGAGTAAAGTCAATCC	CCCTATG	AGSWEGMAHNLAL
TAACAAACATCGAAAACTCCATCAAAATATCTGGGAGTAAAGTCAATCC			ACCCCC	VDEFCLTTGLTVQPK
			CTTCCCA	KCHSFMVRPCRGAF
			TACCGA	TVNDCPPWVLGGKA

148

149

150

151

R2	R2-1_GA	.	Gastero steus aculeat us	CATATTGGGGTCTCAGGAGGAGACACAGGGTCTGTTCGGGCTCCGGTA AACGGTACCGGAGTCGGTTAAGCATCGTTGGGCCCGCCTCCACGTGGT GGTCCGGGTAACACCAATAGGGTGGCTAAGAGGGCCAGTAATTTCCCC GAATTGCTTCCCCCGCGCGGGGGGACCCCTTTAGTGTCGAGC	CATATTGG GGTCTCAG GAGGAGA CACAGGGT	TTGTGG GTTGAG TATAGTC GGAGCA TATGGT CGGTG TCCCGTT GTTTCA CGGCGG GCAAGC GACTAT CATGAT AAAGTA GAATGG GAGACG GGCTCC CTGACA AACCCG GAAAGG CGCCCC CCCGTG GTTGCT AGCAGC TGACGG ATCAGC CTCGAA GAAAAA TGAGTG AGAGGG GACGCC GCAACC AC (SEQ ID NO: 1266)	MLRGGVGTTPAGGA GAVGPGMASPGCS VRFSPGRRLLGHRT GGLSPSVSWRLKRLS

GGTCGCGCTCCGCGTTTGGGGTGTGCGAGGCGTGAGCCCTTCGTCCCT	CTGTTGCG	CCGAAG	VSLRWSGPGILLGA
TAAGTTCAGACGGTCCCGGCTTCTGCGGGCCAAACCCCGGTGCAGCG	GCTCCGGT	GGCCCC	DGAGGGAASPRG
TTCTCCCATGTTGGATCGGCACCAAGCCCGGTGCCATGCGAGTTGAG	AAACGGTA	CCCGTTT	TQVLGSGAGRWLG
ACATTTGTTATGATCTGCTGCGTGGTTGACTTGTAAAGTCAATTCCT	CCGGAGTC	CAGACC	HSGRSPSAARGLR
CCTCACTGCGTCCCGGAGTGTGATCGGTTGAAGAGGATTCGTG	GGTTAAGC	TGATTC	RLTVRLKRLSGLLSP
TTGACCTCGGCGTGAATTTGGGATGATATACAGGTAGGTATAGAG	ATCGTTG	AGGCTA	KACRDAEEGSSSPG
GGGTGGGATGTTGCGTGGCGTGTGTAATCCCGGCTGGGGG	GGCCGCC	CCTGTG	FRNPKGLGGRLTPL
GCGGTGCGGTGGGCGAGGATGGCTCGCCGGGTGTTGCAGTGTCT	TCCACGTG	CCTAAT	GSRRFRLTVSLNRW
CGGTTAGTCCCGAGGAGGCGACTGTTGGCCACAGGACTGGAGG	GTGGTCCG	GGGGGG	RGSLVKLNASSRASG
GTTAGTCCCTCGTGTCTGGAGGCTCAAGCACTGTCTGTCTCTGA	CGGTAACA	GTCCCA	RTPVPKACDSRAGR
GGCGTGGAGGGGCTGGCTGCTAGGTGCGGATGTGCGGGGGGA	CCAATAGG	AAGAGA	GSEHAEGGGVSAAP
GGCGTGGGTGGCTCCCGAGGGGTACGAGGTCTCTGGGAAGTGGG	GTGGCTAA	TGTTGTC	MVLSRRRLTFSVDG
GCCGGCGTGGTGGCTTGGGCACGGGTGCGAGGGTCTTCTCTCTCT	GAGGCCCA	TGTTGTA	DSNSGDRARSGVSA
GCGGCCGGGGCTAAGCGGCTGACGGTACGGTTGAAGCGACTCAG	GTAATTC	GAAGGG	ARPGHLLVDGESASS
CGGTGGCCTGTTGCCCTAAGCGGTGTCGGGATGCGGAAGAAAGAA	CCCGAAT	TTTGCG	RSRPAGDARLAGPST
CTCAGCAGCCAGGGTCCGGAAATCCAAAGGCTCGGGGGAAAGGG	GTCTTCCC	CCACTG	RSRRKGCLPPVDEN
GTTGACGCTCTCGGATCCCGTAGATTTTGTGGCTGACCGTCTCCCTGA	CCCCGCG	ACTGCA	PKKRTRLMAKMTNG
ATCGTGGAGGGGCGTGGTGAAGTTGAACGCTAGTAGCAGGGCCT	GGGGGGG	CGGAAG	NPTSHVPCPAPCSNG
CCGGCCGAGGACCCCTGTAAACCCGCTGTGACTAGAGCCGGAC	ACCCCT	GGTGGG	HEGGRRVAVIEGRLP
GGGGCTCGGAGCATCGGAGGGAGGTGAGTGAGCGCTGCACCTATG	TGATGTC	CCTCGA	ELSGSRISGQIPALPV
GTGTTGCGCAGTCGGCTAAGCTCACCTCTCTGTGGATGGCAGCTA	GGAGCGG	CAGGTA	ETSFVGQSTGRGAD
ACTCGGGGATAGGGCCGAGCGGGTCCGTCTGCGAGCCCGTCTG	TCGCGCT	GGGGT	GDANANSSPPSPNL
GCCACTGTTGGTGATGTGAGAGTGGCTCTCAAGATCTGCCCCCG	CCGCGTT	ACATGA	GGSVGMVPAVRDGT
GGGGATGCCAGTTGGCGGGGCTTACCGGGAGTAGGAGGAAG	GGGGTGT	CTCCGT	PPLRPGEDHSRECA
GTTGCTTCCCGGTCGACTTTGAAACCCGAAAGCGCACACGGTT	CGCAGG	GCTGCT	GGNTPLWMLDSFR
GATGGCTAAGATGACGAATGGTAAATCTACCTCGACGTCCTTGCCCT	GTGAGCT	CAGCAG	CDYCPREFGTRAGRS
GCCCCGTCTCAATGGGCATGAAGGAGGTGGGCGAGTTCGGGTGATC	TCGTCCC	ACCCGC	LHMRRAHLAEYDGA
GAGGGGGGCTGCCGAGTTAAGCGGTAGTAGGATCTCTGGAATACAG	TTAAGTTC	GCCTCT	GFCWGERLSEFAATR
CCAGCCCTGCTTTGAAACCACTTTGTCGGCCAATCGACTGGCCGGG	AGACGGTC	GAGACC	LWSTEETKKLAVFCE
GCGGGACGGCATGCGAATGCGAATAGTAGCCCGCTTCTCTAATCT	CCGGCTC	GGGTAG	RGVPSSECRAIAASL
GGCGGCTCGTTGGGATGGTGCCTGCCGTGCGTGATGTACCCCGCC	TTGCCGG	GGCTAC	GAGKTHHQVRSKCR
GCTTGGGCTCCAGGAGGATCACTCGCGGAGTGTGCGGGGGAA	CCAAACCC	TTGAAC	LVFEAIRRELLEVA
ATACTCCCTCTGGATGCTGGAGGACAGTTTCCGGTGTGACTACTGTCT	CGGTGCAG	AAGCGA	ATERLEKSARRKQPA
AGGGAATTCGGCACAAAGAGCGGGCGCTCGTTGCAATGCGCAGGGCT	CGTTCTCC	CGCCCT	VPPAPVHGVGVLR
CACCTGGCGGATACGACGGGGCAGGTTTCTGTTGGGTTGAACGTCTC	CATGTTGG	GGTGTA	GLLGKRV/PREGGTTG
AGTGAAATCGCCGCTACGCGCTCTGGTCGACGGAGGAAACCAAAAG	ATCGGCAC	TGTCCG	STSARIVRRDDCRQG
CTGGCCGTGTTTGTGAGAGGGGTGTGCCCTCACCGTCGGAATGCAGA	CCAGCCCC	TATCCTA	AVASASLINLRRLGRK
GCCATTGAGCCTCTCTGGGGCGAGGAAAAACATCATCAGGTTAGAT	GGGTGCCA	ACCTGG	ATGRSGRRRVLGRPP

CGAAGTGTGCACTGGTGTTCAGGCCATTGCGGGCGTGAATTGCTTGA GGTGGCTGCTGCCACGGAGCGTTTGAGAAAAGCGTAGGCGGAAGC AGCCGCCGTACACCCGGACCCGTACACGGAGTGAGAGGGGTCTCTGC GGGGCTACTAGGAAGCGGTGCGAGAGAGGGTGTACACAGGC AGCACTCAGCAAGGATCGTCAGGAGAGACGACTGCCGTAGGGGGCA GTTGGTCGGCTTCTCAATCTGATCAGAAGGCTGGTCAAGGCAA CGGGCGTCCGGAGAGACGGGTCTTGACGCCACCCAGGATGG ATGAAGCGTAGCTGAGGATGAGGAGGATGCGCAGGTTCTCTATC GGTTGGCCGGCTGGCTGGGCCAAGTTGGCTATGTTTCTCTGGACG GACAGATGGGGGAGCTGCCCGTTCCACTCGTCGAAGTGTGCGCGG TCTCCGGGAGAGTGGAGCATAGTCAGAGCCTTCTGGGTCTGGTC AGTTCGGGGCTTCGGGACTGCCGACAACGACGATTTGGGAAGCTGA TCGATCCGGCTGAAGTCAGGGCCCATCTCCAGTCCATCAAGAACCGGTC TTCCCCGGGCCGGATGGCATCACCAGGTGGCGCTGTCCAAATGGGA CCCCGAAGGGATTAATTTGGGCACATGTACTCAACATGGTTGGTATCG GCAGGCATCCCTAAGTCTTCAAGAAGTGCAGGACGACACTATCCCAA AGACCGGGACGTAGTCTACATGGTGACGTGGGCAATGGAGGCCCA TAACCATTCGCTCCCTGCTGAGACTCTATTCGCGGATCCTGACGGA AAGGATGACAGTGGCTGTCTAGCCACCCGCCAGAGGGGCTTCATT GCCTCCCCGGGTGTTGGAACCTCATGCTGTTGGAAGTTGCATGA GTCTCAGCAAGGAGGAAATGGCTCCCTCGCGTGTGTTCTCGACTT TGCGAAGGCTTCGATACCGTCTCCACGACGACCTCTGAGTGTCTG GTGCAAAAAGGCTTGACCAACACATGGTGAGTTGATCAAGGACTCC TACGAGAACAGCGTGACCAAGGTGCACTGTGAGGAGGTTGTCCACT GACATGCCATGAAGTGGGAGTGAAGCAGGGTGACTCCATGTCCTCT CTCCTTTAACTGGCTGGATCCGCTTATCCAGCAACTGAACGCGA GGCCGGGGCTTCCAGTAAATGGGAAGTCCATTACTGCGATGGCATTT GCGGATGACTGGCCATAGTAGTACTCTGGGAAGGCATGAGAGCC AACCTGATATCTGTGGACTTCTGCGAGCTTACTGGAATGCGAAGCC AGCCAGTAAAGTCCAGGGTCTGATTGAGAAGAGTGGCAGCAGGT CGTACAAAAGTGAACAGGTGCGAACCGTGGCTGCTGAACGACACAGCTC TTCACATGGTCGGGCTAAGGAATCAATCAAGTACCTGGCGTCCAGGT GAACCGTGACAGGGATCTTCGTGAGGATACGGTTGCCAAACTACG ACAGTGGGTAGTTGCAATCTCAAGACGCTCTACGTCCGCTTGACAAG GTGCCCTGTTGTGCCAGTTGCCGTACCGAGGGTCACTCTCTGGGCTG ATCACTGCATGCTATCTCGAAGGCCCTGACAGAAATGGATAGGAGCAT AAGACAAAGCAGTGAAGAGGTGTTGCACCTGGCAGGTGTACCAAGAA CGGCCTCTCTACTCAAGGAAATCCAGCGGTGCTGCTGGGTATCCCAAAA	TGCGAGTT CAGACATT TTGTTTAT GTATCGTC TGCGTGGT TGACTTGC TAAGCTCA TTTCTCTCT CTCACTGC GTCCCCC AGGTGCTG ATCGGTTG AAGAGGA TTCGTCGT TGACCTCG GCGGTGA ATTGGGA TTGTATTA TACAGGTA GGTATAGA GGCGGTG CGG (SEQ ID NO: 1144)	TTTGGG AAAGCC GATACC GGCAAT GCCCC CACAGG TGTCGC GCACCC CACGGG ATGACG TATGGG CCCCGG GGGACC TCATGG ATACTCC ACTGGA CTTGAC AATCCT GGTGTA CTGGAT GCAGCG ACGTTG GTGACA TAAGCA ATCGCT AAGTCG GGGTAG GGGAGG TGGGA CCTCGG CACGGC TGTAGG AACGGG TGTATG GGCTCC GGCAGC CGTCGT CACTCCC	RMDVRRSVRMRM RRFLYRLARLGWAKL AMFVLDGQMGASC PVPLVEVSAVFRERW SIVRAFLGLGQFGGF GTADNAGFGKLDPA EVRAHQSIKNRSSP GPDGITKVALSKWDP EGIKLAHMYSTWLVS AGIPKVFKKCRITLIP KTGDVSLHGDVGGQ WRPITIASLVLRLYSRI LTERMTVACPSPRQ RGFIASPGCSENLML LEGCMISLKAGNSL AWFVDFAKAFDTVS HEHLSVLVQKGLDQ HMOVEIKDSYENSVT KVHCQEGCSTDIAI KVGVKQGDMSPLL FNALDPLIQQLEREG RGFPVNGKSIAMAF ADDLAIVSDSWEGM RANLDILVDFCELTG MRTQPSKCHGFLIEK SGRSYKVNRCPEWL LNDTALHVMVGPKEI KYLGVQVNPWTGIF AEDTVAKLRQWVVA ISKTLPLDKVSLLC QFAVPRVIFVADHC MLSAKALTEMDRSIR QAVKRWLHLARCTT NGLLYSRKSSGGLGIP KLSMIVPAMQARRLL GLRSKDETVRWMF LETTDHFVAFERAWLR AGGSPDEVPELGPDL

TTGTCGATGATTGTTCCGGCCATGCAGGCCAGGAGACTCCTGGCCTGT CCCGTTCTAAGGACGAGACGGTCAGGTGGATGTTTCTGGAGACAACTG ATCACGTGGCGTTTGAGAGGGCATGGCTGAGGGCTGGAGGGTCGCCA GATGAGGTACCGGAGCTGGTCCGATCTGGTGGAGGGCTCCCTGCG GAGGGAAACGCTGACCCTGTACGACGGTGAGGCCAAGGAAGCGCAT AGTCCGTGTACTGGCTCAAGTCAGTTCGACAGATGGGCCGGTCA ATTGGTCAGGGAAGGATTCCGACGTTTCGAAGCGGACAAGATCA GCAACTGCTGTTGTACGATACCCGCCAACAAAGCTGAAGCCTGGGG ATTTACGGCGGCTGCCAGCTTAGAGCGAAGTTTACCCGACCCGGGA GCTAGCGGTGCGGAAGGACCGATACGATAGATGCTGTTGTCGACA CTGTGGGAGGCCCGACAGACTTGTGGCACATCCTTGCCTCTGCCG AAGGTTAAGCGGTGCCGTATTCAAGAGGCCACCAAGGTGTGCCAGGTC CTCGTCGGGAGGCTGAGCGCCATGGATGGGAAGTGGAAGGGAAAA GCGCTGGATGCTGCCCTCCGGGAGTGTGTCGCCCGGACCTGATCTG CTGGTTGATGAGTGGCGCTATTGTCGATGTACGGTGAGGTACGA GTTGATGAGGAGTCGTAGAACGCGCGCGAATCGAGAAGGAATGCAA GTACCGCCCTCTCATTCCAGTGATCAGGGCGAGCAGAGTTCAGACGAAG AAGGTGACGGTCTATGGCTTCCCTCTGGAGCCAGGGGAAAGTGGCCT GCTAAGAACGAGCTGCTGCCGACCTCGGCCCTGAGCAAGGCTCGG ACTCGAGTTTGTAAACTCTGAGCCGACAGTCTCTTACATCTCT GGATGTTATGAGGACGTTTATGCGTTAAGGAGGGGAGTAGGTCTCTAC TCTGACCCGAAGGGCCCCCGTTTCAGACCTGATTCTAGGCTACCTGTG CCTAATTGGGGGGTCCAAAGAGATGTTGCTGTTGTAGAGGGTTTG CGCACTGACTGCACGGAAGGTGGCCTCGACAGGTAGGGTTACAT GACTCCGTGCTCAGCAGACCCGCGCTCTGAGACCGGGTAGGGCT ACTTGAAACGACGACCCCTGGTGATGTCGCTATCCTAACCTGGTTG GGAAGCCGATACCGCAATGCCGCCACAGGTGTCGCGCACCCACG GGATGACGTATGGCCCCGGGGACCTCATGGATACTCCACTGGACTT GCACAACTCCTGTTACTGATGACGACGACGTTGGTGACATAAGCAATC GCTAAGTCGGGTAGGGAGGTGGGACCTCGGCACGGCTGTAGGAA CGGGTGTATGGGCTCCGGACGCGTCTCACTCCCATACAACACAGGG GCTGCATCCTGGTGGCGGTGCTAGTTGGTTCTGGAAAGCCCGCCGGGC TGTTTCGAGAAAGCAGGGTGCGCCAAGGTTTGGTATATCTGGG TCCGGTGCATACCTATCGATGGCAGCGAGGGCCGCTCGTGACGCG CTGTGTGGAGCTGGAGCCGCGCTGGGTATGAACAGTTCTTCCGGATGT GGCGTAGCTAGATAGTACCCGTGGTTGTGGCGTGGTGTGACCAAAAT GTTGTCCTGTGTGCATAGGCCAAGGTTACGTGGGTGGCAGTCAGA AGCACCCGACCTGGAAGTGATTGCCCCGGGATCCCGGCTCTCTGTGAA	ATACAA CACAGG GGCTGC ATCCTG GTGGCC GGTGT AGTTGG TTCTGG AAGCCC GCCCGG GCTGGT TCGCAG AAGCAG GGTGCG CCCAGG GTAGGT TTGGTAT ATCTGG GTCCCG TGCGAT ACCTATC GATGGG CAGCGA GGGCCG CCTCGT GACGCG CTGTGT GGAGCT GGAGCC GGCCTG GGTATG AACAGT TCTTGC GGATGT GGCGTA GCTAGA TAGTAC CCGTGG	VEGSPAEGNADPVST VRPRKRVPCDWRQ VEFDRWAGQLVQKG GIRTFEADKISNCWLY DYPPNKLKPGDFTAA VQLRANVYPTRELAG RGRTDIDVCCRHC EAPETCWHILALCPK VKRCRIQRHHKVCQV LVAEERHGWEEVER EKRWMLPSGECVAP DLICWLDELALIVDVT VRYEFDEESLERARIE KECKYRPLIPVIRASR VQTKKVTYVGFPLGA RGKWPAKNELLADL GLSKARTRSFAKLLSR RVLLHSLDVMRTFM R (SEQ ID NO: 1389)

156

R2	R2_B M	AB076 841	Bombyx mori	GGGCGATACGCATAATTTTAAATTTCCCGATTGAAATCCAGTCGTCTTAAT CTGGTGACCAAGTGGCGGGTCACCAATATAGTGCACAGGACGTGAATG	GGGCGAT ACGCATAA	AGAGCA TGTTCA AAAGCC CGGGGA ACGTTCC GGGGGC CTGCTT GGGTCG TTGGAC CCACATC CGTAAA ACGATG GATCTC GCGTCG GCGCTC GGGAGA ACTTCCC GCATGA ACGCTG ATTGCAT GTGAGA ACGCC CCACGG CGGCGG GGCAGG CGTCCC CCTGGG TGTAAG GCTCGG GGGGGT CACGGC TCCGCTC TAAAAG (SEQ ID NO: 1267)	MMASTALSLMGRC NPDGCTRGKHVTAA

GCTCCGAGGCTGGCGGAGTCACTCACTATAAGTGTGAGAGACGATGTC	TTTTAATT	AGTCCA	PMDGPRGPSSLAGT
CTGTGCCAAGTATACGTCCAAACCTAACGGGTTAAGTGAAATAGTTGC	CCCGATTG	GCGGTA	FGWGLAIPAGEPCG
TCATAACAGGGACGGTGTACCTGTTTCTCTGTGGCTGGCTATCGAATGG	AAATCCAG	AGGGTG	RVCSPATVGFPPVAK
ACGGGACCAATACACCCCTGTTAGTAATGGGTAAGAGAGAGCGGT	TCGTCTTA	TAGATC	KSNKENRPEASGLPL
CTGAAACTATGGCCGAATACAGACGCCCACTCTACCCATAACCTGCA	ATCTGGTG	AGGCC	ESERTGDNPTVRGSA
CGTGTAACCGCCGACATTGACCGATACGGGAGGAGGGGCGACCTTG	ACCAATGG	GTCTGTT	GADPVGQDAPGWT
AATCACGTAGTCTTGGTGTAGCCATTGCGGGACTACAGCCCTCGTAAGT	CGCGTGCA	TCTTCCC	CQFCERTFSTNRGLG
GCCGCTTAGAACGCAACGGGGCAATAGGTGGGCGGGGCGCTAGCG	CCAGTATA	CGGAGC	VHKRRHPVETNTD
GGGGGAGTAATCTCCCTGTTGGCGTGACCGCACTGTCCCACTGGG	GTGCACAG	TCGCTCC	AAPMMVKKRWHGE
GGCAGTGTATCCGGAACAGGTGGGCGGGGCGGCCACAGGGGGGA	GACGTGAA	CTTGGC	EIDLLARTEARLLAER
GCAATCCCTCTGTATGATGGCGAGCACCGCACTGTCCCTTATGGACGG	TGGCTCCG	TTCCCTT	GQCSGGDLFGALPGF
TGTAACCCGGATGGCTGTACACGTGTTAAACACGTGACAGCAGCCCCG	AGGCTGGC	ATATTTA	GRTLEAIKQRRREP
ATGGACGGACCGGAGGACCGTCAAGCCTAGCAGGTACCTTCGCGGTGG	GGAGTCAC	ACATCA	YRALVQAHLARFGSQ
GGCCTTGCATACCTCGGGCGAACCCCTGTGTCGGGTTTGACGCCCG	TCACTATA	GAAACA	PGPSSGGCSAEPDFR
GCCACAGTGGGTTTTTCTGTTGCAAAAAAGTCAAAATAAGAAAAATA	AGTGTGAG	GACATT	RASGAEEAVEERCAE
GACCTGAAGCCTCTGGCCTCCGCTGGAGTCAAGAGGACAGGACGGGATA	AGACGATG	AAACAT	DAAAYDPSAVGQMS
ACCGACTGTGGGGTTCCGCCGCGCAGATCTGTGGGTGAGGATG	TCCTGTGC	CTACTG	PDAARVLSELLEGAG
CGCCTGGTTGACCTGCCAGTTCTGCGAACGAACTTTTCGACCAACAG	CAAGTATA	ATCCAAT	RRRACRAMRPKTAG
GGGTTTGGGTGTCACAAAGCGTAGAGCCCACTCTTTGAGACCAATACG	CGTCCAAC	TTCCGC	RRNDLHDDRTASAH
GATGCCCTCCGATGATGTTGAAGCGCGGTGGCATGGCAGGAAATC	CCTAACGG	GGCGTA	KTSRQKRAEYARVQ
GACCTCTCGCTCGACCGAGGCCAGTTGCTGCTGAGCGGGTCACT	GTTAAGTG	CGGCCA	ELYKKRSRAAAEVID
GCTCGGTGAGACCTCTTTGGCGCTTCCAGGGTTTGAAGAATCT	AAATTAGT	CGATCG	GACGGVGHSLME
GGAAGCATTAAAGGACAAACGGCGGAGGAGCCCTTTCGGGCATTGG	TGCTCATA	GGAGGG	TYWRPILERSVDAPG
TGCAAGCGCACCTTGCCGATTGGTTCCAGCGGGTCCCTCGTCGGG	ACAGGA	TGGGAA	PTPEALHALGRAEW
GGGTGCTCGCCGAGCCTGACTTCGGCGGGCTCTGGAGCTGAGGA	CGGTGTAC	TCTCGG	HGGRDYTLWKPI
AGCGGTGAGGAACGATGCGCCGAAGACGCCGCTGCTATGATCCATC	CTGTTTGC	GGATCT	SVEEIKASRFDWRTS
CGCAGTCGGTCAGATGTCGCCGATGCCGCTCGGGTTCTCTCCGAATC	TCGTGGCT	TCCGATC	PGPDGIRSGQWRVAV
CTTGAGGGTGGGGGAGAGACGAGCGTGCAGGGCTATGAGACCCAA	GGCTATCG	CTAATCC	PVHLKAEMFNAMM
GACTGAGGGCGCGAAACGATTGTCACGATGATCGGACAGCTAGTGC	AATGGACG	ATGATG	ARGEIPEILRQCRTVF
CCACAAAAACAGTAGACAAAAAGCGAGGCGAGATCGCGCTGTGCA	GGACCAAT	ATTACG	VPKVERPGPGGEYRP
GGAACGTGTAAGAAGTGTGCGAGCAGAGCAGCAGCTGAGGTGATCGA	ACACCCCC	ACCTGA	ILIASIPLRHFHSILARR
TGGCGGTGTGGGGTGTGCGACACTCGCTCGAGGAGATGGAGACCTA	CTGTTAGT	GTCACT	LLACPPDARQRFI
TTGGGACCTATCTCGAGAGAGTGTCCGATGCACTGGGCCTACACCG	AATGGGGT	AAAGAC	CADGTLENSAVLDAV
GAAGCTCTTACGCCCTAGGGCGTGGGAGTGGCAGCGGGGCAATCGC	AAGAGAG	GATGGC	LGDSRKKLRECHVAV
GACTACACCCAGCTGTGGAAGCCGATCTCGGTGGAAGAGATCAAGGCC	AGCGGTCT	ATGATG	LDFAKAFDTSHEAL
TCCCGCTTTGACTGGCGAACTTCGCCGGGCCCGGACGGTATACGTTCCG	GAAACTAT	ATCCGG	VELLRGMPEQFCG
GTCAAGTGGCGTGGGTTCTGTGCACTTGAAGGCGGAAATGTTCAATGC	GGCCGAA	CGATGA	YIAHLYDTASTTLAVN
ATGGATGGCACGAGGCGAAATACCCGAAATTTACGGCAGTGCCGAAC	ATCACGAC	AAA	NEMSSPVKVGRGVR

CGTCTTTGTACCTAAGGTGAGAGACCAGGTGGACCGGGGGAATATCG ACCGATCTTGATCGCGTCGATTCCCTGAGACACTTTCACCTCATCTGG CCCGAGGCTGTGGCTTGCTGCCCCCTGATGACGACAGCGCGGATT TATCTGCCCGACGGTACGCTGGAGAAATCCGCACTACTGGACGCGGT GCTTGGGATAGCAGGAAGAGCTGCGGGAATGTACGCTGGCGGTGC TAGACTTCGCCAAGGATTGACACAGTGTCTCAGGAGCACTTGTCTGA ATTGCTGAGTTGAGGGGATGCCGAAACAGTTCTGGGGTACATTGCT CACTTACGATACGGCTCCACCACCTTAGCCGTGAACAATGAAATGA GCAGCCTGTGAAAGTGGGACGAGGGTTCGTAAGGGGACCCCTCTGT CGCGATACTCTTCAACGTGTGATGGACCTCATCTAGCTTCCCTGCCG GAGAGGTCGGGTATAGTTGGAGATGGAACCTTGTGTCGCTCTGGCC TATGCTGACGACCTAGTCTGCTTGGGGGTGGAAGGTAGGGATGCAG GAGTCCATCTCTGCTGTGACTGTGTTGTTAAGCAGATGGCCTACGCC TGAATTGCAGGAAGAGCGCGTCTGCTATGATACCGGATGGCCACC GCAAGAGCATCACTACCTGACTGAGCGAACCTTCAATATTGGAGGTAA GCCGCTCAGGCAGGTGAGTTGTTGAGCGGTGGCGATATCTTGGTGT CGATTTGAGGCTCTGGATCGTGACATTAGAGCATAGTATCAGTAGT GCTCTGAATAACATCTCAAGGACCTCTCAACCCCAACAGAGGTTGG AGATTTGAGAGCTCATCTGATCCAGATTCAGCACCGTTTGTGCTT GGAACATCTCGGATGACCGATTGAGAAATGCTCGATGTCAAATCCGA AAGCAGTCGACAGTGCTAAGGTACCGCGGATGTGCCAAGGCAT ACTATCACGCCGAGTTCAGGACGCGGCTTAGCGATCCCATCGGTGCG AGCGACCATCCGGACCTATTGTGAGGCGTTTCGGGGGCTCGACTCG TCACCATGGTCAGTGGCAAGAGCCCGCCCAATCTGATAAGATTCGTA AGAACTGCGGTGGCTGGAACAGCTCCGAGGTCAGCCGTGTG ACTCCACAACGCAACCATCTGTGCGCTTGTTCGGGAGAACATCT GCAGCATCTGTTGATGAGCGCAACTTCGGAATCCACAGCACCCCG ACATCCAAAAGTGGATTAGGAGCGATGCGCGAGATAACCGGACGG GACTTCGTGCACTGTCGACATCATCAACGCCCTCCCATCCCGCAT TCGCGGATCGAGAGGCGTAGAGGTGGGGTGAGTCTTCGTTGACCTG CCGTGCTGTTGCAAGTTAGGAGACGACGGCTCACATCTACAACA GTGTACAGAAACACGCGCGCGGATTCTACGACACAACAAGATTGTA TCTTCGTGGCGAAAGCATGGAAGAGAAAGTGACGGTTGAGCTG GAGCCGAGGCTACGAACATCGGTTGTTGCTCCGTAAGCCGATATTATCG CCTCCAGGGATGGTTCGGAGTGATCGTGGACGTGACGGTGGTCTCGG GCCAGCGATCGCTTACGAGCTTACCCGTGAGAAACGTAATAAATACGG GAATCAGGGGAGCTGTTGAGTTGGTCGAGGTAGACTAGGACTTCC GAAAGCTGAGTGGTGGAGCCACTTCGTGCACGATATCTTGGAGGGG	GCCCCACT CCTACCCA TAACCTGC ACGTGGTA CCGCCGA CATTGACC GATACGG GAGGAGG GGCAGCAC TTGAATCA CGTAGTCT TGGTGTAG CCATTGCG GGACTACA GCCCTCGT AAGTGCCG CCTAGAA CGCAACGG GGCAATAG GTGGCC GGGGCGC TAGCGG GGGAGT AATCTCCC CTGTTGGC GTGCACCG CACTGCTC CCACTGGG GGCAGTGT CATCCGA AACAGGTG GGCCGG GCGCACCC AGGGGG AGCAATCC CTCCTG (SEQ ID NO: 1145)	(SEQ ID NO: 1268)	QGDPLSPILFNVVMD LILASLPERVGYRLEM ELVSALAYADDLVLLA GSKVGMQESISAVDC VGKQMGRLNCRKS AVLSMIPDGHKKH HYLTERTFNIGGKPLR QVSCVERWRYLGVD FEASGCVTLEHSISA LNNISRAPLKPQRL EILRAHLIPRFQHGfV LGNISDDRLRMLDVQ IRKAVGQWLRPAD VPKAYVHAAVQDGG LAIPSVRATIPDLVRR FGGLDSSPWSVARA AAKSDKIRKKLRWA WKQLRRFSRVDSTT QRPSVRLFWREHLH ASVDGRELRESTRP TSTKWIRERCAQITG RDFVQFVHTHINALP SRIGSRGRRGGGES SLTCRAGCKVRETTA HILQQCHRTHGGRIL RHNKIVSFVAKAMEE NKWTVLEPLRTSV GLRKPDIASRDGVG VIVDVQVVSQQRSLD ELHREKRNKYGNHG ELVELVAGRLGLPKAE CVRATCTISWRGV WSLTSYKELRSIIGLRE PTLQIVPILALRGSHM NWTFRNQMTSMVG GGVG (SEQ ID NO: 1390)

R2	R8Hm-A			AGTATGGAGCCTGACTTCTTATTAAGGAGTTAAGGTCCATAATCGGGCTT CGGGAACCGACACTACAAATCGTTCCGATCTGCGGTTGAGAGGTTAC ACATGAACCTGGACAGGTTCAATCAGATGACGTCGTCATGGGGGCG GCGTTGGTTGAGCCTTGACAGTAGTCCAGCGTAAGGGTGTAGATCA GGCCGCTGTTTCTCCCGGAGCTCGCTCCCTGGCTCCCTTATATTT AACATCAGAAACAGACATTAAACATCTACTGATCCAATTCGCCGCGT ACGCCACGATCGGAGGTTGGAATCTCGGGGATCTCCGATCCTAA TCCATGATGATTACGACCTGAGTCACTAAAGACGATGGCATGATGCC GGCGATGAAAA (SEQ ID NO: 1544)	TTCAAGTG GATGAAGC TGGGAAG GTAATCTG TAGTTGGT TGAGTTGG TTGCAGAT TACTGCTG TCGATTTT GCTTCTA TTGAAAGC CTGTCTCT ACGGGTCC TGAAGCTT GAATTTG GTAGCTAT AGTTTGT GGGAGGA AAGTGA ATTTGTA CCATCTTTT GTCTCTCG TATCTACT ATAGTAAA TCCGGTCA TGCAGCCT CTACGCGG CGCAACTA GAAACTTG	TAAATG CCAAAA GTTGCTT GGGCTA AATGAT ACGTAC GCTAGA AAAAAGC GACTTG CTGCAC GGATGA CGGTC ATCAGA GCCCCA TATGTG CATGTC AAGGCG GCAGGG AGAATC ACTAGT GTAGCT GTTCTTT CCATTAC GACTTA CGCGGT TAACGT GGCAGG ATAGAT TTACACC	MINLIVTSSIKESDVP SSGKGGVAVNNITAG ASGKDTCVIHPGTD GIWCCTECVIEHNSG KDLKRHLAKRHPSTI SGYKCNLCPFVSERQ LSVGTHLRYCRGVKE VVKREFACASCFSDD TFSGLQVHMQRKHI AEWNDQLKEKTEFA WTDRELRLAEKELT TPSFRYNKIFYAALGT SRTYDAVRKIRYNDR YKSAIEMRSQIADA AAAAQERDVERGLV SAHSDRGKEMLPV ETKSDIQVNDIKKDI ELTPNSRQKQTNLAL ARPAVIEVEDLGRQ DVQYLASLRQDDYT SPAERSIFAYCREETN WSATKRQVLKISRTT RGLRQPKKVRPFEP EGFKPNRNMKWR KYRFLQECYREKRAET VSKILDGTFIDEPEEI RPELEEVQRMVIDRL EKRTQLDITTKIVQTD EVFCLOSYGRITGEV
				TTCAAGTGGATGAAGCTGGGAAGTAACTCTGTAGTTGTTGAGTTGGTT GCAGATTACTGCTGCGATTTTGCTTCTATTGAAAGCCTGTCTCTACGG GTCTGAAGCTTGAATTTTGGTAGCTATAGTTTGTGGAGGAAAGTGG AATTTGTACCATCTTTTGTCTCTGATCTACTATAGTAAATCCGGTCAT GCAGCCTCTACGCGGCGCAACTAGAAAACCTGGATCAGTGATCAAGGCTA ATGCATGCCGGTCTCCTCAGATTAGGAGTATAATACAAATCTGACTTC ATCACTAAGAGGCTATGGGGCTAACGATCCTATAGTCTCGATGAACCTA TTGATTGTTACTAGTAGCATAAAAGAAAGTGACGTACCTCTAGTGGAA AGGGGGGTAGCAGTCAATAACATAACAGCAGGAGTAGTGGAAAA GATACGTGCGTATCATACACCCAGGTACCGATGTTTGGTGTGTA CTGAGTGTAGAGATACATAACAGCGGTAAAGATCTGAAACGACATCT TGCAAAACGTCACCCGAGTGAACGATAAGCGGTTACAAATGCAATCTG TGTCATTTGTTAGTGAACGCCAACTAAGTGTGGGACACATCTGAGGT ACTGCAGAGGCGTAAAGAGTGGTTAAAGAGAGTTTGCATGCGCGA GCTGCTTTTTCTCGGATACGTTCTCAGGACTTCAGGTGCATATGCAA AGAAAGCATATAGCAGAAATGGAACGACCAAGCTGAAGGAGAAACGGA GTTTGTGGACAGACCGAGAAATGAGGGAGCTTGCTGAGAGGAAC TACCACTCCTTCTCAGGTACAAACAAATTTTCTATGCTGCGTAGGTA CCTCCGGACCTACGACGCTGTGAGGAAAATTCGCTATAATGACAGATA CAAATCTGCCATTGCTGAAATGCGATCAGATAGCAGATGCGGCTGCC GCTGCACAAAGAGAGGGATGTAGAGCGGGGTTTGTTCAGCACACTCA GACAGAGGAAAAAGAAATGCTCCCTGTTGTTGAAACCAAAAGTGATATCC AAGTAAACCAACGATATCAAAAAGGATATTGAATTAACACCCGAATTCAG ACAGAAACAAACTAATCTAGCGCTGGCAAGGCCAGCTGTAATTGAGGT GGAGGAAGACTTGGTAGGAGGATGTGAAACAAATATCTCGCATCCCT GCGCCAAAGACGACTACAAAGTCCGGCGAGCGGTCAATCTTTGCATAC TGCAAGGAGGAAACCAATTGGTCTGCGCAAAAAGACAGGTATTAAAG ATATCGAGAACTACAGAGGTTTAAAGACAACCTAAGAGGTTCTGCCAT TTGAGTTCCGGAGGGTTCAAAACCTAACAGAAATATGAGAAAGTGGGA			

GAAAGTAGATTCTTCAGGAATGCTATAGGGAAAAAGAGAGCTGAGA CTGTTAGCAAGATCTGGACGGGACTTTTATCGATGAACCGGAGGAAG AGATTAGACCAGAGTTAGAGGAAGTACAACTGATGTACATTGACCGGCT GGAGAAAAGAACTCAGCTGGATACCAAGAAATTGTCAAAACAGACGA GGTGTCTGTCAAAAGTACGGTCGATTACGATCGGGGAAGTAAGA GATGCACTGGTGCAAGCAAGAGGACTCGGCTCGGTCCTGACGGC CTGCTTACAGGATGTAGAGAGGCTGGACCACTATTATTGTAAACA TCTTTAACATGTGGTACTTACATGGGATCCCTGTGGAAGAAAACAGGTG TCGAACAATACTCTTATACAAGAGTGGCATAGACATCTGGCATCAAA TATAGACCTGTGACAATCGGCAACATGCTGAACAGGCTTTACGCCAAA TCTGGGACAAACGGATCCGGAAGAACGTCGCTCTTCAATGTGAGGCAAA AAGCATTATCCCGTGGATGGTGCTTTGAGAACGTAAAAACGATCCA ATGCGTTCTCCAGTCTACAGAAAGCGTAAGTTGGAACACAAACGTCGTA TTTATTGATCTTGCCAAAGGCTTTGACACGGTCTTGCATGACTCGATAAG GAAAGCGTTGTGGCGGAAAGGTGTTCCGCTCGGGTGTGTTAAAGTGT AGACAGCTTATATGCGGGAGCTGTCAAGACATAAGTTGGAAGAAC GAAAACCTGTTCTATATGATAAACTCTGGAGTCAAGCAGGGTTGCTCT CTGTCACCTCTTCTAATCAACCTAATACTGGATGAACAGCGAGAGAT AGAGGCAACCGCTCGGGTATAGATCTTGTGTCACGTTCTATCATCT ATGGCCTTTGTGACGACTACGTGTTGCTAGCGAAGGACTCCGTGGAGA TGAACGAGTTGATAAGAGTGTAGTACATCTTCAAGAGAGAAAGGCTT ATCTGTAACCCAGGTAATGCAATCGTAAGAGTCTTCCCGTAAAG GAGAAGAAACGGTCAATGAAGTCTTGTAGACCTCATAGATGGTG AGGATAAAAGACCAGGATGTTGACATCCCATCTATGACATATGACAGCT TAGGAAAATACCTTGGTTCGATTGACCCCACTGGTAAGATAGCGCT TCCGATTGAGGAGTGAAGAAATGGATGACCAAGCTAAAAGAGTGTAA GCTCAAGCCCGAGCAGAAAGTTAAATCTGAAAGAAAGTGGTTGCTCT CGGTAAACTACGTTTTCGGATGTCAGAGTGGGATCAGCGAACTTC GGAGTTGACACGATTTGTAAGGAATGGCGAAGAAACATCATTCACCT ACCCATGTTGACGATGAGTGGATACACTCGATCAAGAGGTTAGG CATTCCGACGTTTCAAGGGAAATTGTCAACAACGTATGAGGGCTTCG GAGAAAATGTCTACGTTGAAGACGGTATAGTCCGCTGGTTGGTGCA CGACTTGTTCAGAAAGACAGAGTCTTGTGGGAAAAGCGCGTTTCGAA GGTATCGAACTGAAGGCGAGCCAGGAGGCACTGCGAAGTGGAGAGACT CAACAACATTGGTAACATTACCAACGGCTTGCACCTCAAACTATCGCA GCAGTCTCTCGTAAATCGGACTGATGATTGAAGACAACTTGAAT CCGGGAACAAAGATTCTCGTTTGGAAAGCAATGGCGGGTGCCATTCCAAC AAAGATTAACTTTCGCGGGCGTAGCAGACCAGACCCCTCAAAAAATGT	GATCAGTG ATCAAGGC TAATGCAT GCCGGTC TCCTCAGA TTAGAGT ATAATACA AATCTGAC TTCATCAC TAAGAGGC TATGGGC TAACGATC CTATAGTC TCG (SEQ ID NO: 1146)	AGGAAA TAATAC GTGAAG GGTCC ACCATAT ACTGGA GTTTAG ATCTATG AGGAA ACATTTG TAATAA GTCAGT CTGGTA ACCTGG CGCCGC TGTTGA GTCAAA TTAACTA TGTCAT ACTCAT AAGTTA TCGACTT TGATAT GGCATG GGTGA TCCGC GTTATAT CAAAGT CAAACA TGATGA TTGCAAT GAGAAA CTACCAC GCTTGG TCACGTT TGTGAG GAGAAC ATCTCAT	RDALGASKDSASGP DGLLLQDVRRLGPLL CNIFNMWYLHGIPIVE ENRCRTILLYKSGDRH LASNYRPVTIGNMLN RLYAKIWDKRIKRV RLHVRQKAFIPVDGC FENVKTIQCVLQSYR KRKLEHNVVFDLAK AFDTVLHDSIRKALW RKGVP SGVVKVVDLSL YAGAVTSISVGKTKTR SICINSGVKQGCPLSP LLFNLIIDELAERIEAT GCGLDLDGHLVLSM AFADDYVLLAKDSVE MNELIRVCSTFKEK GLSVNPGKCSLRVL PVKEKRSMLVLRP HRWWRKIDQDQDIP SMTYDSLKGVLGSI DPTGKIALPIEWN WMTKLKECKLPEQ KVKILKEVVCSRNVY LRMSECGISELRSWT RFVRNWAKNIHLPT WCSSDWIHSIKGLGI PDVSKGIVIQMRAS EKMSTSEDGIVRVVG ARLVQKNRVLWEKA GFEGIELKAARRHCE VERLNNIGNITNGVA LKTIAAVSSVNRVW MIEDNLKSGNKILVW KAMAGAIPTKINLSR GVADQTLKKRRCGL TAETDGHILAGCHTS SDAYSKRHNMLCDKL

R2	R8Hm-B	.	Hydra vulgaris	CGTCGATCGGGTTAAACGCGGAAACGGATGGACACATCTTGGCTGGA TGCCATACTAGCAGCGCGCTACTCAAAACGTCAACATCTGCTGTG ATAAACTGCCAAAGAGCTCAAACTCAATGGTGGACCAACAGACGTGT GTGGCGAGAGGAGCTGCTTCACTAGTACAGGAGGCGATATAGACC TGACATTATCGTTAAAGATGACAGTAAATCACAGTCATCGATGACTT GTCGTATGAGAAATCAGAAAGGACACCTGATCCAATGTGAAAGTGGGA AAGTAAATAATACGAGCCACTCAAGCTAGATAAGTATGGACTCGAGA ACTCGAGGAGCAAAATGGTATTGCTGAAAGGTAGAGCTGATGGG ATTGGCAATAGGGCGATCGGCACAATCATCGTAGTACCTTCGGAA ACTCTGTAGTTAAAGTCGGCAGGATCGTAAGACGCTACAAATGATT GCTTGTAAATAAGCGCCAAATTAATAAGGTCACTGTCAAGGGCGGA CTCGGAGGAAATTTGCGGTGATAAATGCCAAAAGTTGCTTGGGCTAAATG ATACGTACGCTAGAAAAAGCGACTTGCTGCACGGATGACGGTTTCATCAG AGCCGATATGTGCATGTCAAGGCGGCGAGGAGAAATCACTAGTGTAGC TGTTCTTTCCATTACGACTTACGGGTTAACGTGGCACGATAGATTACA CCAGGAAATAATACGTGAAGGGTTCCACCATACTAGGAGTTTAGATCT ATGAGGGAACATTTGTAATAAGTCAGTCTGGTAACCTGGCGCGCTGT TGAGTCAAAATTAACATGTCAATCACTCAATTAAGTTATCGACTTTGATATG GCATGGGGTGATTCGCGTTATATCAAAGTCAAAACATGATGATTGCAAT GAGAACTACCACGCTTGTCACGTTGTGAGGAGAAACATCTCATTCAA GCCTCCCGGATGTCGGCACCCGCTGACATCTTCTGGCTTATGAAAATTTT CATTAAATTTTGTAAAGTCATGGCGGCTTGAAAGC (SEQ ID NO: 1545)	AKELKNGGPNRRV WRERTCFSTGRRYR PDIIVKDDSKITVIDM TCPYEKSEGHLIQCES AKVTKEPLKLDKYW TRELEGANGIVAEKV ELMGLAIGAIGTIMR STLRKLCCLKSGRIVR RLQMIACNNSAQIIK GHLSRATRRLR (SEQ ID NO: 1391)	TCAAGC CTCCG GATGC GGCACC CGCTGA CATCTTC TGGCTT ATGAAA ATTTTCA TTAATTT TTGTAA GTCATG GGCGGC TTGAAA GC (SEQ ID NO: 1269)		MSNRITIGDVPVSGK GGLTVNKQTAGADG AEACVVIHPGAKGIW SSPACLRKFTIGKELR AHLAQIHLAPSAVR YRCNKCPYEGDVQLS VGTHLRYCKGIAGVV EEKQFACAICNFSSD TFSGLQVHKQRKHV VEWNEQLKEKTEFA WTDRELRELAVKEVT IPFSVNTETFAVLDI TTRTKDAVRKIRYDTR YKSILAEVRAQVNAV AEEAPQASDESQITLL VNTGRGAELQPAVIN
				CTTGGGGT CACTGACA CATTTTC GGTAGCCA TAGTTTTT GAGAGGA AGAGTGG AAGTTTTT CCATGAGT CGTCTCTC GTATAAAC TGTGGTAA ATCCGGCC ATCCAGCC TCTACGCG GCGCAACT	ATGCC GAGTA GTTGG ATAATG ATGCAC AAGCTC GTAAGG CGACTT GCTGCA CGTATG CCGCTA AAGCT TAGCTC GATGAG TGCATG TCAAGA			

GGAGACCTTTGCTGTCTAGATATTACGACGCGGACTAAGGATGCTGTG AGGAAATTCGCTACCGGATAGATACAAATCTATCTGGCTGAAGTAC GCGCAAGTTAACGCTGTGGCGGAGGAGCGCCGCAAGCTAGTATG AGAGTCAAAATAACGCTCTTAGTTAACACAGGCGGAGGAGCAGAAATAC AACCTGTGTGATTAATAAATACTGATTCAATTGAATTAGTACTGATGC AATGAGGTTGAAATGGTAACATCGAATTCAACCAATGAAGAACAGCCTA TCAACGCGCGGTGGAACCGGCTGTAATTGAGCGGACTTGGGAAGAC AGGATGCGAAACTATATCTCGCATCGCTGCTGCTCAAAAGCATTGCACAA CGCATCTGATCGATGGACCTTGCATATTGCAGGGGAGAAAGTTGATTGG TGTAAGACGAAAAGCAGGCTTTCAAAGTATCAAGACATGCCCGGGTT TAAGACAACTCAAAGGGTGGAATTTGGAGTTTCCAGAGGGATTCA GACCTAACAGGAACCTTCGTAATGGAGGAAGTATTCATCTTGCAAAG TTGCTATAGAACGAAGAAGAACTGTTAGTAAGATTCTTGATGGT ACTTCAAGGACACACCTGAGGAAGAGATTAGGCCAGAGTTGGAGAA GTACAACGTGTGACGTTGACCGGCTAGAGGTAAGAACTCAGCTGGAT ACCACTAGGACAGTGATATAGACGAAAGATTTCGATTAGTAAGCTATG GTCGATTACGATCAGGGAGGTACAAGACGCAATCAGCGCAAGCAAGA AGGATGCTCAGGGGTCGACGGCTTCTCTCAGGACGTGACGATCGGCA AGCGAGCCACGCCAATTGTGATCATCTTTAATATGTTGCTACTTGCTAT GGAATCCCTGTAGTGAAATAGTGCCGGAACAATACTTGCATAAGG GTGGCGAAGCATCTAACGTGCACTACCGACCTGTGACGATCGGCA ATATGCTGAATAGGTATACGTAAGATCTGGGACAGACGGATCAGAA AAAACCTGCAACTTCATGTGACAGAGAAAGCATTCGTCGCTGGATGG GTGCTTGAGAAATGTAACCAATCCAATGCTATTCGATCTCCAGTCTTACAGAA GGAGCAGCGGGAACACAATGTCGATTTGTCGATCTTGCAAAAGCGTT TGATACGATTTGATGATTGATAAAGTGTAGAGAAAGCATTCGAGGAAAGGC ATACCGGAAGTGTGATAAAGTGTAGAGAAAGCATTCGAGGAGCT GTCAGAGCATACGTTGGGAAACAAAGACTCGACCTATATGTATAA ATTGAGGGGTGAAGCAGGTTGTCCTCTATCTCTTGTGTTCAATCTA GTAATAGATGAAGTACGAGGAGGCTGGAGGCAACTGGCTGCGGCTCT GATCTGGAAGGTCAGCTATTCTTCATGGCTTTTGTGATGACTACGT GTTGTGGCGAAAGACTCGGTTGAAATGAACGTGCTAATGAACGTTG CAATACGTTCTTTGAGGAGAGGGTTTAGCTGTAATCCAGCAAAATGT CAGTCGTTACGCGTTTTCCTGTAAAAGGCAACCGTCCATGAAAGTCC TTACGAGGACGCATAGATGTTGAAAAATTAATAACGAGGATGTTGAAA TCCCATCTATGACATACGAAAGTGTGAAAAATATCTTGGGGTAAATGAT TGACCCAGCTGGTAAGATTGCTCTCCGATTGAGGAAATGGAAGCTTTGG CTAACTAGGTTAAGGGAGTGTAAAGCTCAACCTGATCAAAAAAGTGAAG	AGAAACTT GGATCAGT GATCAAGG CTAATGGA TGACGGG ACTCCATG GATAAGG AGATATAA AGATCTTA TTTGAACG CATCTTAA GGGGTTAT GGGGCTA ACACCCCC TTAATTCT GGTGACAA TTTATTGA CCGTT (SEQ ID NO: 1147)	CGGTG GGAGTA TGATCA GTGGAG CTGACTT TCCAGA CAACTC ACGCGG ATTGCG GTGCGG TGGATA CAACAC CTGGTA TAACAT ATGAAG GGTTC ATCTAGT ACAGGG ATAACG ATCCAT GGGAGC AAACTA ATTAGTT GGAGGT AATCCA ACGCGG CTGTTG AGTCAG TTTTAA CCGCCA GTCAAC TCTTGTA GGTTAT CGGTCT TCGGCA GACCTT GGACCG CCTAGC	ITDSIELVTDVNEVEM VTSNSTNEEQPINAP VEPAVIEADLGRQDA KYLASLRQSDCTNA SDRWTLAYCRGEVD WCKTSRLFKVSRHA RGLRQQRVENWEF PEGFRPNRLKWR KYSFLQSCYRTKKET VSKILDGTFKDTPEEEI RPELEEVQRVYVDR EVRTQLDTRTVHID ERFDLVSYGRITIREV QDAISASKDASGGP DGLLLQDVKKASPRQ LCIIFNMWYHGPV VENRCRTILLHKGE KHLTSNYRPVTIGNM LNRVYAKIWDRIIRK NLQLHVRQKAFVPLD GCFENVKTIQICILQSY RRSRREHNVVFDLA KAFDTILHDSIEKALLR KGIPRSVIKVVDSLVA GAVTSITVGKTKTRPI CINSGVKQGCPLSPLL FNLVIDELAEERLEATG CGLDLEGHVSSMAF ADDYVLLAKDSVEM NVLMMVNCNTFFEEK GLAVNPAKCSLRVL PVKGKRSMKVLTRTH RWWKINNQDVEIPS MTYESVGKYLGVMI DPAGKIALPIEEWKL WLTRLRECKLKPQK VKVLEVVCARANYV LRMSGGICELRKWS

				GTGCTGAAAGAGGAGTAGTTTGTGCCCGAGCAAACTATGTTCTCCGGATGT CCGGTGGCGAATCTGTAGCTCCGTAAGTGGTCACGATTTGTGAGGG GATGGTGAAATCCATCATTCATCTCCCGCATGGTGCAATAGCGAATG GATGCATCGAGCAAAAGCTTAGGCATCTCTGATGTAGTGTAGCAAGT GTCATCAACGAATGAGAGCTCGGAAAAAATGGTAAGTCAACAGAC GGAGTAGTCGAGTTGTGGGGCCGCAATTGTGCAGACAAATAGAGTT TTGTGAAAAGGGCCGATTAGCAGGCATAGAACTGGATGCCGCCAGG AAGTTCTGTAGGTTAAGAGGTGAACAAAATGGCAATCAACCAAT GGAGGCCCTCAAGACTATAGCAGAGTCTCGTGAGCGCGCACTGG TTATTGAAAAGAAATATAAGACCTGAAACAAAATTTCTAGTTTGGAAAG CAATGGCAGGAGTATCCAAACAAGATCAATCTGTCTAGAGCGTAGC CGACCAGACTCTCAAAAAATGTCGGTGTGTGTTTAAACAGCAGAACT GATTGTACATCTTGGCCGATGTCCTACAGTCGGGATCGTACTCGA AACGTCATAACTTGTCTTGATAAACTGCCAAAGAGCTAAGACTCAAT GGTGGCCAAAGCAGACGGGTGTGGCGGAGAGGATGTGTCTCTCTGG GAATGGCAGGCTTAAAGCCGATATTGTTGAAAGATGATGGTGT AATTACTGTCATCGATATGGCATGTCGTACGAGAAATCGGAAAGACAC CTAAGTCAATGCGAAGATGCAAAAGTTGTAAGTACGAGCCACTAAGG CTTGATAGGAGTTGGACTCAAGAACTTGAGGGAAATACGGCAGAAAT GCTAATGAAATATCAGTTGTAGGATGTCAGTAGGGGCGATTGGAACA ATTACGCGTAAACCCAGCGGATCTAGCAAGTTGAACTGGCCAAGG TCGGAAGACCGTTACAAATAATGTCATGTAATGAAAGCGCCAAATAT AAGACGACATCTTCGGGATCGAGACTTAGAAATTCGGTGAATGCC GAGGTAGTTGGGATAATGATGCACAGCTCGTAAGCGACTTGTGCA CGTATGCCGCTAACGCTTAGCTCGATGAGTGCATGTCAAGACGGTCGG GAGTATGATCAGTGGAGTGAATTCAGACAACTCAGCGGATTCGCG TGCGGTGATACAAACCTGGTATACATATGAGGGTTCATCTAGTA CAGGGATAACGATCCATGGGAGCAAACTAATAGTTGGAGGTAATCCA ACGCCGTGTGAGTCAAGTTTAAACGCCAGTCACTTTGAGGTTAT CGGTCTCGGAGACCTTGACCGCTAGCGCGGCAACAGTTTGTGCG TCGACTAACATGATGATTGCGAGAGAAACCCAGCTTTGTCTACTATGT GAGGATAAAATCTCTGTCCATATGATCCTTTGAAGGGAACAGCGCTT GAGCTTGTCTCGCGTTGGACCTTTAGTCTGTAATATTTCTTGATATTA TGGACGAAAAGGTAGTATGGTTGCA (SEQ ID NO: 1546)		GCCGGC CAACAG TTTGTCG TCGACT AACATG ATGATTT GCGAGA GAAACC CAGCCTT TGTCAT TATGTG AGGATA AAATCTC TTGTCCA TATGATC CTTTGAA GGGAAC AGCGCT TTGAGC TTGCTC GGCGTT GGCACC TTTAGTC TGTAAT ATTTCT TGATATT ATGGAC GAAAAA GGTAGT ATGGTT GCA (SEQ ID NO: 1270)	RFVRGWVKSIIHPPA WCNSEWMHSSKGL GIPDVVSGIVQMR AAEKMAKSTDGVVR VVGARIVQTNRLVW KRAGLAGIELDAARK FCEVVRVKNIGNQTN GGALKTIAESSVRH WLEKNIRPGNKILV WKAMAGVIPTKINLS RGVADQTLKKCRCC GLTAETDCHILAGCPT SRDAYSKRHNLLCDK LAKELRLNGGSPRRV WRERMCLSGNGRRY KPDIVVKDDGVITVID MACPYEKSERHLSQC EDAKVAKYEPLRLDR SWTQLEGNNGRSA NEISVVGIAVGAIGTI TRKTQRILSKLAKV GRPLQIACNESAIIR RHLSGSRLNLR (SEQ ID NO: 1392)
R2	R9Av	GQ398 057	Adineta vaga	GAAATAGTTGCAATGGTAGGTGATGGCGCCTCTGTCTCTCTTCGCG TGGATATAGTTTGACGATTTTGACAGGTACTGTCTCTCTGAGTTCA GCACGAGTTTGAACAGGCTTAGCGATAGACCTTCGAACCTTGAACACTG TTGTGAAGCTGGCTGGGCCCTGCAGATTTCTCGATTAGAACGTGAGT	GAAATAGT TTGCAATG GTAGGTGT ATGGCGCC	ACTAGT CTCCTC TTCTATT AGTCAG	MNLPIREHAVSVHNI NKFNYLCQLCKSYD TINSVKAHVACRRQ KNASSTAVPTNVIN

GTTACGTCAGAAATGACCCACCAAGTGGTTAGTTCTACGTTGCCCTGGAA	TCTGTGC	TCTAAT	NNQLAINTNQVISRN
AGGAGAAAAGTTGAGCTAAAATCGCACGGCTAGTTGTTATCAAAATAG	TCTCTTC	AATTTT	PLQCECLMKQVDF
GCACGGTGAGGAACTCTTATGTACCTGACTAAAGTACTCACTTGTC	GCTGGATA	CTTACAT	YAKDTKALVTHMRTK
GCTGGTTTGTCCCTCGCATTTGACTTATCTGATCGCACTACCCACCA	TAGTTTGA	TCTACAT	HAAAYEESKKVATRR
AACGAAACATAAATCTAGCTCGTGTATCAGTCCACACGCTGTGCAGTC	CGATTTTG	CTAGTTC	VAWSPDEQILAELE
GGATTGAGGGAGCGTGTAGTGACAAGCAGGATAAATATTAACTAGT	TACCAGGT	CATTATT	VKLKIQKGQLSLRV
TAATGTTAAGCGTTCAACATTCCTTATCCAATTGGAAGAGTTGACTGTG	ATCTGTTT	AAATTG	VEYNKCADSKAPSR
AAGTTTGTATGAAGACATTGGACAAATGAATTTGCCGATTCGAGAGCA	CTTGAG	GTATGA	SKDAIRTRRQHQDYK
TGCCGTATCTGTACAAATATAAACAAATTAATTTATTTATGCCAGCTAT	TTGAGCAC	TCAGTG	LLRSLQSQQPPVGS
GTCTAAGCTTATGATACTAATTAATAGTGTAAAGCTCACTATGTTGCA	CAGTTTGA	CTATCTC	EDSDSISSNNNPLT
TGCAGAAACAGAAAGTGCCTCATCCACAACAGCTGTTCCAACCAATG	ACAGGCTT	TGCTAC	TTHNVTPTDSSNVV
TCATCAACAACAACCAACTTGCTATAAATACTAATAAGTAATATCAAGA	AGCGATAG	ACTCAAT	LLIQKIRESDSIVKIT
AATCCACTTCAGTGGTGGTGTCTAATGAACAAGTTGATTTCTATGC	ACCTTCGA	GCTTAAT	NLKLNTNMLNAASA
TAAAGATACAAAGGCACTAGTCACGCACATGCGTACTAAACATGCTGCT	ACTTGAAA	CGTATG	FINQNNNMIDPLELS
GCCTACGAGGAATCAAAGAAAGTCGCAACAAGAGAGTTGCCTGGAGC	CACGTGTTG	TTATTGA	MRGIEEDVKAIRDE
CCTGATGAGGATCAAAATTTCTGTGAAGTGAAGTCAAAATGAAAAAGA	TGAAGCTG	CAGTCT	LQKPTRNVPSSTSR
TACAAAAAGGTCAATTAATGCTGTGCTTGTGTTGAATATAATAAATGT	GCTGGGCC	GACACT	KPTRNAKRLKSKKY
GCTGATAAATCGAAAGCTCCTCCAGGTCGCAAGGATGCTATTCGTACAA	CCTGCAGA	TGATTAC	GYQHLVYNNKKLV
GACGCCAACCAACATGATTACAACTATTGCTTCGCTCACTCCAATCTCAA	TTTTCTCG	TCTTACG	AEILDGETSGAKPPP
CAACGCCAGTTGGTAGGAAGACAGTGACAGTGACATATCTTCTAGTA	ATTAGAAC	ACATAT	MNLVEDYYRNIWSR
ATAACAATCCTTTAACACAACACATAATGTCACTCCAACGCCAGATTCA	GTGAGTGT	GCACTG	STIDDSPVNNIKTVNS
TCCAACGTTGTACTACTAATACAAAAAGATCCGTGAATCTGTAGATCCAT	TACGTCCA	TTTGCTT	DSIFAPISRDEIKLALS
TGTAAAAATAACGAACCTCAAAATGAACACGAATATGCTGAACGCAGCA	GAATGACC	CAGAGA	NTKSDSAAAGPDVTT
AGTGCCTTCATTAAATCAAAAATAACAACATGGATCCACTTGAACATCTAT	CACCAGTG	AACCAC	KEAKAIIDNLVAYNI
GCGTGGTATCGAAGAGGATGTGAAGGCAATTCGAGACAAAGAACTTCA	GTTAGTTC	TGTTTAT	WLGVDGPIPEQLKLNK
GAAACCAACGAGGACGTTCTTCTTCAACAACCTCGAGAAAGCCAACT	TACGTTGC	ATAGTG	TILIPKGNDSLKLN
CGAAATGCCAAAAGGCTTGAGAAATCAAAAAAATATGGCTATTATCAAC	CCTGGAAA	AAGTTC	WRPTISSILRVYNRL
ATCTGTACTATAATAACAAGAAAAAATAGTAGCGGAAATCCTCGATGG	GGAGAAA	CTCAGTT	LAYRMNKIFKTNDKQ
CGAAACAAGTGTGCTAAGCCACTCCAATGAACCTGGTTGAAGATTAT	AGTTGAGC	TTCTGTT	VGFKPVNGCGINISW
TATAGAAATATTTGTCACGTTCTACTATTGATGATTGCTGCTGATG	TAAAATCG	GATATA	LHSLKHARLNKNSIY
TATTAACCCGTTAATAGTGACTCTATATTTGCTCCAAATTCGCGTGATG	CACGCGCT	TTCTTCT	ACLVDVSKAFDSVSH
AAATCAAAATTAGCATTATCAAAATACGAAAAAGGATTCAGCAGCTGGACC	AGTTGTTT	TTCAATC	QSIVRALTMINGAPSL
TGACGCTGTAAACAATAAAGAAAGCAAAAAGCTATTATTGACAAATCTTTAT	ATCAAAATA	TCGCTTC	LVKLIMDQYTNVNTV
GTTGCATATAATATATGGTAGGTGTTCAAGGAAATTCCTGAACAACCTGA	GGCACGGT	TCCTTTT	ITCSGISNKNISSGV
AATTGAATAAAACTATCTTAATCCAAAAGGAAATTCGATCTTAGTCTA	GAGGAACT	CTACTGT	KQGDPLSLLLFNLVID
CTGAAAAAAGTGGGACCTATTACAATCTCGTCTATTATCTTAAGAGTATA	CTTCTATG	GTTCTTT	ELFDVIKQDQGYTIDN
CAACAGATTATTAGCATACAGAATGAACAAGATCTTTAAAACCTAATGAT	TACCCTGA	TTATCAG	IGTTNARCFADDLTI
AAACAAGTTGGATTCAAACTGTTAATGGTTGTGGTATTAAATATATCTTG	CTAAAGTA	TTTTTTG	SSSRMGMMNKLELTT

GCTTCACTCTCTTGAAGCATGCACGCTTAACAAAAATTCATATATG CTTGCTTGCGATGTCTAAAGCCTTGTATCTGTGTCACATCAATCAA TAGTAAGAGCTCTACAATGAATGGTGACCATCTTGTAGTGAAT AATAATGGATCAATATACGAATGTAATACTGTGCATCAGTCTGGTT CTATATCAACAAGATAAATATCTCAGTGGTCAAGCAAGGTGACCC ACTATCAGTTGTTGTTCAATCTGGTTATAGATGAATGTTCTGATGTA TAAAGGACCAATATGGTTATACAATTGATAACATGGCACCAATGC ACGATGCTCGCGATGATTAACACTAATATCATCATCTAGATGGGTA TGAATAAATGCTTGAGCTCACCAAGAAATCTTCAAGAACGTGGACT AAATGTAAACCCATCAAGTGCATGCTATGGCATGTCCAAAGGTTAT AAAGGAAAGAGTAAATCGAATCTGAACCACTCTTCTCTATCACCG ATGCTCAGATACCGATGTGGCTATATTGATAAGACAACTCGATATCTC GGTGTAATTTACATCTATTGGTGCAATTGATGCAAAAAGAAATCAAAA AAGACCTTCAGGACACACTCGATAAGCTTGAACATCTTAAACTCAAAGC TCAGTGCAAAATGGATCTCTACGAACCTATATGATACCAAGATTCATGT TTCAATTAAATAGGAAATAGCTAAACGAATCCTACATCTGCCATATCAAC ATCTTAATTAGGAAATAGCTAAACGAATCCTACATCTGCCATATCAAC GAGTAGTGAATCTTTACTTACCCTCAAAGAGGAGGTCTTCAACTAA CCTCACTTAAAGAGCAGTTGGTTAGCCAAAATAAAATACACAAGAA GATAATGTCCAGTAAATGATCCAAATGTTATGCTACTTGTAGAGCCAG AGGAGCCGTATTGTCAACATTTTATGAAGACCTTAACTTGGAGATT CTTAAACATTAACGAAATGAATAACATCAAGAGTCTTCATGAAAGA AAAAAGAAATCTATTTGCTCAAAAATTCACGGTGTGGCTTCGAAGTA TTCTCATCAAGTCTTTGACGAACCAATGGATTAATGGCGAAATTAAGA CAATGACAACTAAACATACATTAACCTAACTTAACTTAGAACAAATACT CTAGAAACTCGGTAAACATCTCGGGGACTGAACATCATAAAAACAT GTAGAAAGTCCACGTAGCTGACGAAAGTCTCATGCATGTGCTCCAATG TTGCTCTTACCAAGGTTTACGATACTCTCGTCATCACAATAATGTG CCAAAGTAGCAATAAATGGTAATGAATGGTTATGGTGATTTCTGTGA GAAGAGTTATCCAGATCCAAACAACTCAGTTTCATACCTTCGACCCGAT ATAATTGCAGTAAAAAATGGTCATGTTATTGTTCTTGATGTAACGGTTGT GTACGAAGTAACTGGTGTACGTTTATTAAATGCCTACCAAAACAAAAATA AATAAATATAATGCGATTATGGTACAAATCGAGCAAAATGTTCAATTGTG TTAATGGTGAAATGCGATGGTCTAGTAATGGATGACGCTGGTTCAATTCAT CACAGTCAACTCCACATCTGGCATCAAAATGGGATCTCTTCCATAGAAT TAAATATGTGGCTATAGGATGATGGAGGATTCGCTCAGAAATCATGTCC ACATTTCTCAAAAGCTATCAGATGAACCTAGTCTCTCTCTTATTAGTCAGT CTAATTAATTTTCTTACATCTACATCTAGTTCATTAATAATGGTATG	CTCACTTG TGCGCTGG GTTTGCTC CCCCTCGC ATTGACTT ATCTGATC GCATACC CACCAAC GAAACATA AACTAGC TCGTGTA TCAGTCCA CAGCGTG GCAGTCGG ATTCAGGG GAGCGTG TAGTGACA AGCAGGAT AATATTAA CATAGTTA ATGTTAAG GCGTTCAA CATTCCT ATCCAAT GGAAGAG TTGACTGT GAAGTTTG TCATGAAG ACATTGGA CAA (SEQ ID NO: 1148)	TGGA AATTGA GAATA ATAAG T (SEQ ID NO: 1271)	KFKERGLNVNPSKC MSGMSKGYKGKS KIESEPLSFIDAQJPM LGYIDKTRYLGVNFT SIGAIDAKRIKKDLQD TLDKLEHLKKAQCK MDLLRTYMIPIRFMF QLIHTELYPKLLIKMDI LIRKLAKRILHLPSTSS EFFYLPFKEGGLQTS LKEAVGLAKIKLHKKI MSSNDPMLCYLIESQ RSRIVEHFMDKDLKG DSLTLNEMNNIKECF MKEKRISFAQKIHG GFEVFSSPLTNQWI NGEIKTMTTKTYNSI KLRTNLTETRVTTSRG LNIKTCRRCHVADES LMHVLCQCSSTKGLR YSRHHKICAKVANKL VMNGYGVFREKSY DPNNSGYLRPDIAV KNGHV/VLDVTVVE VTGATFINAYQTINK YNAMVQIEQMFNC VNGELHGLVIGSRGSI HHSQHLHIWHQMGFS SIELKYVAIGCMEDSL RIMSTFSKAIT (SEQ ID NO: 1393)

R2	R2OI	LC349 444	Oryzias latipes	ATCAGTGCTATCTCTGCTACACTCAATGCTTAATCGTATGTTATTGACAG TCTGACACTTGATTACTCTTACGACATATGCACGTGTTTCTTCAGAGAAA CCACTGTTCATATAGTGAAGTTCCTCAGTTTTCTGTTGATATATCTTCT TCATTCGCTTCTCTTTCTACTGTGTTCTTTTATCAGTTTTTGTGGA AAAAITGAGATAAATAAGT (SEQ ID NO: 1025)	CGCACAGG GGACACA GAGCTGC CCAAGTAC CGTCCCG AGGGAGC GGGAAAC GGGGGG TGACTATC CCCTGGG TCCGGCGA GAGCGTG GTCTACGG ACCAAGG GTGGCTGT GGGCGG CTGCTCT CAGGCCAG TTGATTAG TTACGCAT GGGCTGTA CCTCCACG TGCTCCCG CTGGTAAC GACTTGTC GGCTAAAT CAGCCCGC CCACCATC TGGGATAT GGTTGACC GTCTAACC CCAGTACT CAGGTAC	GGGGGA CAGCTG GGAGTC TCGGCA TGATTAC AAATCTT GCGCTG CACTCG GATGTC GTCCCC GTGACG GACACA TTAATCC GGAAAG CGAGTG GTGACT CGCCTC AAG (SEQ ID NO: 1272)	MGTDTVVVGQDYP GLSKRVPARLVAGP MLRERSCHAHVFRA GHMWNWRTSLPSG RWDQPALEKSRVLTR SVATATDPEITSYPGK SVSTSTQVQEEDWC SRESGWISPLAPEE PSVSEITASMVATM RVATEEVVLEPQPEQ VVTILPEHGRNVPPG LAEQDTASPIEVSLL PDLAENCPLCGVPSG GLRLGKHFAVRHAG VPVTYECRKCWRSR NSHSISCHVPKCRGR ARMPSGDPGIACDLC EARFATEVGVAQHK RHVHPVEWNKVRLE RRGARGGGIKATKL WSVAEVELRLIREH GDSGATYQLIADELG RGKTAEQVRSKRLL RIDTASNPDDAEVE EERLESIAVRSSSRP PSLVATRVREAVARG ESEGGEIRAIALIRD VDQNPCLITSASDIIS KLGRRVDPKPRPV VREQTQEKGWVRL ARRKREYEAQYLYS RDQARLAAQILDGAA SQECALPVDQVYGAF
				ATCAGTGCTATCTCTGCTACACTCAATGCTTAATCGTATGTTATTGACAG TCTGACACTTGATTACTCTTACGACATATGCACGTGTTTCTTCAGAGAAA CCACTGTTCATATAGTGAAGTTCCTCAGTTTTCTGTTGATATATCTTCT TCATTCGCTTCTCTTTCTACTGTGTTCTTTTATCAGTTTTTGTGGA AAAAITGAGATAAATAAGT (SEQ ID NO: 1025)	CGCACAGG GGACACA GAGCTGC CCAAGTAC CGTCCCG AGGGAGC GGGAAAC GGGGGG TGACTATC CCCTGGG TCCGGCGA GAGCGTG GTCTACGG ACCAAGG GTGGCTGT GGGCGG CTGCTCT CAGGCCAG TTGATTAG TTACGCAT GGGCTGTA CCTCCACG TGCTCCCG CTGGTAAC GACTTGTC GGCTAAAT CAGCCCGC CCACCATC TGGGATAT GGTTGACC GTCTAACC CCAGTACT CAGGTAC	GGGGGA CAGCTG GGAGTC TCGGCA TGATTAC AAATCTT GCGCTG CACTCG GATGTC GTCCCC GTGACG GACACA TTAATCC GGAAAG CGAGTG GTGACT CGCCTC AAG (SEQ ID NO: 1272)	MGTDTVVVGQDYP GLSKRVPARLVAGP MLRERSCHAHVFRA GHMWNWRTSLPSG RWDQPALEKSRVLTR SVATATDPEITSYPGK SVSTSTQVQEEDWC SRESGWISPLAPEE PSVSEITASMVATM RVATEEVVLEPQPEQ VVTILPEHGRNVPPG LAEQDTASPIEVSLL PDLAENCPLCGVPSG GLRLGKHFAVRHAG VPVTYECRKCWRSR NSHSISCHVPKCRGR ARMPSGDPGIACDLC EARFATEVGVAQHK RHVHPVEWNKVRLE RRGARGGGIKATKL WSVAEVELRLIREH GDSGATYQLIADELG RGKTAEQVRSKRLL RIDTASNPDDAEVE EERLESIAVRSSSRP PSLVATRVREAVARG ESEGGEIRAIALIRD VDQNPCLITSASDIIS KLGRRVDPKPRPV VREQTQEKGWVRL ARRKREYEAQYLYS RDQARLAAQILDGAA SQECALPVDQVYGAF

AAACAAA (SEQ ID NO: 1149)	AAGGGATGGGTAAGGCGGCTTGCCCGGCGGAAAGGGAGTACAGAGA AGCGCAGTACCTGTACTCAAGGGATCAAGCAAGGCTGGCGGCCAGAT CCTCGATGGTCCGCCAGCAGGAATGCGCCCTCCCGGTGGACCAGGT CTACGGAGCGTTCCGTGAGAAATGGAAACCGTAGGGCAGTCCACGG ACTTGGTGAATCCGGACGGGTGCACGCGCAGACAACTGGGAGTTCTA CTCTCAAATCTGGCGGTGAGTGAAAGAAACCTAATGAGAATGGCT AACGGCAGCGCCCGGACGACAGGATAGCAAAAGGCTCTGCTT GACTGGACCCCGGGGTGACAACTGGCAGCGGTACACGACGTGG CTGATCGGTGGGTATACCAAGGCTTCAAGGAGTGACAGGACTAAG CTGTACCGAAATCAGCAGCCCGTCTGAGTTGACGAGACATCGGTGGA TGGAGCGCGTGACGATTGGTCTGATGGTACTAGGCTGTTCAAGTCGG ATTCTAACGATGAGGCTAACCCGAGCCTGTCGATCAATCCGAGGCAGC GCGGTTCTTGCCCTCTCGAGTGATGCGCGGAAACCTGTTGATCTT TGACGAGATCGTCAGGCGCTCGAGGCGGACGGGGCGCGCTGGCAG TGGTGTGTTGGACTTTCGAGGGCCTTGAATCCATCTCACATGAACAT ATCCTGTGTTCTGAAGAAAGCGGGCTTGACAGGACGTTATCGGGT TGATCCGAAACTCGTACGTGGATTGCGTGACCAAGGTTGAGTGAAGCAGGA AGGGCATGACACCAACAATAAAGGTTGGAGTGAAGCAGGGA GACCCATGTCCCTTGTCTTCAACCTGGCTATGGATCCCTCATCCAT AACTCGAGACGGCGGAAGTGAAGTGAATGGGCGATCTTCAATC GCCAGCTGGCTTTCGCGACGATCTGGTGTGGTGAAGTCTGAGG AAGGCATGGGAGGAGTCTGGGATTTTGGAGAAATTTGCCAACTGA CTGGGTGAGGGTTCAGCCAGGAAGTGCACGGTTTCTTATGGACAA GGCGTGGTGAAACGCTGTGAACCTGGGAAATCTGTGGTCAACCGAT CCACATGATTCCTCCGGGGATCAGTTCGTTATTTGGGAGTCCAGGTA GGCCGGGGCGGCGGTGATGGAACCGGATCTTATCCCTACGGTCCAC ACGTGATCGAAAGGATCTCGGAGGCTCTTAAAGCCCTCACACGCA TGAGGGTTTGAATCAATCGCTCTCCCGGATAATTTACAGGCCGAT CTAGGGAAGTTACGGTAACCAATGGCCAGATAGATGGGATTGTC CGGAAGGCTGTGAAGAGTGGCTCAATTTGTACCATCCACGTGCAATG GACTGTGATTCACGGAACCGCAGCGGTGTTTGGGCTCTCTAAAGCT GGAAGACTAATCCCATCCGTGCGCAGCAAGCGTATCTATCGGATGTC AGGTCTCCGGATATCTGGACACGGCGAATGACCAAGCCTTCTGTGTCAA AATCTGACTGGGAGATGTTGCGGTCCAAGCGGGAGGTGAGAGGGGC AGTGCACTGTAAATGGGTGCGGTGAGGGCTGCCCGACCGATGTGGAG AGATCGCAGACTACCCAGACTGCGGCGTGAAGAAACCTGGCATGG TCGGCCCTGCGGGTGCAGGGTGTGGGTGCAGACCAAGTTTCGAGGCGAC AGGACCAGCAGCTTGGATCGCGGAGCCCGCTTCGGTTGGGTTCCGCGC				REKWETVQGFHGLG EFRTGARADNWEFY SPILAAEVKENLMRM ANGTAPGPDRIKKA LLDWDPRGEQLARLY TTWLIGGVIPRVKEC RTKLLPKSSDPVELQD IGGWRPVTIGSMVT RLFSRILTMRLTRACP INPRQGFASSGCG AENLLFDEIVRRSRR DGGPLAVVFDVDFAR AFDSISHEHILCVLEE GGLDHRHVIGLIRNSYV DCVTRVGCVEGMTP PIQMKVGVKQGD MSPLLFNLMADPLIH KLETAGTGLKWGDLS IATLAFADDDLVLVSDS EEGMGRSLGILEKFC QLTGLRVQPRKCHGF FMDKGVVNGCGTW EICGSPHMIPPGESV RYLGVQVGPGRGVMI EPDLIPTVHTWIERIS EAPLKPSQRMRLVNS FALPRIYQADLGKVT VTKLAQIDGIVRKAVK KWLHLSPTCNGLLY SRNRDGGGLLKLER LIPSVRTKRIVRMSRS PDIWTRRMTSHSVSK SDWEMLWVQAGGE RGSAPVVMGAVEAAP TDVERSPDYDWRRR EENLAWSALRVQGV GADQFRGDRTSSSW IAEPASVGFQQRHWL

R2	R2_LP	AF015 814	Limulus polyph mus	AGCGCCACTGGTGGCTGCCCTGGCGCTGAGGGCTGGGGTGTATCCGA CTCGGGAGTTTCTGGCTCGGGGTAAGGAAAGTCAGGAGCAGCTTGCA GACGTGCCCCGCGAGGTGGAATCATGTTACACATACTTGGGCAATG TCCGTTGTTACGGCAACAGAAATGCGAGGCACAAAGGTGTGTGT GCTTTGGCCACGGAGGGGAGAGTTCCGGTGGAGCGGTAAATAGGG AGTTCCGTCTTGGAGCGCGCTGGCGTCTCAAGTACCCGACCTGGT TTGCAAGAAGCGGACACAGTTCTATTGTCAGCTGACCGTCCGGTAC GAGATGGATGGAGACGCTAAAGGGCCGCTCGAGAGAGGTGAA ACACTATCTCCAGTAGGCAACAGATAACGGACAAAGTCCGAGGGCG TTGCTTTAAAGTATGGGTCCCTGTAGGTGCTAGGGGAAAGTGGCG GCGAGCAACAACAGTTTTGGCTGAGTTAGGCTCCCTGCAGGTCGG ATGAGGACCTTTGCCAGGCTGGTGAGCCGGAGGACTCTTCTTATTTCTT GGATATATTGAGGGACTTCATGCTGAGCGCGCGGAGGGGAACTCG GGTTGCTCTCATCCCTGCGGCAACGGGTGCCGGAATTGAGGGGACA GCTGGGAGTCTGGCATGATTACAAATCTTGGCTGCACTCGGATGCG TCCCGTGACGGACACATTAAATCCGGAAGCGAGTGGTGACTCGCCTCA AG (SEQ ID NO: 1026)	TGGGAGG AGACCCAA ACTATCCT AGGATGG GGCGGAA CCGACCAT ATGAGCCA TATTAACA TTGCCAC ACTATCCT CTGGAGGT ACCTCCTC GTGGTACG GCTGGATA TAGGTAAA TCCTGTAA CCAAATCC TCCAACCC GTGAAGG AGAAACAT AAAACCCA	ATTTGT CTCTTC CCCAAT GATGTC TACTAG CACGCT GCCGAA GCTAGA TAGATT GAGGAA TCTGCG TAATCTG TAATGA TTACGCC TCATGG GCATCT ATCGGT AGCGTC GACCTT GACGTT AAATTG	AALALRAGVVPTRF LARGKEKSGAACRRC PARLESCSHILQC VQANRIARHNKVCV LATEARFGWTVIRE FRLEDAAGGLKIPDLV CKKADTVLIVDTV YEMIDGETLKRAASEK VKHYLPVGGQITDKV GGRCFKVMGFPVGA RGKWPASNNTVLAE LGVPAGRMRTFARL VSRRTLLYSLDILRDF MREPAGRGTRVALIP AATGAAN (SEQ ID NO: 1394)	GIDGYMFGYARASG STSVSIQSSMTEGET NERATPRADSSSVSI QSSCVTEGECLPTD NCNPSVENQLPCVTE GRFERSVSLVTVRLP FRKVACDLCSEFLTY SKFAVHQANFHNS QACCTYCGKSDGNH HSIACHVPKCPWRRT VTFAANLSNFLCDLC NDSFKTKSGLSQH KHPCSRNAERILSLG VRTPSARPRQVWWS EEETRLREVEWVYSG QKNINVLCAHLPK TSKQVSKRRDLHRI RSSNVHGTPTTQSRG DPVEQVEEELDWE GMHPFPDPSKFC LDQLRQKGLTEPV

AACCGTCTGTAGAGAACCAAGTTACCGTGCCTAACTGAGGGTAGGTTTG AACGGTAGGCTCACTGGTGACGGTGCCTGCTGCCCTTCAGAAAGGTGG CATGTGACTTGTTCTAAAGAGTTCTTGACATATTCGAAGTTTCGAGTC CACGAGCAAACTCCACAATCAGAACTCAGGCATGCTGCACATATT GCGGTAAAAGTATGCAATCATCACTCTATAGCTGTACCGTCCGAA ATGTCCCTGGCGGCACTGTTACGTTTGTGCGAACTTAAAGCAATTTCT TGTTGATCTTTGCAATGATGTTTTAAGACCAATCAGGGCTTCGCAA CATAAGCGTCATAAGCATCTTGTTCAGGAATGCTGAACGCATCCTTC TCTTGAGTCAGGACGCCGTGGCCGCCCTCGCCAGGTAGTGTGGTCC GAAGAAGAACACGAAACCTCCGGGAAGTGGAAAGTAGTATTTCGGGC CAAAAGAACATTAAATGCTCTGTGCGGGGCATCTACCTGTTAAGACTT CCAAACAGGTCTCGGACAAAGCGCGAGACTTGACAGGATACGGTCTTC TAACGTACATGGTACACCCCACTCAGAGTCGTGGAGATCCTGTTGAA CAGGTCGAGAGTACGAGGAGTTGGACTGGGAAGGAATGCATCCTTTT CCGACCTGACTCTAAGTTTGTCTGTACCTTGATCAGCTGAGAGATCA GAAGGACTCACTGAACCGGTATGGCAGGAGATCGAAATCGTGGCACA AGAATGGGTAGAAAACCTGCCCATGTTCAATCGTCTTGGAAATCATGAG AGAACAAAGCAGGTGCCAGAAACAAATACACCTGCACGAAGACCA TTTAAAGGCGTCTCCATCGTGTGGAACGTTAAGCGGTTTCAGAGAA TGACGACCTCCAGGAAAGCGCTGGCTGAGGAAATCTAGACGGCC GGGAAGCCGTACATGTAACTCAAAAGGAGGAGATCAAAAGACCCT ATGATCAGGTCTACGGTGTCAAAATGATAGAGTTTCTTAGATGACTG CCCAGGCCACCGGGGCAATAACACCGACCTCTGAAACCGTTTACG CCAACCGAAGTATGAGTCACTCAGGGTATGAAGAACGGGGCGCT GGCCTGATAAGATTACCTACCGTTCTCCAAAACGTTTAAATGG CATCCATGTTTCTGGCAATGTGTTAACCTTGGCAATTCGCGGTC GCATCCCGAATGCATGAAGTCAAAATAGGTCACTCTATCCGAAAGG GAAGAGCAATCTCGGGATGTAGAACTGGCGGCCAATCACAATCTC CTCGATTGTTGCGGTATACACAGGATCTTGGACGCCGCTCTCGAG CGGGCGTGCAATTAATCCCGACAGCGAGGCTTCGCTCCCTCAGGCTG GGTGTAGGGATAATATCTCTGCTTCACTGCTATGAGGAGGGCTAA GCGAAAGGGAACCTCGGCTCTGGGGCTTCTGACTTGTGCAAGGCATTT GACACAGTTGGTCACAAACATCTTCTGACAGCCTAGAAAGGTTTCGCTG TCCACCCGCAATTCGTCGCAATTTGGAGGACATGTACAGTGGTTGTC GACGTCCTTCGAGTAGGACGCCAGTCTACTCGCCCATCTCTGATGA GAGGCGTCAAAAGGGGACCCCATGTCTCTATATTGTTCAACATCGC TCTGGACCTCTTCTCGTCAACTGGAAGAGGAAAGCCGAGGCTTTATG TTAGGGAGGGGCGGCGCCCTGTCTCATCTCTAGCATATGCCGATGATA	TATAGTGG CCTCGCCA ACCACTAT ATGTCCAA CGGAGG AGAACTA TCTCCGG ATGGAA GGAAACC GTAAACCG TGATGGGA ACTTACCG GCCCATC AGCTATTG GGTACCCG GTAGGGA CTTGCAAC CCTACCCT GTATTTGC ATTTTATA GGGAACC GGTCGGCC CTATATCA GAGTAGAC CGTTTATT AAATATGG GTGAAAT ATTAAACG TAAAGCT ATGTTTG GCGTCCGT GTGTGCC AGGGCGG CGGCAAAA CCCGAGCT ACTTGCA CCAACTGG GGATGGTA	GGTAAT AAGAAA TATCGA (SEQ ID NO: 1273)	WQEIIEIVAEQWVEN LAHVQSSWNHERTT KQVPENNTPARPPFK RRLHVERYKRFQR MYDLQRKRLAEELD GREAVTCNLKKEIK DHYDQVYGVSNDRV SLDDCRPPGANNT DLLKPTTPTVEMDSL QGMKNGAPGPKIT LPFLQRLKNGIHVSL ANVFNLWQFSGRIPE CMKSNRSVLIPKGS NLRDVRNWRPITISSI VLRLYTRILARLERA VQINPRQRGFVPOA GCRDNIFLLQSAMRR AKRKGTLAIGLLDLK AFDTVGHKHLTSL RFAVHPHFVRIVEDM YSGCSTSRVGSQST RPIVLMRGVKQGD MSPILFNIALDPLLRQ LEESRGMFREGQA PVSSLAYADDMALLA KDHASLQSMGLTVD KFCSGNGLGNIAS AGLLRGANKTFTVN DCPSWLVNGETLPM IGPEQTYRYLGASICP WTGINSVPVKPTLEK WIANITESPLKPHQR VDILCKYALPRLFYQL ELGTLNFKELKELDS MVKQAVKRWCHLP ACTADGLLYSRHRDG GLAVVKLESIVPCLKI KTNLRLVHSTDPVISS

TGGCACTACTGGCTAAAGATCAGGCCAGTCTTCAGTCGATGTTGGGCAC TGTGGATAAAATTTGTTCAGGGAACGGACTTGGCCTTAACATCGCCAAA AGTGCCGGACTTCTGATTAGGGAGCGAATAAGACCTTCACTGTCAATG ACTGCCCTTCTGGCTAGTAATGGTGAACGCTCCGATGATCGGTCC CGAACAACTTACCCTTATCTGGGGCAAGCATCTGCTCCGTGGACTGGG ATAAACAGCGGGCTGTTAAACCCACCCTGGAGAAATGGATAGCCAATA TCACAGAGTCTCCCTCAAGCCACATCAGAGGGTCGACATACTCTGTAA GTACGCTTACCCTGGCTGTTTACCACCTTGAGCTGGGCACTCTGAAT TCAAAGAACTGAAGGAACTAGACAGCATGGTCAACAAGCTGTCAAAC GTTGGTGCCATCTACTGCTGTACGGCTGACGGCTGCTATACTCCCGT CATCGTGATGGGGTTTAGCTGTAGTAAATTAAGAGTCTCTTGCCCTTG TCTAAAGATCAAGACAAATCTCAGACTAGTGCAATTCGACCGACCCCGTC ATATCATCTTTGGCGGAATCCGATGGTTAGTGGGTGCCATCGAGGTA TTGCTCAAAAGGCTGGGCTCCGATCCCTACGCCTGACCAGCGATCTGG AACATATCAATCTAAATGGAGAGATATGGAAGGAGAGAGCTGGGAAAG GTTGGCCCTGCACGGGCAAGGTGTGGAGCTCTTCAAAGGCTCAAGATCT GCCAACCACTGTTGCTAGGCCAGTTGGTATGAAGCCACACCACTGGG TGAAAGTGTGGCAATGAGAGCTAATGTATACCTCAAAAAGAGGCT CAGTAGAGGGAATCTATTAAGAACAAAGATTCGCCAAGTGTGCGGG ATGCATCAATGAGGGAGACCCTATGTATCAAGTGGTCAATGCCCG AAATTGAAGTCGATGAGAATAAGGCGCCACAATAAGATCTGTGAGCAC TTGATCGCCGAGGCCAGCTTAAAGGCTGGAAGTTCTGCAAGAGCCTA CCTTGGTTACAGACAAATGGTGAACGTGCGGACCTGATCTGATCTTCCA TCGTGATGATAAAGCGTGGTTGTGACGTGACGGTTCGCTACGAAAT TCGAAAGACAGCTTGAGAGAAGCTTATGCTTCTAAAGTTCGAAGGTATG GATGTTGACCGAACAAATTAAGACCTTACAGGGGCTACCTCCGTTGT TTTTCATGGATTTCCAATGGGTGCCGCGGTGCTGTTTCTGAAAGCT CGGACGTGATGCCGACCTGAACATTCGGTCAAAATATTTGAAGAGTT CTTGAGACGCCACCTCTATATACACTGGACTTATATGGAATCGA ATAACGAACAATATTTAGAAAGCTTGACCAATAAATTTGTCTCTTCC CCAATGATGCTACTAGCAGCTGCCGAAGCTAGATAGATTGAGGAATC TGCCTAATCTGTAATGATTACGCCTCATGGGCATCTATCGGTAGCGTCG ACCCTGACGTTAAATTTGGGTAATAAGAAATATCGA (SEQ ID NO: 1027)	GCTTCCGA GCGATTCC CTGGCGAC GTGGGACC GATCGACG ATGGAGTC CAAAATC CGGAATAG AGGAATTG AGAAATAC CTATTCCA CCACCGGC TCACATAC CCAAAGTG AACCCTGT GCAACTAG AGTACAAC CTATCTGT GGCGGTA GGTGCCGA ACCACTCA GGTGACG GGCTTGT TATTGATG TCTCCCTA CGAGACAC GAATTGTG ACAAATCC ACTCCGT GGACAAT ACCCGATC TATGAACC TGTACCG ATATTAGA CAAGAAAA TAAAGAAC TGACAACG CCTAGAGC	LAESDGLVGAIEGIAQ KAGLPIPTDQSRGT YHSNWRDMERRSW ERLALHGQGVLFKG SRSANHWLPRPVGM KPHHWVKLAMRA NVPTKRGLSRGNLS KNKDSAKCRGCTSM RETLCHLSGQCPKLS MRIRHNKICEHLAE ASFKGWVQLQEPTLV TDNGERRRPDLIFHR DDKAVVWDVTVRYEI SKDTLREAYASKVRR YGCLTEQIKDLTGATS VVFHGFPMGARGA WFPESSDVMADLNI RSKYFEEFLCRRTILYT LDLLWKSNNEQYLER LAP (SEQ ID NO: 1395)

NeS L	NeSL-1	Z8205 8	Caenor habditis elegans	GCTCACTTTCTATCGTGTAAACCGTAGCTTACACTCCAGTGAGTGTA TAAAGGTTATTCGATAGAGGGTGTCTCCCTCTTCTTGGGTAATCTTCG GCGGTCGGGGTCTCTCCCTCGTCTTTTAAACTTTTCTTCTCATCC ACTCTTTGCTCTTTTACTAACTCTGTACTCTATAGTCTTTCTCATCC CCCATCCGCCGTTGGCAAGTTTATTACTTTGTAAATCCATATTTAT CTCTCACCCGTACAGAAAGCGTCTCTCTCAACGCTTTCTGTACTT TTTCTATATTTTCAATTAACATAATTTTCTGTATATACTAACCTAACCTCC ATTGTCAATTACTAACTAATTTGTACACGGATTTTCGATGTCGCCGAA AAGGACGTCAACCAATGGTTATGGTCAATCTGTCAAAATGGCAACCCAG TGACATGCTGAAGCAATTGGAACAGGAAGTCTGGGCCACACAGCG GTCCAGGCATCCGAACACCGCTGGCAATCAATGCAATGTTTGATCCC CCCAACAGGATCTCTCGCCAGAGACTCATGGTCTCAACCAATCAAC GCATCTTCAAAATCAAAAGGAGCGGATCAGGATGGGTATAAGACCTGG TCAGGTAAGGAACAAATATGGTGGGGGTGGGCTCACAGAGCAGGGG ACCCAAAGCGTCGTGTCAGCTGGTCAGCATACAAGGAAGCGAAGTGA CCGTCAAGAAATCTACCCGTCGGATGAAATATTCAGTTGTACTCCAA TCATGTGATATCAAAACAAAGCTGGCTATGGCCCTGAGGACCTAAAGC ACCTGACTCGTCATATCAAGAACGAGCATGGTCTCAAAAGCTCGCTGGGC ATATCAATGTGGATTGTGCAATGAGAAAGTCGACCAAGTGTATCGGA AGGCCACAAATGGATGGAGGCACACATGGTCGCCGTTACCAAAAGCTC TGCGGAAAAAGGATAAAGTCTATCAGAAATGCACGGGTGCAAGAGT TGCAAGACAGCTACAAGCTGCTGCTCCATCGCTTACTGTGCGGGGAAG CACAAATCAGGCTCTAGAGACGCTGCCAAAGATTTCGATGACACCAACAA AGGATGATGACCCGAAACACAGGATCTATCAGACACGAAAGCTAGTGA AAAAGTCGACTCAGAAACACAGCAGAGCCAAACAGATGAAGGGTCTAGAG GCCAAAGTAGCGATCCATTTTTCAGAAATCCGTCAAAAGCAAGGAAGAG	TTCAGGCA GCATGTCT GTAAGTAT CCAGTCAT CGAGCGTG ACTGAGG GCGAAATT GATAATAA CTCTGAAA CTGA (SEQ ID NO: 1150) GCTCACTT TCTATCGT GTTAACCG TACGTTTA CACTCCCA GTGAGTGT AATAAAGG TTATTCGA TAGAGGGT GTCTCCCT CTTCTTGG GGTAATTC TTCGGCGG TCCGGGGT CTCTCCCT CGTCTTTT TTTAAACT TTTCTTTCT CATCCACT CTTTTGCT CCTTTTAC TAACTCTT GTAAGTCTA TAGTCTTT TCTCATCC CCCATCCG CCTCCA GGGCAC GCCGCA CGCCAA AAGTCC TGGCAT AACTCT GCAAAAT AACATC AAACGT CAATCA ACTCCAC AAACTCT CCACTCT CTTCAA GTCTTCT CGGTGC TTCCAAAC ACCACA ATGGTG AAAGCT CCTTCAC CTTTTCC CTCCAA AATCTT CCCATGT MLRRKGRHRMVMV NSVKWQPSAHAEAI GTGKSWAPQRSQAS EHGWQSNAMFDP NRILFARDWSLNQS THLQNRSGSLGIR PGQVRNNMVG HRAGDPKRRVELVSI QGSEVTVRTYPSDEI FSCYSKSDIKTKAGY GPEDLHLTRHIKNE HGLKARWAYQCGLC NEKSDPSVSEGHKW MEAHMVAVHQSSA EKRIKSYQKCTGARV AEQLQAAAPSLTPVG KHKSGSRDAKDSM TPTKDDDPKTRIQ RSVVKSTQKTAEPT DEGSRGPYASIFQKS VKARKSLALLCELSP KPMNPLPTNLTKE GNSRELAKEEAPSEGI DDIVIDDESESPR RKRFNTWCLDHESRR EAWLDDTIFWVYIS

CTTGGCGCTTCTGTGAATTAAGCAGCCCTAAGCCTATGAACCCCTTC	CCGTTGGG	GGGGA	LCRGSTKYSALDPCL
CTACAAATGAGCTAACTCTGAAAGAGGGAAATCAAGAGAGCTCGCCA	CAAAAGTTT	GTCTG	WSMYKVKGSRYLDR
AAGAGGAAGCACCATCTGAAGGTATAGACGACATCGTCATCATGATCT	ATTACTTT	TTCTGT	LESSITYFFPICEEDH
GGACGAATCGGAGGATCGCCACCAGAGGAAGAAACGATTCAACACCTG	GTTAAATC	AAGCTC	WTLVLKDNSYYAN
GTGTCTGGATCATGAGTCAAGCCGTGAAGCATGGCTGGATGACACAGC	CATATTTT	TCCGGA	SLHQEPGPGVRDFIN
AATCTTCTGGTACATCTCTATCTCTGCAGAGGAAGTACAAAGTACTCAG	ATCTCTCT	GGCTGC	DSKRARKEFKVQVPL
CTTTGGACCCATGCTCTGGAGTATGTACAAAGTCAAAGGCTCAAAGATA	CACCCGTA	AAGAGC	QRDSFNCGVHICLM
CATTCTGACCGTTGAAAGCTCCATCACATATTTTTCCCGATATGCG	CAGAAAGC	AGAAGA	TNSIMAGGKWHSEE
AGGAGGACCATGGACACTGTTGGTATTGAAAGACAATTCATACTATTA	GTCTCCTT	AATCTT	DVRNFRKRLKTLQE
TGCAAAACAGTCTGCCAAAGAGCCAGTGGCCCGGTGAGGACTTCATC	CTCAAACG	CTTTCTG	EGELYSVNSLGIPIQ
AACGACTCAAAACGGGCTCGGAAGGAGTTTAAGGTGCAAGTACCTCTTC	CTTTCTGT	ACAAGG	APTEQMDYKETRCK
AAAGAGACTCCTTTAACTGTGGAGTGCACATCTGTCTAATGACCAACTC	ACTTTTCT	TCAGAA	RSYASVLTQISPPAKR
GATTATGGCAGGAGGCAATGGCACTCTGAAGAAGACGTCAGAAACTT	TATATTTTC	GGAAGT	PDCKPDNNIFVPTKD
CAGAAAAAGACTGAAGAAGACACTCCAGGAAGAAAGGCTATGAGCTTTA	ATTAACAT	CCTGTTT	CAAEGNPQEKGRNE
CTCGGTCAATAGTCTGGGTATACCATTCCTCAAGCCCCAACGACTGAGCAA	ATTTTCTCT	TTGAGG	SPEEINTEHIVVAGKP
ATGGACTACAAAGAAACAAGATGCAAAAGAAAGCTATGCCAGTGTCTTA	GTTTATAC	CGTCCAT	ANNISPRCSTSEML
CTCAAAATAAGCCCGGCCAAAAGGCGGACTGCAAACTGACAAACA	TAACTTAA	CCCGGG	FEMVKATTSSGRSSL
ACATATTCTGACCAACCAAGGATGTGCTGCCGAAGGTAACCCGCGAGGA	CCTCCATT	CGTCAT	GTMTQDEFIRSTIAE
AAAAGGCCGAAATGAATCTCTGAAGAGATCAATACGGAACATATCGTC	GTCAATTA	AGGAGA	AVPLMSIKLPPMELP
GTGCGAGGAAACCTGCAAAACAACATCAGTCCAAAGGTGTCGGAGCACC	CTAACTAA	GATCAG	RKILPIPPRKPTQTN
TCGGAATGCTGTTGAGATGGTGAAGCCACCAACGACGATGGAAGA	CTTGATCA	ATGCAC	GGQKGKQQRVPTGK
AGCAGCTTGGGCACCATGACGCGAGGATGAGTTTATCCGAAACCAACACA	ACGGATTT	CTTCTAG	PDTLNKVRNWFNN
ATCGCCGAGGCGAGTCCCTAATGAGCATAAATCTCCCAACCAATGGAGT	CG (SEQ ID	CAGGAG	QLESYAMEGRSFQRL
TGCCAAGGAAATCTGCCACCAATTCCTCCAGAAACCAACCCCAAC	NO: 1151)	CTAGAA	EWLTEVLTAIQKAA
CAATGGAGGTCAAAGGGAAAGCAACAGAGGGTGCCTACAGGAAAC		GGGCTG	AGDEGIVDIICKRNPP
CAGACACCTAAATGCTAAAGTCCGGAAGTGGTTCAACAAACCACTTGA		CCCTGTC	LEVAKGEMCTQTEN
GTGATGCGATGGAGGGTGCAGCTTCCAAACGACTGGAATGGCTGAC		TTGAGA	KRKTNNAAARIADPI
GGAAGTACTACTGCTCGATACAAAAGCAGCAGCAGGATGATGAAGG		TCCCCAC	QSSKGAGDVKASYW
AATAGTTGATATTTTGCAACGGAAACCCGCTTGAAGTTGCGAAG		GGGGGT	KERARTYNRIIGSKKE
GGTGAATGTGACCCAGACCGGAAACCAAAAGGAAACGACCAACAAT		CAATAG	LCKIPIDQLEDFFKKST
GCAGCAAGAAATTGCGGACCCCAATCCAGAGCAGCAAGGGAGCTGGTAT		ACGGGA	SRTNVQESIMIKESS
GTGAAGGCATCGTACTGGAAGAAAGGGCTCGCACTTACAACAGGATT		GGGGCT	KIPALKIGNWWMIEKFF
ATTGGTAGCAAGGAGGAACCTGCAAAATTCCTATCGATCAACTGGAGG		GCTGGC	IGKEVAFALRKTKDTA
ATTTCTTCAAGAAATCCACGTCCCGCACCAACGTGCAGGAGTGCATCAT		TTTTCTT	QGADGLRYHHHLQWF
GAAGGAGAAAGCTCCAAAATTCCTGCTCTCAAGATAGGTAAGTGGATG		TTTTAAG	DPSGELLAKVYNECQ
GAGAAGAAGTTTATCGGAAAGGAGGTGGCGTTTCGCTCTCGGGAACA		AGGAAG	RHRKIPKHWKEAETIL
AAAGACACCGCGAGGCTGACAGCGGACTGCGATACCAACCACTTCAA		CACCAA	LFKNGDQSKPENWR
TGGTTTGATCCCAAGTGGTGAATTTGGCGAAGGTATATAACGAGTGCC		TCCGGA	PISLMPVIYKLYSSLW

174

AATGAAGATTGATGAAGAACTCCAGGATTCAGCAGACTCCGTCCAGAT ATCTGCCTCAAAAGCCCTGATGAAAAACAAATCATCTTGGCAGATGTG CATGCCCATATGAGCATGAGTAGAAGCGATGGAAGGAGCTGGCAG GCAAAAATCGACAAATACGAGACGGGATTCGCCACCTCGCGAAATCG GGAACCAAGCTGACCGTCTTCCGATTATAATCGGTCACCTGGATCAT GGTGAAACCGACAGGTGACAGTCTCAAGGAATTGGGAATCAAGGGA AGCGTGATCAACAGTGCCATTCAGAACTCTGTCTACTGTCTCGAACA CAGTAAGAAATACGTACTGGAATCAATCTTCGGTGAAGCGTACATACCA AATCCAATGCGAAACGGACACGCAAAACCTGCTGGAAATGGATGGAAA AAGGAAAGATTGCAGAAAGCCCTGTGAGGCTACCAACTAGCCTCCA GGCACGCCGACGCCAAAAGTCTCTGGCATAACTCTGCAAAATAACATCA AACGTCAATCAACTCCACAAACTCTCACTCTCTTCAAGTCTTCTCGGTGC TTCCAACACCACAATGGTGAAAGCTCTTCACTTTTCCCTCCAAAATTCT TCCCATGTGGGGAAGTCCTGTTCTTGAAGCTCTCCGGAGGCTGCAAGA GCAGAAGAAATCTTCTTCTGACAAAGGTCAGAAAGGAAAGTCCTGTCT GAGGCGTCCATCCCGGGCGTCATAGGAGAGATCAGATGCACCTTCTAG CAGGAGCTAGAAAGGCTGCCCTGTCTTGAGATCCCCACGGGGTCAAT AGACGGGAGGGGCTGCTGGCTTCTCTTTTAAAGGAAGCACCAATCC GGAGATCCTTAGGGGTCAAAGGATTAAAGGCAGCAGGTCCAATCTC CTCACTGACTTCGGTCAGAGAGGAGTCCCGCTTGAGACCTCCCGGG GAGTTGCTGAAGAGCGGAAGCTCTCTAGCAAGAGCTAGAGGGA GTTCCAGTCTGAAACCTTGCGGTGATGATGGAATGGAAGAGTACT TCGGTACTGCTCGTTGCTCTCTGCGTTTACTGCCGAGGGCCGATTT GCTCGAATCGGAAAGGTCTCAATCGACCATTCAGATGACGGCTTATC TAAGGTCCGAAAGCAGTTGGGAGAGTAACGTGTTCTCTACCTTTCAAG TTGAATGGTCGTTTACTGTTTGGATAGCTGACTGATGCTAGTACGCT TCATCTGTGGATGACGCTCCCAAGCAGTCAAGTAGACTTGAAAGGTGC CCTCGCCCTAGTTAGCTCTTAGACCTTATGGTCCGATGGTTGTGGACG GGTATGCTTCCGGAGCCGAGTCTGTTCTTAGAACCAACCTCGACGA GGCGAAAGCTTGCAAGTTAGCAAAATGTGGTAGGGCCGACTAGAA AATGAGTCCCTTAGGGGTTACGCTTGGCGAAAGTGAGGACAAATTGG CATTGACGGGTGCTTCGGCACTAGGCAAAAGGCGCCACCACTGTCCAA TCTCTAAAAAGTTCACTTATCGAAAGAACTACCGGAAACCAACACACAT GTGTTGAAACCTACACGGTGGAAAGGAAAGGAAAGCTTCGCTGGAACG AAAAGAACGGATAGGTTCCCTTCTTATGGCTGTGAGGCTTAGGATGG ACGGGAAGGCCGTGAGGCTCAGGCGGGTAACCTCGGCCAGACGCTAGT TGATCTTCGGATCACGACAGCCCTGGCTAAGAGGAAACCTGGATGGAG TGTGAAGGATGGCGGGTAGGGGGTTAAGCTGTTGACAGACCACCGA	TTTGCTC GAATCG CGAAAG GTCTCA ATCGAC CAITCAA GATGAC GGCTTA TCTAAG GTCCGA AAGCAG TTGGGA GAGTAA CGTGTT CTCCTAC CTTTCAA GTTGAA TGGTGG TTTTACT GTTTGG GATAGC TGACTT GATGCT AGTAGC CTTCATC TGTGGA TGACGC TCCCAA GCAGTC AAGTAG ACTTGA AAGGTG CCCTCGC CCTAGTT AGCTCTT AGACCT TATGGG TCGCCAT	LPGFSRLRPDICKSP DEKQIILADVACPYEH GVEAMERSWQAKID KYETGFAHLRKSATKL TVLPIIIGSLGSWWKP TGDSLKELGIKSVIN SAIPELCATVLEHSKN TYWNHIFGEAYIPNP MRNGHAKPAGNGW KKERLQKAPVRPTN (SEQ ID NO: 1396)

	GGTGT GGACGG GTATGC TTGCCG GAGCCG AGTCGT GTTTCT AGAACC AACCTC GACGAG GCGAAA GCTTGC ACAAAGT TAGCAC AATTGT GGTAGG GCCGAC TAGAAA ATGAGT CCCTTAG GGGGT ACGCCTT GGCGAA AGTGAG GACAA TGGCAT TGACGG GTGCTT CGGCAC TAGGCA AAGGCG CCACCA CACTGTC CAATCTC TAAAAA GTTTAC ATTCATC GAAGAA		CTGCAGTCACAAAATCAGTGATTATGCGGGTGGACCAATCTGTTGGCGG GTGTTCCCTCTACCTGACCCCGCAATATGGTATGTACGATCCTCGGATC TAAAAATTCATAATGGCCCAACCAACCAACCTCCCTAGCAGCTGGTG GTCCCGATAATTCGGGTCTTGCCCACTACTGCGACCCAGGCTCGCC (SEQ ID NO: 1028)				

177

178

CRE	Cn1	.	Cryptococcus neoformans	CCCTCTTAATACCCCATAAACACATAACAACACCCCTTAATCAACGTTCTCTGC ACCTTAAACACCACCAACATGTCCTGTCAGAGGGCCAAACACGCCCGTG GAGATCCTGTCGGTGCAACCTATGCTCTGCCGACTATAGGGACCTCAA AGATCATCTCAATAAACAAACATCCACCATTCTTCGTCCTCCGACCT CCGTGGCTCTCCCTAGTCGCTTGCCCTCGCTGCGGACACCCCTGCTCAG CTGGCACTGGTTATCTCGTCACCAGAGCCGGTATTGGGTCACCCGCT CCTCGAATCCGCCGAAATCGCTGGGAAACTCAACAAACACATCTCGCT GCCCTCCCTCAATACTGAGCTTCACCATCGTTCTTCGCTCCGAAACC AAGCCCAAGCCCCCTCAGCTGCTGAAGTTGTTGCCAGTCTCGAACC ATTGCTGAAGCCGAGGAGGTGCTGGAGGTCGCCAGGTGATGCCGA GACTGTTGACACGCTGGAAAGGACCCGGAGAGCTCGGAAATCCGTTCC GAGATCTGCCGAGGAAGTAGCACGCGAGTTAGGGAGCTAAACATGAC AGCGCGGAGGAGGAGCATGCTGGGAGGAGGAGAGTAGTCATACCA ACCAACTGCCCGACGAGGCTCGAAGACGCGGTGAGCTCAAGCTGG GGCTTCCCTGGGACGTTGCCCTCTTACTTCCGTCCTCCAGAGTGCT AACGAAAGATTCTGTACTTGCACCTGCTGCTTCCGGAGCAAGCCTC TGCCCAACAACTAGTTACCGACTTCATGGACGCGCTGAGCGTTGTGC TCTTGCTACATTGCAAACTCCGACTCTACACTGCTGGCATTTCTCG CCCTTCAAAGGTGCGCTCACCCAGGCGCTCGCTCCAGAACAGCCCT CAGGCCGTCAACCTTCTTAAGCAGTTCCCGCATATCCCTGGCCAGAAC AGCCACCCGCTGCTGCTCTCCAGCAATATTCGTCAGACACCAACCAAA CAAGTCATCAAACTCGTTGAGAAATGGCGCCTAGGTGCGGCAGAGAG GTGTTGGAGGAGGATGCTTCAGTAGCCGAACTCGATCAAGGGGTATC GACCAGCTCATCAACCAAGCACCCCAAGGGCGCTCTTGCTCCATTGGCA ATGCAGTGGGTCCAACCTCTGGTAAAGCTCCCGACATCGACACCATCCA AAAGGCCCTCGACTCTTCAAGCCCGACACAGCACCCCGGCTTAGTGGC TGGTCAGTCCCTCTCTTGAAGACGGCTGCCAAGAGGGAGCCGGTCAAG CAGTTTCCAACTCTCTGCGCCGCATCGCCCAACAAACACCCGCCCTGG TCGCTATGCTCCGACTCTCTGCTCATCCCTTGAAGAAAGGACGATG GCTCTATCCGACCTATCGCTGTTGGTGAACCTATCTATCGGCTGTGCG AAAGCTCTCATCTCGCATTTCCAACCCGACTTCTCTCCCGTTCCAG CTCGGGTCAAGTCAAATCGGTGGTGTAGAGCCGATCGTGAGGCTGACA GAGAGACTTGGAGGGTCTGCCGGCGCTGAGTTCTCTTTTAGCCT CGCTCGATGCTTCAACGCTTCAACCGTGTAGATAGGGCCGAGATGGC AGCAGCGGTCAAGACCCATCGCCCGACGCTTGGAGGACATGCAAAATG GGCTATGGCGACTCGTCCGACCTGTGTGTGTGACAAAAATCCTTCAA	CCCTCTTA ATACCCCA TAACACAT AACAAACC CCTAATCA ACGTTCTC TGCACCTT AAACACCA CCAAC (SEQ ID NO: 1152)	NO: 1274)	MSLQRAKNARGDPG RCNLCSADYRDLKDH LNKQHSHTFFVPSDL RGSSLVACPRCGTPC SAGTGLSRHQSRVCG LTAPRIRRNRVGNST NTSRCPSPNTAASPIV PSPSPERPSPQPAE VVASLEPLSEAEVLE VAQVDAETVDLTLEGT RRAPESVPRSAAEES TRVRELNMTAPEEEH RGEESHTNPTAPA GLENVVSSTLGPSPG TLPSSLPSQECANERF LYLAHLPVRSKPLPN NLVTFMDAAERCA LAVIAQSPDSTLLAFL ALPKVGLTQALAPEQ PLRPSTFLKQFPHIPW PEQPPARRPPSNIRP DITTKQVIKLVENGRL GAAERVLEEDASVAE LDQGVIDQLITKHPK GPSCPFGNVAVGPTP GKAPDIDITQKALDSF KPDTPAGVSGWSVP LLKTAARKREPVKQFL QLLCAAIANNNTAPGR SMLRTSRLIPLKKDDG SIRPIAVGELIYRLCAK ALIISHFQPDFLLPFQL GVKSIGGVEPIVRLTE RVLEGSAGAEFSFLAS LDASNAFNVRDRAE MAAAVKTHTLWR

				TCCTCTCAAGGTGTTGACAGGGTGACCCCTTTGGCCCTCTCTCTCTC GATCACCTCCGACCAACCTTGAATGCCCTCAGTCAATCGCTAGGTCGT CTACGCAAGCACTCGTTACCTCGATGACATCTACCTCTCTCAAAAGCAG TCGCAAGTCTCAGCAAACTACCAATTCCTCGCCGACAGCAGCACA TCATCAAGCTCAATGAAAAGAAATGCAAGTTAATCAGCTTCGATGAGAT CAGGAGGAGGGCTTCAAGATGCTAGGACGATGGTAGGAGGTAAAGG AGAAGCAGCGGAGTTTCTGAAGGCAGGATTCGGAAGGAAATGGCA AAGTGGGCAAGCTCAAGGATCTTCCACATCAACACGCGCTCCTCTAT TAGCTTCTGCAATCAGCAAAATCTACGACACCTGCAGAGAACTGCGG CTCGACGACCTTGAGACCTATGGAGAGGCTGGACACGATGCTATG GGAGAGGTGAAAAGGATGAGGATGAGGCAGCGGAGGATACGGCG GAAGAGGAGGCTAGGGAGATCGTTGACGAAGCTACCAGCGGACTG GGCGGACTAGGTCTACTTTCCTTCAAAGATGTAGCCCCCTTGCTTACCG CTCGCAGCGAGGCTCCGACACTCTCTCGATAACCTAGGTCTCCTTT CTTGGCAGAGGAACCTCCAACCTCGATCCCCCAACGAACCTGATGCGC AGAACTCTGGAAATCGCAACAGGAAGCCATCTACATAAACCCTGGCGAC ACTGAACGCAAGCACTACCGAGAATGCCTCCAGACTCGGCCGAAGTT GGTTATCAGTTATCCCTTACCTTCAACCCCTGCGCTTCCAATGTCGAG ATTGCTCTGGTCTCATGACCGCACCTGTCGGCTCTCGATCCCTGT CTGTGCTTCTGTGGTGGACTCACCTTTGGGTACGACGAGCTTTGCC GCGCCGCAACCCCTGGACCCAGCGCGGCACAATGCCATCAACCCGCT CATTATCAACACTCAAAACAAATTCAAAGGTGCCACGGTTGAGATTGAG CCCCACGCTGTCGGGACAAAGGAGAAACGACCTTCGGGTCAGAGGT TCCAGGCTCTGGCTTCACTGACTACGACTGAAGGTTTACTCCCTCGG GGACCGAGACGCGAGAAACCGTCAACCCCTGCGCCCAACGCGCAA GCTGGCCGACTTCTGCTTGACCCGGTGCCTGAACCTGGCTCGACAAGGT GGGTGAGGTGCTCTAAGAACGCTCCGAAGGTCACTGGTGGGGTCTTT AAACCAATCATCTTCCACTGGTGGCTTGATGAGCAGGAGCACAGCAG ACGAATGGAAGGACTGGAGGGACGCGATGCCGGTGGGGGGGTTTCGAG AAAAATGGAGAAACGGATTGGTGTGAGTTAGTAAAGGCAAGGGCGAG GACGCTGGTCTTATGAGGAAGAGGAGGTTGGATTATTTTCTTTCTTT AATAAGTTGTTTATTAAGTAGTTTCTTTCATTTCGGGCAACCCACACGAC AACCCCAATAAATTAACAAACGAAAAATGCAACCTCTATAACCC (SEQ ID NO: 1029)				
CRE	CRE- 12_CC ri	.	Chondr us crispus	ACGCCCTATCCATTTCTGCCAGCCTCCCATCGGCTCGCCGTCTCCGCA ACCCCTCTCTCGGCTGTACCAAGTTCGGCTCCACAACTCCCTCGCCAC AATGAGTCAACCTAATATTCTGTCGCTGAGACCCCTTGTCTCAGCTGC	ACGCCCC TATCCATT TCTGCCAG	TAAGTC CTTGAC GCCTGC		

CCACGCTGTTCCACCCCGCTTCTCCCTCCAATCCCTCTCTCTCTCTCC TACTGTGGTGACCTCTCTCTGCCCCATACGCTCTCTCCATGTTTACTC ATCCATCCCTTCTCGTGTACACAGTTTACAGATGCTCTCATCAAGAC TGTCGCGCTGCTGCGGCAATGACCCCACTGAATCMCATCGCGCA TTCATTCTACACATTTCTCCATCGCTGCTTGGACGCTGTTCCGT GGAGAACCGGATGGGTTCCGCACAGGCAACATCATGCCCTCCGC CAACGATTCCCGCGCTCTCGGTCGGCATTTGGCTGCTTGTGGC ACGAAGCCTTGTGACATCAGTCGACCTTGACTACCGCACCGGTCA CAGCGTCGTTACCAAGCTCCGCTACATCGCGCACCGCATCGGCCGT GCCATGCTGTGGCGCGATGCCCAATATGGAGCTGCAATGTCGGCC TCAAGGCCAAACCGCTGCCGATCTACATGCTGCCGCCACCGCGACAC ACTCACGCGCTTACCCCTCTCTCTGCCAGCCGGTTCAGCTCTCTCACC GACTGACCTCCCGCGTCCCTGAAATACGGAAGGTCAAGTCTCTCCGA GCGGACGCGCCCTTAATCCACATCCGCTGCAGGACCGGACCATCTCT CTCCCGTATCTGCAGCTCTTAGCCCGCACCATATCAGCCCGGAAGCT GGGGTTACGGGTTGTCGGCATTGACGAACCTGGTTGACGCTCTCGCC GAGGTGACATTCGGGATAGAACTGGCTCTCTCTGCTGCTGCCACTCT GATCCCTCCAAACCCCGCTCACAAAATACGGCCGATTGCTGTAGGG CAGGCTTGGCGCTGTGTACGAAGTCTCTTCTGCCCGCCCATCCA GGACCCCGCATCACCTTCTCCAGAACAGCTCGCCAACTCGGTTGCC TCGGCATGGACGCAATCGTCCATGACACGCGCATGCTTATGCATCGTC ACGGTCAAAACCCAGACTACATCATGCTCCGTAGACGCGGAATGC CTTAACACCTTCTACGTCAGTCCCTGTGGATGCTCTCTCTGCAGAC TCCTCCCTGCGCGTTTCTCAATCAATCTATGCGCCGACCGTCTCTGA TCTCGTGCTGCCCTCTCTCGCGGTTCTGATGAAAAGTCAGGAGGGC ACCAACAGGGGACCGGCAAGTATGCTCTTATTTTCGCTGGCAATCC AGCGCTCTGCTGCTCTACCCGCGAGTGCCTCTCGACCTGAACCG CTGTACGCGGATGACGGCACTGTGTCGGGCAATCTCGGAGGTCTAT CAAGGCACTCCGAATTTCTGTGATGACGGCCGCGAGTCCGGATTTCAC GTCAATATCAACAAGTGGCGGCACTGCGCGACCGTAATGCCAGAA AGTTGTCCGAATTGCTCGTATCTTCCCTTACGTCGAGTGGCGGCA GGCGGTGCTGCTGCTGGTGCCCGCTCGGCACAGATGCTTGTGTC GCCGACATCTCATGAACAAGTTCAATCATGCCATGCCCTCCCTCAGCCTC CTTGATGAAATCCCGACGCGGTACGCGGTTTTCACCTTACACCTCCCGC AGGCTCGGTATGCAAAAGTCGAGCATGTTTTCGCTTACACCTCCCGC TCTCCCTCCAGCAGCTACAAAATTTGATGAACAACAAATCGCTGCTTAT TCTCGGTTGAATGATGTGGCGGTTTCCACATCCATCGGCTACACAAATAG GCCTCCGTTTCCGCTCGGTGGACACGGCTTACCCCACTGTCAACCATTC	CCTCCCAT CGGCTCGC CGTCTCCG CAACCCCT CTTCTCTG GCTGTACC AGTTCCGC TCCACAA CCTCCCTC GCCACA (SEQ ID NO: 1153)	CCCGTG ATACAG CATCGG TACCCCT AGCATTT GAATAA AAAA (SEQ ID NO: 1276)	VYSSPSCLHSFTML LIKTVRAASATMTPT ESHRAFIHLHPIAVL RRSFRGTGWSRT GQHHLRQIRASS GRHWAALWHEALA AHQVDLDYRTRHSR RYQASATSRHRIGRA MRLAADAQYGRAM SALKAKPLDLHAAA TRDTLTALHPPASP VQPLSPDLPPVPEIT EGQVLRAARALNPTS AAGPDHLSPRILQLA RTTISPEAGVTGLSAL TNLVRRLARGDIPDR TAPLAAATLIPLQR PHKIRPIAVGQALRRL VTKVLLPPAIQDTRD HLLPEQLANSVASG MDAIVHDTMLMH RHGRNPDYIMVSV ARNAFNTSRQSLD RLPQTPLARFLNIY GRTVPDLVLPSSPRFL MKSQEGTQGGDPAS MLFSLAIQPLRLT RECRDLNRWYADD GTLVGPISEVIKALRIL RDDGPQSGFHVNIIN KCRAYWPTVMPEKL SELLRIFLHVECEG GVALLGAPLGTDAFV RRHLMNKVQSCHAS LSLLDEIPDARTFHL HRVTGSVCKVEHVFR LTPPHLSLPAATKFE QQIAAYSRLNDVAVS

CRE	CRE-13_Cri	.	Chondrus crispus	ATCCATGCGTCTACGCTGCCAGTTTAATTGAGGGCGGCACCTGTTCTGTG GAAGGGCCACATAACCCCTCCGAGTCGTTTTATCGCCGCATGGCCCGT CGTCATATCGTCCAGTACTAGGGCCCTTGAACCTGAGGTCCGACCC GAGGCATTCTGGGACCACTTCTCCCTCGGACCACTTTGAACAGAGGC CCTTTGTAGACCTGAACGCTACACACACACATTAATCAGGCTATGC AGGGGGCACTTCTGACTACTGGGAACACACCGCGTGGACCTTGA CCCTCTCCCTCGAACACAGTCCGCTCTGTCGCGGACGTGCCCGT ACAAATCCCTCGTCTCGGGCGCGGCTGTTTCTATGAGCCACCCC TCACTCACTTCTGAGTCCCTCTGCGGTGGTCTGCTGATGACGCCG GCATCTGGATACACCGTCTACTGTGACTCTATTGCGCTCTCATATGTT CGCATGCTGTAAAGCCAATGACGCTCGCGGTGATGCTGCAATATG CCGTATGGCTTCGGGCTGTTACCGTCAACACTGTACGCAACCTAC TCGCCGTCACGCTTCGCGCGCGGCTCTCTGCTGCGACCTTGAGGT CCCTTCTCTTGCCGAATACCGGAACCGCCCGGATATTCTCGTCC AGCCGCGCGGCTCTTCGGGGCTCTCCCGGACCGCCCACTGCGTA CGACGTAACCGTTCGTTCCCTCTACTGTGCTCTACAATGTCTCTGCTG CGAAAGGCTCGGGGTGCAGCGGAAGTGTGATTTGGACAAGCTTC GGTCCATTCCCGTACAGTGGTACGCAATTCACCTCCAGCTGACTCC CCACTCCCTACTGACTGGCACTTTGTCGCTCGCATTTGATACTCTC GGCGGACCACTCTCGACGATGCGAGTCTTGAGTACCTCGCTCACC GCATTGCCAACCGGACATATTCTACGGGACCGCCCAAGATACGTCT ACTACAACGATCAGTTTCGCTGTTTGGTCCAGTTTGGCTCTGCCACC TTTCCGATGCCCCTATACGGGCGGCTATCGAGCCCCCGCCCAAGT GTAAGTCTTGACGCTGCCCCGTGATACAGCATCGGTACCCCTAGCAT TGAATAAAAAA (SEQ ID NO: 1030)	CNCCAGCC AMCGATC CCGCGGCC ACTCGCM GCCCGGCC GTCTCGAC CGCACCT CCCCGASG CCCCAGCC CATC (SEQ	TAATTCA CCTTCAT ATCTGCT AGTGTC TCTGTAA GCGCAC CCCTCAT GCATTG ATAAAA TTACCCC	TSMATQJGLPRLGG HGFTPLSPFIHAYAA SLIEAAPVRVKGPHN PSEFYRRMARRHIV HVLGALNPEVTRGI LGTHSLPGPFEALL SRPERVHHTLIQAM QGATSRLYWEHTAW DLPLPRNHSAASVR RRARYNSLRAPGAAS FLCSHPSLTSRVPSAV WSCMLRRHLDTPVY CDSIRPLICSHCKPM DARGDHAAICRHGF GVVHRHNTVRNLLA RHAFRAAGLCCDLEV PSLLPNTANRPADILV QPAPPPSGALPDRPT AYDVTVRSPLYCRSTM SLAAKGLAGAAEAD LDKLRVHSRTVRDAF HLQDPSLPLLDWHF VPLAFDTLGATSRT MAVLEYLAHRIANRT YSSYGTAKIRLLQRISF AVWSSLASATLSRMP YHGAALSSPAQV (SEQ ID NO: 1398)

CCASAAAGCGCAAGGGCGCCCTACCAAGCGCCCTCTGCCTCTCAGGG CCCAGCCGCTCTCCCATGACKAATGACCCSCGGACGCCACCTCMA CCCCGCGCCGCTCCCTACAGGTTCCCTCCCCWACTCGACACGA CTMTTCGCCCTCTCWTMAGCCTACCCCAACCCGCTCCCGGGCC ACCTCMCTTKCCGTCGTTGAGGCACTCTTCTCCCGGATCGGCC STCAGACATTTCTCCATTCCGGGACCGGCTTCATAGCACTGATTAC TTCTATCCGCTCGTCCGCGCTACAGCCCAACACTCCGAAACC ACCACACTGMATCTTGCACTTCACTGCTCCGACTGCCSTTCTCGA GAGGCTTCCGKGGGAGSCTGGTGAGGTCTCGCGGGTCAACTTC ATGCTCTCCGTTGCKGATACGGAGAGCGTGTACKGGACGAGAGTGG GWCTCTTATGKAAWGAAGCMCTAGATGCCACAGCKCCAGGACAGAA TGGACGACACGCGATGCCGGGCGCTCGCCCTGCAAGCTCAATACGCGG CACGTGCGCGCGCTATGCGCTTGCTCCCAAGCTCAATACGCGCG CGCATGCGCACATTTACCAACCCCTCTAGCTGACCTCAACGACCCGG CCACGATGGAGCGGCTCCAAAGCCTTACCCCACTCTACCGTCCCGTC GTGCCCTGCCACCTCCGACAGCTCGACCCCAAGTCAACCGMGG AGCGGCTGWTGCGGTTGCTGCGCTCAATCCGAAGTCAACGCGCG GCCCTGATCGCATGTCGCCGAAATGCTTCACTCTCGCTCACACTCCC ATAAGCCAGAAAGCGGCTCACCGGCTCTCKGGCTAACCAACCTCG TCAGCCGCTGCTCGGCTCCCTCCACCTGTACGATCCCACTGGCC AGTGGGCGACACTCTKCGTTGAGCCCGACCGGGAATAACGCC CGATCGCTATTGGCAAGCCTWCGCGGCTGTACAAAATMCTTCT TCCTGCGCCATCGACGACTGTCGGGACCACTGCTCCGAAACAMTG GCMACGGCATACCMACGGCATTAAGACCCACACTACTAATGGTGTCTA ATGCTAGTACGAGCCCGTCAATAATTTCTACGSCAACAACTCTGAC TTGAGCTTCCATGCTTCAATATTTCTACGSCAACAACTCTGAC CAGTGCCCACTCGAGCACCCTGCTCTACGATATTTGGATATGGTGA CGACGCGCCCTCCCTCGTCTGCTTCACTSCCGCTACCATCT CCACGCGGGGAMGGATCAACAAAGGGGACCTGCAAGCATGCTCT TTTCTGCTTCCCTCCAGCGCTCACGCGCTCATTTACGTGAGTGTG AMCTWKAATGAACGCTGGTATCGGACGACGGAACATCATTTGGAC GGATTGACGAAGTTKCAAAAGCCTTGATATCATCACTAAAGAGGGCC CAGGTTCCAAITCTTCTCAACCTTCGAAGACACGCTCTCTGGCCAA GCAGGACGCMAGACCTCTCAGCCCGCTCATGACAGTGGTCTCTGCG GMGTATCGATGAAGCGGTGTGCTGCTCGGCGCCCATMGGG TCACCAAGCTWTATGGCACAGTACATTCGGGAAAWTTGAACACTTGC AAAACCGCMCTCGCCCATCTCGACCATATCCCGAGGCGCGCATGCGCT TTCACCTGCATCGGGTGTCTGCTCTGCTGCTGCGGCTTGACGACCTCTC	ID NO: 1154)	CCA (SEQ ID NO: 1277)	SASQGPAAASHDXPR TPPXPPRPPVRFPP PTLDQHXFALSAXP HPXRRPPSPXRRLR HSFPRFGXQTFSSIP GPRLHSTVLLRLVR AATAANTPETTLXS CTFTCSRLPFFERPSX AXLAGGPRAVNFML SACXYGERVXDESGX SYXXKXHCITSHPRP RRYSRQHSHRNHHTXF LHLHLLPTAXLREAFR GEXGWRSSRSQLHA LRLXIRRACTGREWG LLXXEALDAHSXRTE WQHTHARRPSPVVS PSARAAARMRLASQ AQYGRAMRTFTNPP LADLNDPATMERLQ ALHPTPTV/VVPLPPS AQPRPPEVTEAVXR AVRRLNPNSSAAGPD RMSPKLLHLLAHTPIS PEAGVTGLSALTNLV SRLARGSLPPCTIPLA SAATLLPLQPRPGKIR PIAIGQALRRLLVTKXL LPAIDDCRDHLAPE QXANGIPNGIDAIVH DARMLVRRHGNDDP HYXMV/SIDASNAFN NFSRQVLDQLPTRA PSLSRYLDMVYARAP SPLVLPSPPTILHSRX GSQQGDPASMLLFLSL ALQLTRLISRECLX MNRWYADDGTIIGRI

CGTTGGTCCCCCGGATTTCGGWGTGCCGTTTGCAACAATTCGACC GTGACCAACTCGMAGCCTATGMCGCTTTAATAGTGTGACTATGTCGC CAAGAACTGTCGCCAATACGGCTGCTTTTCMCACGGWGGCCACGG CCTCACCTCATTGGCATCTACCATACAGCCTCWTACGCTGCTAGCCTCA TCGATACCGCTCCAGCAGGCTACAAGTCCCACTTTCCGCGCTCTCT CAGTATCAGCGTTTTCACGAGGCCGTTGCGGGTCTTTCGAAATTT ACCTTCATTGCTKCAACCCGCACACTCTCGATGACGGAACGGACCTC GGMTGCCTTGAACCAKCTGCGCTACTGGCGGACCTGAACGCATACAC ACCTTCTACTCAGGCGCAATACAGTGACGAGCAGWAGCTGCTACTGGC AAWACCCCTCTGGAGTCTTCCCAACCTGCTGATCAGCAGCGCAGC CTCGCTACGCAACGAGTACGCTACAACTCCCTGCTTGCCCGWGGGCC ACCAGTTTCTCACTGCACACCCSGCCGACCTCTCGGGTCCACAACGC AACTTGGTCCACGATGTTACGTGCGCACCTCGACGCCGCCCTGACCAAC GATTCATATCGCGTTGCGATGKCTCACTGCTCCAAAGCCTATGGATGC CCGCGGACACCGGWCATTKGACGCCACGGGTTTGGTACGTTGCA CCGGCATAACACCGTCAGGAACGCTCTGCCMGGCAGTTATTCGGMGT CGCTGGCTCGCCTACTCGCTCGAAGTACCTTTCTGATTCCTCAACACCG CCGCCGTCGCCGAGATATCTCGTCAACACCTCTCCAGCCCTGGC CTACCTCCTGACMAACCCACAGCCTATGATGTACGATTTGTAGCCCTTT TCGCCGCGGAATGTTATACCATGCCGCCGTACCGCGGGAGCCGCC GACGCCGATCTGTAGGAAGMCAAGCCCTCGAGCGCACTATCCGC MACGCTCTCTTATCGAGGACGACAATCSTCCGCCCTCTTGACTGGC ACTTCAACCGCTTCTTCGACGCTGCTGGMGCCCTCTCAGTCTAC TGACACGTTATCGAAGATCAGCTAAGCTCATGGCCCTCCGCAACTCGT GCACMATTCGCACTGCCAATACGGCATCCAACACGCTCAGCTTTGC TATATGGTCCAGTGTCCCGCTATCTCTCTCGCTACCGACACAG CCGCGGACATCTCATACCCGATAGAAATTAATTCACCTTCATATCTGT AGTGTCTCTGTAGCGGACCCCTCATGCATTGATAAAATACCCCCCA (SEQ ID NO: 1031)			DEVXKALDIITKEGPR FQFLNPSKTRVFWP SRQDILLSPLMTVGP LRVIDEGGVXLLGAPI GSPSXMAQYIREXLN TCKTALAHLDHIPEAR MRFHLHRVSASACRL QHLFRLVPPDFAXPF AQQFDRDQLXAYXR FNSVTMSPRIVPKYG CXFTXATASPHWHL PYTPXTLLASSIPLQH GYKVPTFPSPSISVL HEARCGSFEIYLHSX NPHTSRCDDYVAKNR AQIRLXFSHGGHGLT SLASTIHASYAASLIDT APARLQGPHEPAVS QYQRFARGPLRWLR NLPFVQPAHFSMTE XDLGCLPEXALLARPE RIHTFLLQAQYSAAS SYWQXPLWESFPNP GDHSAASLRKRVRYN SLLAPGATSFLTAPHA ATSRVHNATWSTML RRHLDAPVTNDSISPL RCXHCKPMDARGD HAXIXSHGFTLHRH NTVRNVLARQLFRVA GLAYSLEVFLIPNTA ARPADILVQPPPPAP GLPPDXPTAYDVTICS PFRRGMLYHAAHRH GGAADAASVRKXKA LERTIRXALLIEDDNX PPPLDWHFQPLSFDA LGAPSQSTVHVIEDH

CRE	CRE-1_ACa_s	.	Acantha moeba castella nii	TAACCTAACCTCTCCCTCGGGCCCTCTACCTAAAGCGCCCTAATCGA CCGGCAGCGCCTAATCGCTACCTCTACGCCCTAATCGACTTTGGCGCC AAAGCGACTTCCCGGCGATTTCTTCTGCTGCTTTTCTTTCTCTCCA AGCAGCGCCTTTACTTACCTGCGCGTCTGTTTTTCTTTCTCTTTGC ACTTCGTTCTACTCACACCTCTCTCTCTCTCGACCGCGCGGCC TCGAGCGACTTGCTGAGCGGCTCCGGCTGCTCTCAGTCTCCCGCTGCTA CTCCGCTTTTAGACGCCCGGCTTGTCTCAGTCTCCCGCTCGGAA GCGGTAGTGGGTAGTGCTCAAGTACTCAAGCTCTTTTCAGCTC GGCGTCTCTCAATCGGCTCAGTTAGCTTTCAAGTTGCTCGATTAC GCTCTGAATCGTCTCTCTCAGTCTCAGTCTCAGTCTCAATCTCGATC TTGCCTCGCCTTGTCTCGACGCTTGTCTGTGACCGCGACCGTTTGCAA GTGCCAGCTTTTGTGCGCTTGTCTGTGTCGACCGCGACCGTTTGCAA GCCCTGCGCTCGTACCGCGCTCTGTAGCGGTTCTCATCGCTGAAGCT TCTACGCGCTGGCAGCAAGCTCGGCCCTAGCTTTGTAGCGCGCGCGGTG GCCGCTGCCAAATGGCTACGACGACCATCTCACGATCCCTTCGTCTT CTTCTTCTTCTTCCGCTCGCTCGGGGCTGAGCTTCCACGTCGCGCT CAGTTGCTGATACCGGCTCTTCCGCGATGCGCGCTTCCACTGCGCT CTCGCCACTGTACAGCCGACGTCCACATGGCAGGACCTCTCCGCGC ACCTCACAGCATGCACGACGGTGACGTGCCCGGTGACGTGCGCGCGCG CTGCGGCATCGTGAGTGCTACACGAGGGGTGCCGCAAGTGGTTTCGC GGAGCTGAGGACTTGCTCTCATAGGGGCAAGGCCGCTACGCGCCCG CCACAGCCCCCGCGCGCTCGCGCTGCGCGCGCTGCGCGCGCGCG ATTCTCGGGTAGGACCCAGCCCCGACCCCTCGGTAGCCCCGCGCWA CGCGGTCCTCCGCGGAGCTGCGCCACGCGCGCGCGCGCGCGCGCG CCGTGCCGCTGCGTCCGCAACCGCGCGCGCTGCTGCTCTCTCCAC CTCAGCGTGACGTCCCGTGCTCCCGCGCCACGACCCCGCGCTCGCAG CCTTCCCGCGACCTGTTCTCGGGTTTCGGAACGCGCTACACGCGCATC GCCGCCCTCCACGCGGWCWTGCTGCGCCAGCAGGCTGCGCCATCCGCG TGCCAGACGCTTGTCTGCTGTGGCCAGCGCTACCGGCGCGCGCG CCGCGTGTAAAGGCCCAAGTGTGCGCGGTGCGCGCGCGCTTCGTCC CTAAGGCGGAGCGGAGCATACCTGAGGCGTCTCACCTGTGACGC CTCAGGACCGCGCTCTACGCGCGAAGACGCCGCGCGCGCGCTTC GTCCGCGCGGCGCTCGGCTAGCAGACGAACACAGGAGCATGACAC	TAACCCCTA ACCCTCTC CCTCGGCC CCTCTACC CTAAAGCG CCCTAATC GACCGGC GACGCCCT AATCGCTA CCCTCTAC GCCCTAAT CGACTTTG GCGCCAAA GCGACTTT CCCCGGCC GATTTTCT TCCTGCGT TTTTCTTT CTCTCCAA GCGACGC GCTTTTA CTTTGCCG CCGTTCTG TTTTTCTT TTCTCTTG CACTTCG TTCTACTTC ACACCTCC TCCTCCTC CTTCTCGA CCCGCGCG GCTCTGAG CGACTTGC	TAAGCC GCGCGA CGAGGA CGGCCA GGACGA CCAGGA CGACGG CGACGG CGACGG CGACCA CCTAGC ACCGCA CGCGCC ACGACA TATTGTC GCGCGC TGTACA GGCGGC TAGGTC GAGCCC AGCCGA CCGTTCT GAGCCT CAGTCG GCTTGA GCCCC GGCTTC CCAAGG CCTACC GGGGG GCTCTT TTCGCC TGGTTT TGCCCG	AKLMALRNSCTIATA KSRIQQRLSFAIWSSA AAAILSLPHTHAADIS YPIEV (SEQ ID NO: 1399)

GTACGGCGGTGACACAAATCGCGCTCACTGCCCGCCACGCGCCCGCTGAG ACCGCGCCCCCTTCGAGTTCGAGGCGTGTTCCTCGAGGAGGAAGCCC CAGCCACCGCGGACCTCCCGCTTACGCGCGCCCTTCCTCGCTTGC CCGTAGCCCCGTCTCAGGAGATCCGCGCCGCTCAAGTCCGCGTGGC AGCCCGCCCAAGACCATCGGGAAGCCGCGTGATTGCCACACCG CGGGGACACGAGGGCTACACGCCCACTTGGGCTCTTATCGAGCT GCCGGCGGGGTGGAGTGCCACCAACTGTCGAGGCGCCGCCG CACCAAGCTTCAACGAGAGCGCTGCTCGACATCGCCGCCGGAAGGAT CCCGGCATCCCGACCCCGTGCAGACCCCGGGCGCGACGACGCG CTACGCGGTTCCCGTATCGGGACGACACGCGGCGACGTCAGCAA CGACGACACGCGAGGTGTGCAGATCGGCCCGCGGCCACCGCCAG CGCCGGCAAGCCAAACGGCTAGTGAGCAAGGCTCTCTCCCGAGC CCTCGTGCCCTTGAACGGGCGAGCCCGGCTGCGCTCGGCAGACAC CCTCGGCGTCTCGAGGCGTCCACCCGCTAACCCACCGACAGAGGA CTATGCCCGCGCCCGAAGCGCGGATCCCGCGGGTACCGGCCAAG CACCTGGCCCAAGTGCCAAAGAGCTCCCGCGGGTAGTGCGCGGGT CCCTGGGCTGGACGTTCCCGGTTCTCATCGACATGGCGGCCACCGCAA CCACGGGCACCTACACTGGCGGGATTGCTACCGCGAGCCGCTTGT TCCGGGTGAAGAGCCCGACGGCGGTGTACGACCATCGCCGTAGGCGA GGCCCTCTATCGGTATCGCGCGCTTGTCTCAAGGCGGACAGGGTG ATGTCAGCGCCGACGCCACCAATATGTGGGCGGCCACCATATGGC GTGGCTACCCCGTGGGTTGAGGCCCCGTCACGCCCTCCCGCAA CTGCAGACAGCGGCGAGCTCCGAGCGGTGCTCGCTCGACTGGCGTA ACGGTTCAACTCGCTGACCGGTGCACACGGCCCTGCTCATCGCCGA CCGCGACCCGCTCTCGCGGACTCTACGAGTGGTCTACCGTGAGGAC TCAGTCTCGTGTGCGCGCGGTTGAAAGGCGGGGCTCCCGGCC TCCCTGCTCTCCAGCGCGGTGCGCCAGGGGACGCTCTGGGACCC TCTTCTCGCCATCGCGCTGCCCGGTCTCGACGAGATCGACGCCATA CCGTACGTGACCGCGAGCTACCTCGACGACATCTTCTGTCACGATAC CCACGGTGTACGGACGCGGACCAAGGCCCGCTCGCTGCCACCTT CGCTACGGCGGAACGCGAAGCGCGCGGCTGGCTTGGGCTCAACCG CTGCAAGTCGGCGGTGTGGCGCGGACGCAAGGCCCTCTTCCCCC CAGCGCGTGGCGCGGAGGAGCTCGAGAGCTGCGCACGAGTGGC CGAGGCGCTCAAAATCTCGCGCGCCGTTGGGCTCGCCGCTTCGTC GCCAAGTCGCTGACGGCATCATCAAGCGGCCCATCGGCACACTCGACC TCGTGCTGACGCCGAGCTACCGTGCAGCACAAAGCTGGTGTGCTACG GCAGTGCGTGCCAGATACCCACGTTCTGGGCGCGCGCTGCCGAC	TGCAGCGG CTCCCGGC CTCCCCCA CGCGGCT GCTACTCC CGTTTCT AGACGCC CCGCTCT GCTCTCAG TCTCCGC ATCGAAGC GGTAGTCG GGGTACGT GCTCAAGT GACTCAAG CCTCTTTC AGCCTCG CGCTCT CAATCCGC CTCAGTCT TAGCCTT CAAGTTGC TCGATTAC GCTCTCGA ATCGCTCT CTCTCTCA GTCTCAGT CTCAGTCT CAATCTCG ATCTTGC TTCGCTT CGTCTCGA CGCCTTGC TCTCGWA ATCGCTGC CACTACGT GCCAGCTT TTTCGTGC	CCTGTT TTTCTCT CCCCCT TTCCCC CTTTTC ATTGTA CTTAGTT TTTCTT CGCGC CGCAG CTTGTG CCCGC ATAGT TTAATAT GTTTAA AAAAAG TGTAAT TAAATA ACTGTT AACCCT AACCCT AACCCT AA (SEQ ID NO: 1278)	SSRALALERGEP ASADTLGRLEALHPP NPTDRGLWPGAPKA AIPRVTAHQAQVAK ELPGSAPGSGWTF ELVQAIDRQPTGT AAFLIDMAQRALGT LHWRGLLTASRLVAL KKPDGGVRPIAVGEA LYRVIGRLVLKADRV MSSADATQYVGRHQ YGVAYPGGVEAPVH AVRELHDSGQLRAV SLDWRNFAFNSLDRV HTALLIADRAPALARL YEWSYREDSVLVLP AFEKAGLPASLLSQA GVRQGDVGLPLFFAI GAAPVLEIDAIPVY PRAYLDDIFVTIPHV TDAATKAAVAATFAT AEREGAAAAGLNR KSAVWAADAEALLP PHAAAGAREDEVESCA PVREGILGAPVGPSP AFVAKSLDGIKRAIG TDLVADAELPLQHK LVLLRQCVAQIPTFW ARAVPDAGPALAVW DTALLRRTGALVGLD VRDGLQADIAARLPV RLGGLRSMKDTAP RAFVASILFAAALANT RRSELTCSASTARRLR AALPELARTDACNDE AAWRRSIARGVFPD VDKLTQQLQRLVQ GMADSKSAHRTRRQ

GCAGGCCGCGCCCTCGCCGTCGGGATACAGCGCTCCTCAGGCGCAGC GGCGCGTGGTGGGACTTGACGTGGGGACGGGTCCCTGCAAGCCGAC ATCGCGCGCTGCCTGCCGCTGGCGGTCTCGGCTCCGTTCATGA AGGACAGCGCGCCGCTCGTGCCCTCGATCTGTTGCGCGCGC GCTCGCAACACGCGCGATCCGAGCTCAGTGGACGCGCAGTACGGC CCGACGCTCAGAGCGCTGCCGAGCTGGACGCGACGCGCGTG CAACGAGAACCGCTGGCGGCTCCATCGCCAGGGAGTCTTCCC CGAGTGAACAAGTGGGACCAACAACATGACGCGCTCTCCAGGG GATGGGGACTCAAGTCGCCCATCGGACCGCGCCGCAAGTGCCTTC CTCTCGCGCGTTCGAGGACCGCCACGCGCGGATCCGTGCCT GGTGGCGCCATACCTCCGACCCGACCTCGTCTTGCCAGACGCCGA ACTGGCCGAAGCCGTGCGCATTAAGTTGCTACGACGACGGCCAATGC GGCGGCGTGTCCGGCATGCCAACAGACCGCATCGACCGTCCAC GCGTACACGTGCGTTTCGTATCCATTGCGCACAGCACGCCACGAGC TGGTGTGCGCGAGTCGAGCTGCGCTGCAAGACCGAGAACGCGGTCC GCGAACAGTGTGCGCATCCCCCGCTCGCGCCACCGACACAACAA CAACGGGACGAGGACGGCAGCCAGTACCACCGCCGACGACACGC AGCGGCCACGCGTGGGACCAACGTCGCCCGGAGACCGCGCTTC AGCCAGAGCGCGCGCGCCACCGCGCGCGCGCGCGCGCAAT CATCAACGACAACAGCTCTGAGCGACGACGACGACGACGACCA TGACGACAACTGCCAGGAGGAAAGAGGAGGAGGAGAAAGGAAC GTACGTCCCCGCCACTACCGCCACACCTTCGCCGCGGACGACA CGTCGACAAACGCGAGGACGAGGACGAGGACGACGCTCAGGAGC ATGACGAAGACGGTAAAGTGACAAACGACGACGCTACAACT GCAACAGCAGCAGCAGCGCGGGATGAAGCGGTGACGACCTGGAC TAGAGTACAGTACAGAGCGTCACTGAAGCGTGAAGCGCGCGCAGC GGAGAGCCCCAACCCGAGCGCCCTACCGCCACCGCGCGCGCTAC TAGCGCGGACCTGTGGTACCGCCACCTCCACCGCGTGGACGTGAT GGTGGCGCGCTGCCGTGGTACGCGCAAGGCGCTTCGACCGAGC CGTCAGCGCAAGGCGCGAAATAGCGCCTGCGGTAGCGGCGCTC GATCGCAAGGTGGTGGTTCGTGTGCGCTTGGCGTACTCTCG AGCGCGCAAGGCTTCCTCAAGCGCGCATGGCGACACGACGCGG GCCAAACAGGCAAGCGCGTCTGCGCTCGCGCGCTGCGCGCTCATCG GGACGCGCGCTCTCTACGCTGGGCGCGCTGCGCGCGCTCATCG TCGGCGCAACTAAGCGCGGACGAGGACGCGGACGAGGACGAGGAC CGACGCGGACGCGGACCACTAGCACCGCACGCGCGGACGACATATTG CGCGGCTGTACAGGCGGTAGGTGAGCCAGCGCGCGCTTCTGAGC CTCAGTCGGCTTGAGCCCCCGCTTCCCAAGGCTACCGGGCGCGCTC	CTTGCTT CGTGTGCA CCCGGACC GTTTGCA GCCCTGC CTTCTGAC CCGCTCT CGTAGCCG TTCTCATC GCTGAAGC GTTCTACG CGCTGGCA GCAAGCT CGGCGCTA GCTTGTAG CGCGCGC GTGGCGC TCGCCAAC (SEQ ID NO: 1155)	VPLFAAVFEDAATP GSGAWLAAIPSDPTL VLPDAELAEAVRIKLL TTTANAAGVCPACH KTGIDPSHAYTCVSL HLRTARHDVVRVE LACKTEKPVREHLAI PPVAPTDNNNGDE DGSPVTTADDNADG HAVATKRRPETRASA RAAAAAATAAAAAAI INDNLSDDDDDDDD HDDNCHGEERGE RNVTCPGHYTATPFA ADDTLDSDEDED NAHEDDDEDGKDD NDDDVNNCNSSS DGDEGGDDLDYEYS DQSVTRSDAATGES PNPERPTPTALLRA DLWLPATSTAVDVM VAAACRRSRKAFDR AVSRKAAYGPAVA DGSIAKVVPFVSPF GVLSRPAKAFKRAM GDTTAAKQAKARLRL AVAAVRGTARLSYA WGACALIVGGN (SEQ ID NO: 1400)

CRE	Cre-1_FCy			TTTTGCGCCTGGTTTTTGGCGGCTGTTTTTCTCTCCCTTTTCCCCCT TTTCCATTTGTACTTAGTTTTCTTCGGCCGCGGAGCTTGTGCGCGG ATAGTGTTAATATGTTTAAAAACGTTGTAATAAATAACTGTTTAAACCT AACCTAACCTAA (SEQ ID NO: 1032)	ATCAATCT AATACTGA AGGCAATA CCAAACTC AAACCGAA ATCAAAAT CGTTAGAA TCAATATA CGACCCCC GCTGCTGT ACATGTCC AGCCGGAT CTCGTTGT AAAGAAG ATTGCAGC TGTAATAA AGTTGGAC TTCTTTGT CTTCCTGT GAAGAAG TTGATTGT GGCTGTTT AAATCTT TTCATAAT AAAGAAAT A (SEQ ID NO: 1156)	TAGCAC CACCATC TATTCAT ATCCAC ACACTG ACCACCT CCACCTT CACAAAC TCCACTC TCAATTC CCCTGA CTACTAA GAAATA TTTCATG GTGGTT ACATTG GAGGTA TCCCAAC ACCAAG CACAAA ATGAAC CCACTA ACCTCT CATCCTA TCCACG GGGACC ACCTTTG AGCAGA ACACCA TTCATAT TACAAC CTTTAGC TAGATT AAGATA	MAPLPWNAATSSPP SPVPLTNDKKDSTLP TATSKNLKNNNNK NNNTNRINNIKNND NTNDGSNKINLKLPP AAVKITNPYKNKKKN KKKNAGKSNPKTN QNPNSPLSDNDDD DTDSSNITINRRLKFG TDDLAPPNPPSNTNT IGTATAATAATATT ATAATATTATNTTT TTTTNTTGDNLASN INNNNNNNNSGNN SNTNNINTDNGS NNRPPRVTVDPRS DLPGAIEAANKMLD EYGDHVDNPGSH LSGLSSQDLWQGG YFRRIPHNSLYDCP KGKLGKIDITNEYNLF EAIMNGKCNMEKLL VFPVVVLQRRHGVTK NADVRRLLSRLTAW KEGFKYLVEDTHRD LIAQSKARGDTTPA HRAKVYSKLMRGHL QSAVNYITDREGGI LYPYDVDEKSGHTVS RVLQDKHPSMRDPG PTAMPAYESPELPT LEITADTVEIVAGKLS GGAGLSGVDSIQLKH LLLLHHGQASQRLRNV

TCCATCAGGTCAGCAAGCAACGACTGCGCAATGTTTGTGCAAAATT TGGTAGATGGCTGCCAACGAGCACCCCTGGGCTCGTACCGTGCC ATGCTAGCAATAGGCTTATGCGCTAGACAAATGCCGGAATTCGAC CAGTCGGTATAGGTATACATGCGCTGTTCTTCGCCAACTGTTCTA GCAGTCTATGCTTATGCTACTGACTGTTGTGGTCAGACCACTCTG TGCCGACTAAGAGCCGGAGTTGATGGTCCATACATGAGACTACGGC TATGTGGAGGAGATGGAATCTGAGGAAACACAGGTTTCGTAATT GACGAGACAATGCAATCAATGAGGTCTACGCATCAACATGTTATGGA CGATCCGCCACGAATGGCTGCTGGAGTCGATTGCGCTTCAACTGCTA TCGGCACACAGCCTACTAGTGTGACGGAATCCAGGCGGGAACCTTC ACTTCTTTCTAAAGAGGTGTCAACAGGGCGACCCATTTGCGATGAT AGCATATGGTGTCTCTACCACTCATCCGCAAACTGAAAGAATTA ATGTATTATTAGTTCAATCTTGGTATGCAGATGATGCTAGCGCAGCTGG CAAAATTTGATGAAATACAGCCCTTTTCAAGATTATTACGAATGGAC CTGATTTGGGTACTTCTTAATGCATCTAAGAGTATCTCATCACCCATC CCGACAATGTGGTTCAGCTCACCACTTCTTCAACGAGACCCATGGCT AGTTTCAAGATCAGCACAGGAAGTCTTCTGGGTGGTTCATTGGA GATACCAAGTCGAGATGAATACGTATCAACAAAAATCGCCGACTGGA TCCACGGCACCAAGGAGTAGCAGCAGTAGCAAGATTGAAGTATCCAC ACGAGCTTACACAGGCTTACCAAGTGTTCAGCACAAGTGGAGTTT TACTCAACGTGTTTCTGCTGCTGCTGCTGCTTCCAAACCACTGGAGG ATGAATCACCATAAATTTGCTCCCGCTATTTGGAGACCCCTATCC ACTATGGATGACAAGCTCAGACTTTGACCGCTCTGCCAGTCAAAATG CTGGGCTGCTCCCGAATCCAGTTACCTCTCCGCAACCACTACAAG AATAGCACTCTATGAGTCTCATCTCTGTTGGCTGTTCAAGGCAAGAT CAACTCAGTTTACAGGACCACAGAGATACCTGTCAATCCTCTCTCCG CGTCCGAGAGCTCCGACAAACCGAAATGATCTTCAATGACCAACCTC CTTGAGCTCTCCCTCCAGCTGCTGAGGTCAACCAAGCACAACAAGAG CAATCAAGCGTCTGGGAAACCGGTCTTGGCTTACTACTATCCCTAAT CACATCAACGGTAACATCTCGGATGTGACGAATTTATGATGCTATTG ATTGAGATACCAAAAAGTGCCACAATCTCCCTGCCAAATGTGATGGC TGTGGCTCTGCAATTTGATGTAGGGCACGCGCTCCAATGCAAAATCCGGG GCCTAATCATTAGACGTGATGAACCTCAATCTTGAGCTTGCTATTTTA GCAAGATGGCTTGAGAGAATCTGCAATACGTGCTGAACCTGAAATCA ACCCAGCGCTCTATTATGATTTCTCCACCAACCATCACAGCCATCGAC ACAAACGGAGACCGAGGAGATTGTTGATCAAGGGCTTTGGGACAAT GGAATGGACGCTATCATGATGTCAGAAATACAGACACAGATGCCAAAT CCTATCGAACAGAGACCCCAAAAAAGTCTCTACAGTCAACAGAGAAGG					ATTATTT AGTACA TATTTTA TACTATT AAAAAA AAAAAA AAAA (SEQ ID NO: 1279)	CAKGRWLANEHPP WASVRAMLNRLIAL DKMPGIRPVGIGDT WRRFFAKVLAVSM SYATDCCGSDQLCAG LRAGVDGAIHLSA MWREMESEENTGF VLIDADNAFNEVSRI NMLWTIRHEWPAG ARFANCYRHHSLV VRNPGGKPTFFSKE GVTQGDPFAMIAVG VALLPIRLKELNVLL VQSWYADDASAAGK FDEILRLFQDLRMG PDFGYFPNASKSILIT HPDNVAAHHFFNE THGLGFKISTGRFLG GFIGDTSRDEVVSTK IADWIHGTKELAAVA RLKYPHAAVTGKCL QHKWSFTQRVIPGID DLFQPLEDELNNLLP ALFGDPPSTMDDKLR LLTALPVKHAGLALP NPVTSSATNYKNSTL MSSHLLAVQKINF SLQDHRDTCQSSLSA SRELQRTENDSSLNL LAALPPAAAGQPSTT RAIKRAGETGLWLTTI PNHINGNILGCDEFID AIRLRYQKVPHNLPA KCDGCGSAFDVGHA LQCKSGGLIIRHDEL NLELASLAKMALRES AIRAEPEINPSASIMD SPTTTAIDTNGDRG

CRE	Cre- 1_HM	.	Hydra vulgaris	AGAAAAAGAAATACCTCGATCAATGCTACTCCAACGTCGAGCCTT TACCCCTTTTGTGCTCTGTGGACGGCTGATTGGTTACGAGGCCAGCA ATGTGCTAAAGCAATTATCAAAACGTTTAGCAGATAAAATGGAATAAGCC TTATTAGTTACATGTGAATAGTCGCTCACGTATCAGCATTCATGTG CGGAGCTTCCAATCAATGTCTGAGAGGTTCTCGAATACCATTCAAAAC AATGAGCAGACAAAATCAATGGAGGACGGTGCAGGGCCGCGCTCTA TAGAATTGTCGCTAGCACCACTATCTATTCATATCCACACACTGACCA CTCCACCTTCACAACTCCACTCAATTCCTGACTACTAAGAAATATTT CATGGTGGTTACATTGGAGGTATCCCAACCAAGCACAATAAATGAACCC ACTAACCTCTCATCTCTATCCAGGGACCACTTTGAGCAGAACACCAT TCATATTACAACTTTAGCTAGATTAAAGATAATTTATTAGTACATATTTTA TACTATTAACAAAAA (SEQ ID NO: 1033)	TTTCTAAT GTTACGTG ATATGATA TGGTAGT TCATGGT AGTTATG TTATGCT TAGTTAT GGAAATC GTTATTT ATGGACA ATATTGT TGCTGTT TAAATTT ATGTAACG TGTGCAAT TGATGTAT ATCTTGA ACTTTTA ATCTGAAT TTTTACT GGTTAAT ACGTTTAT TATATCTT CGATTGAG CAATTTAT	TAACTTG TATTTT AAATG TTTTAT AGTT (SEQ ID NO: 1280)	DLLIKGFWDNMGMDA IIDVRITDTDAKSVRT RDPKKVLSQSEKEKK KKYLDQCLLQRAFT PFVSV DGLIGYEAS NVLQLSKRLADKW NKPYSVTCGIVRSRISI ACARASNQCLRGSR PFKTMRSRQIQWEDG AGAGLYRIVR (SEQ ID NO: 1401)
CRE	Cre- 1_HM	.	Hydra vulgaris	TTTCTAATGTTACGTGATATGATATGGTATGTTAGTTCATGGTTAGTTATGTTT ATGCTTAGTTTATGAAAAATCGTTTATTTATGGCAAAATATGTTTGCTG TTTTAAATTTATGTAACGTGTGCAATTTGATGTATATCTTGAACCTTTTA ATCTGAATTTTACTTGGTTTAAACGTTTATATCTTCTCGATTGAGCA ATTTATCTATCAAAAGCAATTTATCTTCTGATTGAGCAATTTATCTCTG ATTCGAGCAATTTATCTTCTGATTGAGCAATTTATCTCTATCAAAATGAGC ATATATCTGCAATTTCAAAATATCTACGAAATAGTTCACTTACTGAA AATCATTAAGTAAAGAAAGAAAGAAAGAAAGAAATAAAAAAGTA GTAATCTTTTATCAACAATATCATCTTATTTAAATTTAAAGGAATAT TTTGGTTTGTACTAAATCATCGCTTATTTTCAACGAAAGGGGGCT GCTATATTTTGTGAGTTGTTTATCTTAAACCTTAACTTGTGTTCA ACCAACGTAAACATGATTGCTGCTGCTCAAAATATCTACAATATAA AATTTATCAATCTTTTTCGTTACGGTAAACAATAACAATAAAATAACT ATAGTTATTTTATGTTTACCGCATATTTGTTAACTATAGTTAAACAAAGT ATTTGTTATGGAAACATTACCGATATCTCTTTAAAGGTAAACAAACAAA CATAGACGGCATCTTTTAAAGGTAAATTAAGTATACGGCTAATAATAA AATATACAGCTAATAATAAAATCTTCAATGAACATGTTCTATATGCAA AAGATGTGATCGTAGCTTTACTACCTTAAAGGGACTAAATATTCATAAA GGTCAATGTAAGATCTTTGTTTCCAATAACAAATAAAATAAAACAAATGT AGTTAAACAAATGAATTAACAACACCGAATAAAACAAAGGTGGAATTAAT ACGATATTAACCTGCGATGAGATATCTGTAGAACACTATTCACACCAACA CACCTTACTTACCCAAAATAATTTGTGAATCTATATAGATCCCAAC GACTATCTATGGGTCTATGCGTTTGTAGCTTCTTCTCAACCATGTCAA CACAATATACGATGAAATAGTATTTTACCAATAAAACCTTTTAAAGTGC CATCAGGAAAGGTGTAATTTTATAGAAAGAACTGACCTTTTGGCT AAAAAGTTTAAATATCGAACCAAAATTTGAATGGAATAGCCATGAAATGT	TTTCTAAT GTTACGTG ATATGATA TGGTAGT TCATGGT AGTTATG TTATGCT TAGTTAT GGAAATC GTTATTT ATGGACA ATATTGT TGCTGTT TAAATTT ATGTAACG TGTGCAAT TGATGTAT ATCTTGA ACTTTTA ATCTGAAT TTTTACT GGTTAAT ACGTTTAT TATATCTT CGATTGAG CAATTTAT	TAACTTG TATTTT AAATG TTTTAT AGTT (SEQ ID NO: 1280)	MINMVSICKRCDSF TTLKGLNIHKGQCKIF VSNTNKQINNWN ELTTPNKNKVEINTIL NCDEISVEHYSTNTPY LPKINICESIIDPNYL WGHMPPFSELLNHVN TIYDEIVFYHKNLFKV PSGKGKGMFIEELTF WLKQFNRTKLNGI AMKCFMIVPSLMLQ KPSIRSKAKEHAECV RRITLWRNGNFSLEMI REIRYQSKINTSKKKR TFEDISRIFAKLMME GKVAALKVLDRESS GILQCESEVLKELKSK HPDETVPQDNCLLYG PLQNTPECLFDSIDEI SIFNSALQTKGSAGPS GMDADLYRRVLCSK CFGPSCKTLREEIATF TKNIATKSYQPDIVQP YIACRLIPLDKNPGIRP IGIGEVLRRIVGKTISH HCQKEIKEAAGPLQT

TTCATGATAGTCCCTCCCTTAATGTTACAGAAAGCCCTCAATACGGTCCAA AGCCAAAGAACATGCAGAAATGTTTAGTAAGACGAATTACATTATGGAGA AACGGGAACCTTTAGTGAATTGATCGGGAAATTAGATATATTCAGAGCA AAATTAAACACCTCAAAAAGAAAAGGACATTTGAGGATATCTCAAGGAT ATTGCAAAACTAATGATGAAGGTAAAGTTGCTGCCGCACTGAAGGTT TTAGATAGAGATCATCTGGCATCTTGCAATGCTCGGAAAGTGATTGA AAGAAITGAAAAGTAAACACCCAGACGAAACTCTGTACAGATAAATTG TTTACTATACGCCCGTTACAAAACACTCCAGAATGTTTATTCGATTCAA TTGATGAGATAAGTATATTAACTCAGCTTTACAGACTAAAGGATCTGCA GGTCTCTGGAATGGATGCAGATCTTTACCGTCGAGTCTCTATGCTCAA ATGTTTTGGACCTCTGTAAAGACTCTACGAAGAAATAGCAACATTTA CAAAAATATTGCAACAAAATCCTACCAACCGGATATAGTTCAACCCCTAC ATTGCATGTCGACTAATCCCTTAGACAAAATCCCGGAGTTCCGCCCAT AGGAATTGGGAAAGTTACGTAGGATTGTAGGTAAACCATTAGCCA CCATTGTCAAAAGAAAATCAAAAGAGGCAGCTGGACCACTACAAAACCTGC GCAGGACACGGTGCAGGAGCAGAAAGCTGCAATACATGCTATGCAAAAG ATATTTTCATCAGGAAGATACAGATGGTGTGTTGTTAATCGATGCTAGGA ACGGTTTTAACTGCCTAAACCGTTCTGTGCACTACATAATATACAGATA ACTTGCCCAATCTTAGCTATGATTTAGTCAACACTACCGTAAACCGGC AAAAATTATCATCTACGGTGGAGAACTATTTTTCGAAAGAAAGGCACA ACGAGGGCGATCCCTCGCCATGCCATGGTACTCCTAGCACTGCTGTA CAATCAATAATACATTGAAACTAGTAATCTCTGATGTAACCAAGTATG GTTAGCCGATGATGCTACCGCTGCAGGAAAATACAGTCTTTAAAAAAG TGGTATAAATGCCTAGAGGATGCTGGTGTGTTGTTATGTTATATGTA ATCAGTCAAAATGCTGGCTAATAGTAAATCTGATAACCAAGCTGAAGA AGCTAAACTTATATTGGCAACTCCATAAATATACTACTCAGGGAAAAA GGCACTTAGGAGCTGCATTTGGTTCGGAAGCATACAAAAAAGTGATTG CGAGGATTTAGTAAGTAAATGGTCTAAAGAACTTAAACAATCTCTCGGAA ATCGCCACACGCAACCAAGCTGCTTATTCAGCTTTTATTAAGGGTA CAGATCTAAATTCACCTACTCTTACGCACAAATTGAAGCTTTTGAATAATT CGTAAACACGAGTGGAAAAATTTTATCAGAAAAATTAATACCTGTTTGT TTGGAACTGATTGTTCTATAATCAAAAGAAAATAGGGATTTATTTGGCGCT AAATCCATCGGAAGGAGGACTTGGAAATTTGTAACCTTAATAACTGAGGCC AAGGAACAGCATACTGCCTCTAAGAAAAATACTAACTTGCACATAAAAT CAATACTCGATCAGTCAGATGTTATGAAAGAAAAAGATGATTTCCGGGAA AACATTTTCAGAAAATAAAAAACAAAAACAAATATGGATAAACTAAAAAA AAAAAGAAAGAGGTTAAAAAAATACATGCAGGACTTCCAGAAAAACCTT AAACTTCTGTTGAACAGGCTGTGACAAAGGTTGCCAGCAGCTGGTTAA	CCTATCAA AGCAATTT ATCCTTCG ATTCGAGC AATTTATC CTTCGATT CGAGCAAT TTATCCTTC GATTGAGC AATTTATC CTATCAAA ATTAGCAT ATATACTG CAATTTTC AAATAATC TAGCAAT AAGTTAC TTACTGAA AATCATTA AGTAAAAAG AAGAAAG GAAGAAA AAATAAAA ATAAAAAG TAGTAAAT CCTTTCAT AACAAATA TCATTCTA TTATTAAG TTTAAAGG AATATTTT GGTTTTGT ACTAAATC ATGCGTTC ATATTTCA CCGAAGAA GGGGGCT GCTATATT	CAGHGAGAEAAIHA MQKIFHQEDTDGL LIDARNAFNCNRSV ALHNIQITCPILAMYL VNTYRPAKLFYGGGE TIFSKEGTTQGDPLA MPWYSLSTVTIINTLK LVIPDVQVWLADD ATAAGKLQSLKKWYK CLEVDGGLGYVYNQ SKCWLIVKSDNQAE AKLIFGNSINITTQK RHLGAALGSEAYKKV YCEDLVSKWSKELNN LCEIATTQPQAAYSAF IKGYRSKFTYFLRTIEA FENFVTPVEKILSEKLL PVLFGTDCSIKENRD LLALNPSEGGLGICNL ITEAKEQHTASKKITN LHIKSILDQSDVMKEK DDFGKTSEIKTKTN MDKSKKKKEEVKKIH AGLPENLKLVEQAC DKGASSWLNTLPIKE QHLDLNKEEFKDALR LRYNVPLANLPSYCA CGEKFDELHAMSCCK GGFVCNHRHNDRL TVCLNKVCTDVQAEP HLIPLTNEKFNFKTAN TNDEARLDIKAKGF RKGETAFFDVVRVTHV NSKSKKKQPTKHIFRR HEDAKKREYLERVLE VEHGTFTPLIFGTNG GFGDECKRFTALLAQ KLSLKMGERYGAVIN

ACACCTTACCAATTAAGAAACAACATCTAGATCTGAATAAGGAAGAGTT TAAGGACGCACCTTAGATTGAGATATAATGTGCCACTTGCCAAATTTACCAT CCTACTGTGCTTGTGGAGAAAAATTTGACGAGCTACACGCAATGTCATG CAAAAAAGGTGGCTTTGTTGTACAGACATGATAACATCAGAGATTTA TTAACTGTTTGCCTAAATAAAGTTTGTACTGATGTTCAAGCGGAGCCGCA TTTAATCCATTGACAAATGAAAAATTTAATTTCAAACTGCCAATACCA ACGACGAAGCTAGATTGGATATAAAGCAAAAGGGTTTGGAGAAAAAG GAGAACTGCAATTTTGTAGTTAGAGTAACGCACGTAAACTCCAAATC CTCCAAAAACAACCAACAAACACATATCCGTAGGCATGAAGATGCA AAAAACGTGAGTATTTAGAACGAGTTCTAGAGGTTGAACACGGGACA TTTACCCCATTAATTTTGGTACGAATGGTGGGTTTGGAGACGAATGCA AACGCTTCACGGCACTACTCGCACAAAACTGTCCTTAAAAATGGGTGA GCGGTACGGAGCTGTATAAATTGGTAAGGACACGCTTTCCATGGAG ATTACTAGAGCCTCCTACTCTGCTTAAGAGGGTACGAAACCCCATTTAG GCATTATAACACTGACGATGTGGCCTGGAATAATGTGCAATGTGGACTT ATTTAACTTGTTATTTTAAATGTTTATTAGTTT (SEQ ID NO: 1034)					TTTGTGTTG AAGTTGTT TATCTTAA AACTTTAA ACTTGTGT TCAACCAA CCGTAAAC ATTAGTTC GCTGTTCG CTCAAAAT ATCTACAA TATAAAAT TTATCAAT CTTTTTCG TTACGGTA AACAAATA ACAAATAA ATAACTAT AGTTATTT TATTGTTT ACCGCATA TTGTTTAA CTATAGTT AAACAAAG TATTGTTT ATGGAACA TTACCACT ATCTCTTG TTAAGGTA AACAAACA AACATAGA CGGCATCT CTTTTTAA GGTAATTA AGTATACG GCTAATAA TAAAAATA TACAGCTA	WLRRLSMEITRASLL CLRSRTPFRRHYNTD DVGLENVQCGLI (SEQ ID NO: 1402)

CRE	CRE-1_LSa					ATAATAAA ATCTTCA (SEQ ID NO.: 1157)	TGAACT ATATTTT ATATATT AAAAAA A (SEQ ID NO: 1281)	MASSSTSSSDICLPF RSFHCCPNGEVGSKG IXRMISHIKRHHLLTE DRKCVLREALSSDVG LFMAVEETLKAFGQ WMCGKCMTHALS RYCHHPDGRVXFVT GADGSSRYIVGILKPS TKESVTNALGGLVFD VGLLDRVFEKEPIITVK SIPHSCLRAFSQALKT ALYKVIAQPGSVDA WICLLLPRLTQVFR PKNRQECRSNRSKL QQSSILKSLDTWKE DGIRKLQNMILNIP EVGAMGQGGGILQK ESTSNTNIRQCLRKV ADGHFTAQVAVLCSS GVAPYNGDTIKALED KHPFRPPSPMPSIIS EPLVADDFCVFGCIK SFPKGTSCGRDGLRA QHXLDALCGECSAIA TDLIRAITSVNWLWA GRCPTIAEFVASAPL TPLIKPDNGIRPIAVG TIWRRLLVSKVAMKG VGKEMAKYLNDFQF GVGVSGGAEEVLHS ANRVLSEHHADGSLA MLTVDFSNFNLVD RSALLHEVKRMCPSIS LWVNFLYQQAARLYI
						ACATTAATA TTAGAGAG GTTGATGT TTCAATGG AAGAAAT GAAATTC AAGAACT ATTTTGT GCCACCA AGTGTG ATAAAATG TCCAACT AATTTTCT CTGTGTC AGCTTAT TGTTCAAG ATAATGA GTTTGCT AGTTGAG CGTTCCT GTGCACAC CAACAGTG TGTTGGTG TGCCATT CCTTTCCT CCTTTTA ACTATTGC TTCATAGC TTAAGCTT CATCTCGA GGCTTGT CTCTGT (SEQ ID NO.: 1158)		ACATTAATTAAGAGAGTTGATGTTTCAATGGAAGAAGATGAAATCCCA AGAAGCTATTTTGTGCCCACCAAGTTTGATAAAATGTCCAACTAA TTTTCTCTGTTGACGCTTATTTCAAGATAATAGTAGTTGCTTAGT TGAGCGTTCCTGTGCACACCAACAGTGTTGGTGCCATTTCCTTC CTTCTTTTAACTATTGCTTCATAGCTTAAGCTTATCTCGAGGCTTGT CTCTGTATGGCTTCTTCTACAAGTTGAGTGATATTTGCTGTGCC GTTCAGAAAGCTTCCATTGTTGCCCAATGGTGAAGTGGGAAGTAAGGG GATTGKCCGTATGATTTACACATCAAAAGGCATCATCTACTACTGAAG ATCGTAAATGTGTTTACGTGAAGCTTCTAGTGATGTTGTTTATTT ATGGCGGTGGAAGAACTTTGAAGGCCCTTGGTCAATGGATGTGTGG AAGTGATGACTTTGATGCTCTTAGCCGTTATTGTCAATCACCCGGATGG TCGTGTGAGKTTTGTACAGGGCTGACGGCTCGAGTCGTTACATTGTC GGTATTCTAAAGCCGTACTAAAGAGTCGGTGACAAATGCTCTTGGAG GTTGGTTTGTGATGTTGGCTCTTGTGATCGTGTGTTTAAAGAGCCTATC ACTACTGCAAGAGTATCCCCATAGTTGTCGCTTGTCTCTCAGGC TTTGAAACTGCTTTACAAGGTGATTGCCAACCTGGCTCGGTGATG CATGGAATGTTGTTACTTCTCGTGCACACTGCAGGTGTTAGG CCCAAAATAGACAAGAAATGAGTCTGGGAATAGAAATCCTTACAAC AAAGTCCATCCTGAAGTCTGGATACATGGGGAAAGAGGATGTA TCAGGAAGTTAGTTCAAAATATGTTAGACAAATCCGAGGTGGGCCAT GGACAGGTGGAGGCATCTTCAGAAAGGAGTCTACATCAAGTAACAC CAACATCAGGCAGTGTCTCGTAAGTTGCAGATGKCATTTTACCGCA GCAGTAAAGTGTATGCTCATCGGTGTTGGCCATATAATGGTGATA CTATTAAAGCTTGGAGGACAAACCCCTTCAGGCCACCCCATCCATG CCGAGCCCATAAATTTCTGAACCTCCCTTGTAGCAGACTTTGACTGTG ATTTGGTGCATCAAAATCCTTCCCTAAAGGAATTCWTGCGGGAGAGAT GGCTTGAGGGCTCAACACWACTAGATGCCCTTTGTGGAGAGGGTCT GCTATAGCCACAGATCTCATACGTGCTATCACCTCAGTGGTTAATTTATG GTTAGCGGGAAGATGTCGACCATTTTGGCAGAGTTTGTGTCATCCGCT CCTCTACGCCCTGATTAAACCTGACAAACGGGATCCGTCCTCAATGGCAGT AGGCATATATGGAGACGTGCTGTTTCCAAAGTTGCCATGAAAGGTGTG GGTAAAGAAATGGCAAGTACCTTAATGATTTTCAGTTCGGGGTGGTG TGTCGGGGGTGCTGAGGTGTGTTACACAGTGCCCAATAGGGTGTGA GTGAACACACCGCTGATGGGTCTTGAATGCTGACAGTGGATTTCTC

194

CRE	Cre- 1_MB	.	Monosi ga brevicol lis	CATCTGGCGTGAACACGCTTGTGACAGACAAATCTGCAACCCCGCTCTTT GCGGCCGCGTTTTGGCGGCGCTCGCTCCACCGTGTCCGCTCGCTT GCTCGCTTGTGCCCCGCGGACATGGCCACTGAGTCCGGCGGCGAGG ATTCTTGACCCAGTCCGCGGTGCTAAACGCCGAGTCCGAATCACC TCCAAGCAACACCAACCTCGCTTCCAAACTCATCTGTTCTGCAAAAC ACACAAACATGGACGCTCGCCAGACCGTAACATGTTTTCCTGA CCCCATGACACCCGCTTCCCTCAGCCCGCCACTGTCCCTACCG CCCGTCCCTGCTCATGTGCCCTGACGTCCCCGCTGCCGTGCCGACC GAGTCTCGGACGTCGCTGTCGCCGCGGCGGAGCACTCCATCATCGGAAACACTTC ACGTCAAAACGCGGCGGAGCAACATCCCATCATCGGAAACACTTC AATGCTTCTCCCGCACCCCTCGCACGCCATCCAGCCCTGCTCCTTGCT CAAGTTGCTCGGCAATGCTGCGCGCCGCACTGCCACATCTTCGGCCC CTATGACCGAGGATTTGTGACGATCGGTGCCCTCTGAGCCAAATGGCAG CGGGAGCAACAAACCTCGCCGAGTCCACAGGGCAGACACATCATTC GATTCCTAACACACATCGGATTTTGTGACCATGTCTTCGGATGAAAGCG ACTCCCTCTCGCTCCACCGCACTCGCGCGCCACCCCTATCGCCCTC CCGGCATGATGGTGACGGTGACACAAACGGCAGTCCACCGCTGAGC CATTGGTCAATCACCTACACCCGCTCAAAATGGTGTGCCATATCCATCG GGTACAAACAAACCACTCCGATCCCTCTCCGCCCTCTGCTCACCCCT GCCACTACATTTGCCCGCTGCCATTCACATCTCTGCGAACACAGTGA GCATGCAAACTCAGCCCACTGGCGAAGTCACTGAGAGTGAAGCAACA CAATACAGCGGCGAACACAGTGAAGTGAAGCAAGATGTTCTCTCAG CGATCCCGCTCCGCCATCGTCCCAACGTGCTGGATGCCAGCGCAAG GTCTGCTGAAGACATCGCCACAGGCAACTCTCGCTGCCCATTTG GGCTTTGCAATGCAAGGGGCCCCGCTTGACCGCAAAAGCTGGGTCA ATCATGTACTACGAGACCCCTACGACGAGCAAGCCACTGATCTGGT CAAGCAGGTGATGGAGGCCAAATGGTCGCCAGTGCAACAAAGTCA CCTCTTTCGAAGCTGCTGGTATCAGTCAACACCGCTCCCGATGGTG CCAATCTGAAGCGAGCGACGAGGCGTTGTTTCATCGCGCTGGACAG ACCTGCTGAGATTATCGTGGCGCTTGCCCAACAGTGTGAGGCTC TCGATCAGTGTCTGCGAGCTGCTAAGCTCGCCCGCATCCACTGATG CAGCGACCGCTACCCATCCACGCCACCGAGACCAAGCTGATGGCTG CCACCTGAGCCAGCTGTATTGGTCTGCCGTCCACTCGGACTATACCGCT GAAGAGCGAGAGATGTCTGGGCTTTGATTTGGCCTTGCCTAGCATGT TGTTGCTGCTCCCTCGACCGCACTGTCTACGATTGACCTGCGCAATATG TTTCAGCATCGTCCGTTGGCTTGTGACGGGGCAACTAGGTCGGGTG TGGACGCCATGCGCAAGGCGAGTCGACGCAAGCAGAGCCGTCGAGGAC AGCTGAACGCGGCGGCGGCCACCCGCAACGACGAGTCGAGCCAGAC	CATCTTGG CGTGAACC ACGTTGTC AGACAAAA TCTGCAAC CCGCTCT TTGCGGCC CGCGTTT GGCGCG CCCTCGCT CCCACCGT GTCCGCTC GCTTGCTC GCTTGCTT GCCCGCG GAC (SEQ ID NO: 1159)	TAGTA GGCACC GTCTCG GGGTC CCTCTGT GGGAT CCCTGT GTGAC CTGTC CTCCCTA GGTGT TCCTGT TGTGCT TTTGATG GCTTGA CTTGAT TTTTGT TTAATT TGCTTA ATTTTG CTGTAT TGTGT GTATTT TGCTGA ATTTTG TAAGT CCTTGT ATGATG TCCTGT CTTCTGT GGTCGG TTGTTT CCTCAAT CCGACG TTGTGTC TCGTTG GATGTG AGCGTG	MATESGGEDSWTQ VRGAKRPSAESPNS TTTSPQTHRSAKHT KHGSARHNRHVFP DPMTPPLRHARHS VPTARASHVPTSP AAGATESARAVWA AEPVTRTSNGGGEQ HPIGNTSNASPTPR TPSPRSFAQVAAA MPAAATATSSAPMT EDLSASVPSEPNSSG EQQSPSESTGQTHHS IPNTPSDFLTMSDES DSPRSTALRAPTPA PPAHGDDGDTNGSA TPEPLVQSPTPAQM VLPVPSGTQTHSDP SPSASPPATTILPAAI SHPVEHSEHANSAPL GEVSEETHNTAGEH SESEQDVLLSDPAPPI AANVDAQRKVLLKT SGHRQLLACPFGLCK CKGPRLDRKAWVNH VLREHPYDEQATDLV KQVMEAKLVAQCCK CHLFEAAGISQHS RCGANLKRATEALFH AAGHDLLEIMRGAW PQQCVGSRSISVCELLK LARHPLMQRSRYPSN ATETKLMAATLSQLY WSAVHSDYTAERE MCWALILALPSMILLS APSTALSTIDLNMIF HDLRLWLVTGQLGR VVDAMIRKAVARKQS

GCCTCAGGTCGCTCGTCCGCGACCCGCGACCTGGCGGACGAAAGCCTGGG CAAACCACTCAGAACCGTCTGAACCGAGGTGAGATTGCCAAAGCATT TGATGCCGACAAGGCTCGTCCGCTGATTGGTAATCTGAGGTTCAAGGCC GTGCGGACCTCTTGATACCGCCCGGCTGACCCCGTACATTGCTTCCA CAGCGCTCCAGCTACACTGGCAGCAGCCAGGCTGTGAGCTCCCC AACCGTCTCTTACCAAGGTTGAGTCCCAAGGCTTGGCGGCCACC AAGGTGTACCGACCCCTATGGTGGTGGTAGCTTCTTGCTCCAT CTACCGCATCAAGGACACTTCAAGTCTTGGCCACCGCAGGCT TCTACGACGACCGACTGCTCTTCTGATGGAGACGCGCTCAGGGCC CCACACCGCACTGAGTCTCAGTTCGCTTGAACAAGATCTTTCAC CACATTGCCAACAAACCGTCCCGAGTTCGATTCGACATGCCCTTGTCTC CATCAACTACACTATCTGGAGAAGGCCAATGGCAAGTTTCGACCGTG GGCAGGATTCCATCTTCAACAAGGTTGTCAACCGCGCTCTGCTCGAGC AACAGCAGCCCCATATTGCCACCTTGTACAGGCCAGTCCAGAGCTGGC CGTCGAGTCAAGGACGGCATTTACAGCAGCGGTTGGCATGGCCTTTGG TGAGCTTCAAGCCTGTGAGTCTACCCCGGCTGGACCATGCTCTCCCTC GATTTCAAGAGTGCCTTCAACTACACCGACCGAGCAGCGTGCACGAGA TTGTGCCGACAAGGTCCTGGCTCTTGGCGCTTTGAACGACACTA TGCTGACCCCACTACCACTGCTTGTGACAAGTTCTTCAAGGTGATTG ACATTGATGTGGCAAGGCTTGTGAGGGCAACGATCTCGCCCTT CTTCTTGGCCTGACTCTGTAGTCTCTGGTCTCTCGACGCCACCA CTGACTACCGCTGCAAGGTCTCAAGTACCTCGACGACATTGACTGAT GGTCCCGGAGGACGTGGCGCCGACGTGGAGATTGTCAAGGCTC GTGAGAGTCTGCTGGCTTCTATCTGCAGCCAGTAAGAGCCGCTTCTA CATGCTCGCCACCATTCGGCTTCCATCACTGCTATCAAGTCTGATTGC CAGATGCCGTGCGGAGACGGCCCAACCGGCGATGACGGTCTTGGGAA CGCGATTGGCGTCCGAGTGGATGAAGAAGCAGCTGAACGACAAG GCAAGCAGATTGCTGGCAAGCTCAATGACATGCTGACGACCGGTGTCT CGTTCAAGCCCTCTCAGGCCATGCACTGCTGCTAGCTCATCAAC CACCTCAACGCTGCCCAAGTCTCAGTGGGCTTGTCCGAGCTCTT GAACCGTGTGCAAGGACACTTGTCAAGCCTTTTGGCAAGGTA AACCTGTCTGACCGGCTGAGCTGAAGGTCTATGACGTACGCTGGAAC AGCTCTTGAGGCTCGCTCTTACACGGGCAACACCGGGGCTTTGG CCTGACGACTTGGTTGAGCGGGTCCGGTGGCTTATGCTGCAACATG GCCAAGCTGGCCACTCGTACCTCGGCTACGATCGACTTTTGGAGG ATGCAATGAGGGCTGCCGACTTTGAGGCCACCGTGCAGCGAGCTGGCT TCCAGATGCCACGGTCAAGGACGGCGGACCCAGCGACCGAGCTGAGA TCATTGCCCTCCGCTCCAAGGGCGCACTGGACGACCTGATGGCCAAAGTG					CCGTGG TGTTCTT TGTTT GTGCTG TGATGG CTTGTA GTTGTG ATGTGT GACTGC CTTTTG GGTGTG TTGTGT TGAAAT GGCCGT ATCTCTG GTTATAC TTGGTC GTTTTGT ACGATTT TTGTTT TATGTG CGTGAT TCTTCG GCTTGT ACTTCT GGCATG ATAGAA GCCAAT GAATGT GTCTGT TCTCTG TGTTGT TTGCGT GCCGTC GTGATTT TGATGT CGGGT TGCACA	RRQLNAGAGAHPN DAVDQSLSLVRDPD LADEAWANHVTRL NRQJAKAFDADKA RAVIGNSEVQAVRDL LVPPGLTPYIASPAS TSTLAPATAVSSPTVS FTKGELPKALATKG VTDPYGWSGELLASI YRIKEHFSQVLGPRQ GSTSDPTAPSDGDAP QGPTTATGGPQVAL NKIFHHIANNTVPESI RHALCSINYTILEKAN GKFRPVGTDSIFNKV VNRALLEQQQPHIAH LLQASPELAVGVKDG ISAAVGMAGFELQAC ESTPGWTMLSLDFKS AFNYTDRARLHEIVA DKVPGLLRAFERHYA RPTTHCIVDKFFKVIDI DVGQGVQGNELSPF FFALYSCEVLGLLDAT TDYRCKVIKYLDIVL MGPAEDVAADVEIV KARASAGLHLQPSK SRFYMPRHHSASITAI KSVLPDAVRETANTG MTVLGTPIGRREWM KKQLNDKAKHIAGKL NDMLTTGVSLQALLT AMQVYPSLINHLTYL PPSLTSGLSSELLNRAC KDTFVKFAFFAKVNLS APAGAEHGHVLEQ LLEARLFRANTGGF GLHDLVERGPVAVVC

197

AAAAGAAATCCTTTACAAATACCCAGTCTACAAAAAACTAATACCTCTTTA ACACTCGAACCAAAATACTAAAGTAATACCTCAATACAAAACAAAACG ATATTATAATAGCATCCACTTCGCTAACAACTTAGCGTTTAAATCAAAAA AAGGACTACACATTAACACCTACATATTCTAGAAAAACGACACCCGTAA GCATACTGCTTCTATGAAATGACACCCATAAGATAACATCACATATA GTTGCGAAAAACTACCTGAGCTCCTTCTCAACAACAAATCATTTATT TAATGAGAAATTTATAAATGTTCCCTCTTGCTGAAATAATGAACCATC TACCAGTCCAAATAACAATGTCATGTGGGAGTATCTCATATCAACA ATTTAAATTTTCAACATCCATCTGGAAGGCAGGTAAGAATTTATAGA GAAATATTTTCAACATCCATCTGGAAGGCAGGTAAGAATTTATAGA GGAGCTAAACCTTTTGTTAAGAAAGTTTAAATCCACTTCTAGTTTAAAT CAATCGCGCTGAAAGTTACAATGATTTTGCGAATCTCTTTTGCAAAAA CCCTCCGCCAAATCAAAAGTCCAAAGAACATACATATGTTTAACTCGTAG GATTGACCTTTGGAAAAAAGGAGATACTAGTTTACTGTAAAAAGAAAGTT CGAAATATACAAAAAAATTTGTAAATTCAAAAAATAAGATCTATGG ACGATATATCTAGAAATTTGCCAAATTAATTATGGAAGGCAAAATCACT GCAGCGCTGAAATTTTAGAAAAAGAGGCATCATCCGGCATACTACCAC TATCAGACAAACACATTAAGAACCTTAAAGCAAAACCCCTGAACCTC CCGAGTAGAAGATTATAGCTTACTGTTTGTCCGATTGATTAATCCCAA AATGTTCTTCGATTGATTGATGAGCAACTAGTTATGAAAGCAGCATTT GCAACTAAAGGATCTGCTGGACCATCAGGAATGGATGCCGATATTTATC GCCGATCTTATGTTCTAAAACTTCAAAAGAAAGTAAAGAACTCCG AAAAGAAATGCTAAATGACACAAACTTACTAACAGAAACATATGAA CCAACATTTCTAGAAAGTTTCACTGCTGTCGATTAAATCCTCTAGATAAA AATCCAGGTTATAGGCCAATTGGAGTAGGAGAAATTAAGGCGTATC ATAGGTAAGTAATTAGCTGGAGTTCAACAGTGAGATAAAAGAGGCA GCCGGCCATTACAAACATGTCTGGACATGGGGCAGGAGCCGAAGCG GCTGTACATGCCATGAAAGGAATATCGACAATGTGCAAAACAGATGCAA TACTTTGATTGACGCAAAAGAACGCTTTTAAATGATGAATCGACAAGTC GCCTACACAACATCCAGATCATTTGCCATTAAATTTCAATTTACTTAATC AATACTTATCGAAATCCATCAGGCTCTTTGTGGCAGGGGTAAAGAAA TATCATCCCAAGAGGCACAACTCAAGGTGATCCCTTGTCTATGCCATG GTACTCTTGTAAACACCAGGTTATTATATAGAACACTTACTTGTAATTAACC CACAAAGTTAAGCAGGTGTGGTTAGCAGACGATGCTGCAGCTAGTGGAA GCATTGCAAACTTACATAGCTGGTATCAACACCTTATTGATGAAGGATG TAAACATGGCTACTATGTAACCAATCTAAAATGCTGGTTAATTTGTAAT CCCCCTCGTTAGCAGAGAAATGCAGGCATAGTATTTGGTAAATCGGTCAA CATAACTACAGAGGGTCAACGACATTTGGGTTCAGTAATAGGTTCCGCAA	GCTGAAAT TTAAGATC TGAGCTTT TCAATCAG AGTTTTTT ACCTAAA ACATAAA TTTTATCAT AACAAAA TCGTTCTA ATATTATT AAACTTAA AGAAATTC GTTCTTAT ATCAAAATC TTATTTCA GTGTTTCA CAGACGAA GGGTTTTA CTAGATTT TTATTTTT CAACTTTT GAATTTGT TTATTATA AAACTGTA AACTAGTG TGCAACCA ACCGTAAA AAWTAGT TAGCTGTT CACCCAAA ATATTATT CAGTATGA AAATATTT AATCTTCT TTATTCGC AGTAAAAA ATAAAAAA	STYDEIVNRRNIFNI PSGKAGKEFIEELTFW LRKFNSTSLNSIALK VTMILPNLLQKPSAK SKSKEHTLCLTRRIDL WKKGDTSLLLKEVRN IQKKEVNSKXKRSMD DISRIFAKLIMEGKITA ALKFLEKEASSGILPLS DNTLDLKSKEHPEPS RVEDYSLLFGPIDLIPK CFEDCIDEQLVMKAA FATKGSAGPSGMDA DIYRRILCSKNFIKEGK ELRKEIAKMTQNLLTE TYEPTFLEAFTACRLIP LDKNPGRPIGVGEVL RRIIGKVISWSFNSEIK EAAGPLQTCAGHGA GAEAAVHAMKEIFD NVQTDAILLIDAKNA FNCMNRQVALHNIQ IICPLISYLINTYRNP RLFVAGGKEISSQEGT TQGDPLAMPWYSC NTTIIIEHLLVNPQV KQVWLADDDAAASGS IANLHSWYQHLIDEG CKHGYVYNQSKCWLI VKPSLAENAGIVFGK SVNITTEGQRHLGSI GSQNFKNKYCTEKVA KWLTELKQLCKVAET QPQAAAFIAFTKGRS KFTYFLRTIPKFEQYLA PVDEILSHLLPTLFGK DTPFEDHIRKLTTLTP RDGGLGIPILVEEAPH

CRE	MoTe R1	JQ747 487	Magnap orthe oryzae	AATTTTAAAGAACAAATATTGCACTGAGAAAGTAGCAAAATGGTTAACCG AGTTAAACAACTTTGTAAAGTAGCAGAGACGCAACACACAGCGCGCTTT TATTGCGTTTACAAAAGGATTCGTTCAAAATTTACATATTTCTAAGAA CTATTCAAAAATTTGAACAATATCTAGCGCCGTAGACGAAATACCTAGT CATTTGTGTTGCCAACTCTTTTGGAAAAGATACGCCCTTTGAGGATCA CATTAGAAAACTTTTACATTAACCTCTCGAGATGGAGGATTGGGTATAC CTATACTAGTTGAAGAAGCGCTCACCGATTTTTTATCATCTGTTAAATTA ACTAAAACTCTGTACAGCAAAATTATAGATCAAGATAAAATTTTAAAAAC AAAAAACTCTCCGGGAATGTTCTAGAAGATCTTGAAAAAATATTAACT ACTGACAGACTTAAGCATCGCAAGAGAAAAATAATTGCAGTTGATTCAA TGCAACCAGATTCAATGTTAAGAAACATACAGCAAAAACAGAAACGAAT GTGCTAGCACCTGTTAAACGCTTACCACCTAGAAAAACCAAGTTTTGTT TTAAATAAGAAAGAAATTTGAGATGCACCTTTGCTTACGTTATAATTTTGA CTTGAAAAATATCCCTCGTATTTGTGAATGCGGAGAACCTTTCAATGTAA CTCATGCAATGTCGTGAAGAAAGGTGGCTTCAATTCAGTCGTCATGAT AACATAAGAAAAATTTGTTACCACATGCTTAAAGAGGTGTCATAAAATGT TCAATCCGAACCCACCTCATACCCCTTGATAATGAGAAATTTTATTTCCTCA TACAGCAAAATAAAAGTAAACCAAGCTGTTTAGATATAAGCAAAATGGT TTTTGGCGAAATGGACAGACAGCATTTTTTTGACGTAAGAGTTACGCATG TCAATTCATGAGCAACAAAACTAGATATAGCTGCTATATTCAGAAAA CACGAAAAGGAAAAAAGAGAGATATGGCGAACGAGTTCGTGAAAGT CGAACATGGCAGCTTGACACCACCTAGTGTGTTGGTACAAATGGAGGTATG GGTAAAGAAATGTCATCGGTTTGTTAGAAGGCTAGCGGAGAAACCTAGCG GAAAAACAAATGAAAAATACTCAGTTGTTATGACATGGTTAAGAACAA AGTTATCTTTGAGATACCTCGCTCAACTATCCTTTGTTTAAAGAGGCTCAA GAACACCCCTGGACTAAAAAACCAGATTTTGAATCGGTGTTGATTTTAA GATGGATGCTCTAGAAGCAAGAAATCTGAGCTCTTATAAAATTTATATTATA GCATTTTGTTTTA (SEQ ID NO: 1037)	TCTAGTTA AACAAAAAT ATTTCTTA ATAATAAA AAACAAAA ACTTTTTCT TAACAAGT ACA (SEQ ID NO: 1160)	QFLSSVKLTKNLVQQI IDQDKILTKNSSGNV LEDLEKILTTDLKHR KEKIAVDMSMQPDSM LRNIQQTRESEASTW LNALPLENQGFVLNK EEFDALCLRYNFDLK NIPRICEGEPFNVTH ALSCKKGGFISSRHD NIRNLFTLLKRVGIN VQSEPHLIPLDNENFY FHTANKSNQARLDIK ANGFWRNQGTAFF DVRVTHVNSMSNKN LDIAAIFRKHEKEKR EYGERVEVEHGSLT PLVFGTNGGMGKEC HRFVRRILAEKLAQKQ NEKYSVVMTWLRTK LSFEILRSTILCLRGSR PWTKKDNDFEIGVDFK MDALEARI (SEQ ID NO: 1405)
				CCCGAAACC CGAAACCCA AACCCAAA CCCAAAACC CAAAACCCA AACCCAAA CCCAAAACC CAAAACCCG GAGGGTTC CCAAAGTCG	TAATAG GTAAACG TCCCTAT TTTTGTC TTTGGTT TTGTTTT TATCTTT GTTTTTG TTTTTGT TTTCGTT	MVCPTCNGVADYN DHIRKKHPDERYTAL QLQLGLTPCPICKTA CKNDLGVKTHLSKIH KISGASKISTQPRIITE NTDNTNSVPTSSFNP VLPEIQLTPGLNNSR WADNPKRRADTPS PTRGRNTRPRFSYT DIDLNDDEPADNPRA

GGGGCTATTTATGCGCTTTAAATTTGTGCGGGCTATTAAATGCGCTTTAAC	CCTAAACC	TTTGT	NNPRVNNPRVNNEP
TTTACAAATTTATTTATGCGCTTTAAATTTGTGCGGGCTTTAAATGCGCT	CGAAGGG	TTGTTT	PSSNSLPSISEHTP
TTAAATTTACAAATTTCAITTAATGCGCTTTAACTTTTATATTTACTAATGCG	TTTAGGAT	CGTTTT	GTLPLTNSNISLKDQH
TTATTTATATAATGCTATTATATGCTTGCTATTATATTTACTATTAT	ATTATTC	GTTTTT	DKITGPILQKPLIQKLI
TATCGTTATTTATTTGCAATTTTATATATATAAACCTCGTTGTGCTCCTCG	GTTTATTA	TTTTGT	EYSKIPIEHLHARQ
ATTTATCCGTTTCTTTCCATCCATCGCGGTTTTCGTAAGCTTTGGT	GAATTGGA	TTTTGT	AKIFADAANRIAKNFI
TTGAGGATTTGCTTTCTAGGCTTTGCTTTCTGAGGCTTTCTGTCAGCTT	TAATATTT	TTTTGT	QSPTEKTLFNLLIPRI
TTACTGCTTTATTTTCTTTCTTTTATTTCCCTTTTCTTTTCTTTTACCT	ACCCCTGT	TTGCTT	FGILINGKVTIMQ
GGTTATTAGCGGTTACCTGCTTTTATACCTGCTTCCCTTTACCTTTACCTGT	TGGACAG	TGTTTT	NFPSQPIPIKIDFPSE
TTATTAGCGGTTACCTGCTTTTATACCTGCTTCCCTTTACCTTTACCTTTAT	GGGGGT	GTTTTA	KTDSDPVLNAKILLE
AAGCGTTTACCTGCTTTTATACCTGCTTCCCTTTACCTTTTATAGCGG	GACGGG	TCTTAT	KGYIGRAAKAIIDPTP
CGGTTACCTGCTTTTATACCTGCTTCCCTTTACCTTTTATAGCGG	TTAAATTA	TTTTGT	VAPETESLNILREKH
TTTACCAGCTTTTATACCTGCTTCCCTTTACCTTTTATAGCGGTTA	AGGTTTT	TTTTGT	PIQNNPNTKSQPI
CCGTTTCTATTAGTGGCAITTTATTTCCGTTTATTTATAGCAGTTAAAT	TATTATTA	TTACTT	SGRQITEKAILLAISSI
TACCTTTTAAAGGTTATTTACCTGCTTTTATTTACAGGGCACCCCTGTTT	TGCGCGT	GTTTTAT	GREKAPGLSGWTRSL
TACTAGCAGTTAAATTTACCTTTTAAAGTTATTTACCTGCTTTTATTCAC	TTATTTGT	TTGTTT	LDAAIKIPTQNDVIPA
AGGGCACCCCTGTTTTACCAGCAGTTAAATTTACCTTTTAAAGTTATTT	TACCCCC	ATATTA	LRLTDMIRQGTAPG
ACCTGCTTTTATTAACAACCTTTATTTTCTTATTAACGGGTATTTATTT	CAAAATTT	CCTTTG	RELLCASRLIGLSKPD
ACCTGTTTATTTGGAATTTACCTGTTGACGGCATGTTTGCCCAACCTG	ATAAAAGC	ATTTTT	GGVRPIAVGDLLYKIA
TAAACGGGTTTACGCCGATTACACGACCATATCGGAAAAAACACCCG	GCGTTCCA	CTATTT	FKAILNTLWSPNCLLP
GACGAACGTTATACCGCCCTCAACTCAACCATTTGGTTTAAACCCCTG	TCCTCTTA	TCCCAC	YQLGVNSIGGVEPAIF
CCCTATATGCAAAACCGCTTCAAAAACGATTTGGCGTTAAACCCAC	GGAAAA	CTTATTA	TLEEIMGPNINGIKS
CTATCCAAATCCACAAATATCCGTGTCATCGAAATTTCAACCCCAAC	CGAAGCTT	TTATAAC	ITSLDLKNAFNSVRA
GCGTATACGAACGGAAATACGGATAATACCAATTCGGTCCCAACGTCG	TTCTTGT	CCCAAC	AIASSVAKYAPTFFRS
TCGTTTAACTGCTCTTCCGAAATCCAAACGTTAACCCGGGGTTAAA	AAAAGTCG	CTACTAA	TCWAYNQPSILITEN
TAAACAGCCGTTGGCCGATAACCCAGAAAAACGACGGGCCGATACCCC	CTAGACTT	TATTTT	GSVLSAQGIRQGD
CTCCCCAACACGGGGTGGGATACACGCCACGTCGATTTTCAATACG	TTACTATA	TCTTTT	LGPLLSLAFRPTLETI
GATATCGATTTAACAAACGACGAACCGGGGATAACCCAGGGGCTAATA	AAAGTCGC	TCTTTT	QKSLPTYIAAYLDDV
ACCCAGGGTTAATAACCCAGGGTTAATAACGAACCCCTCCAGCCCC	TAGACTTT	TCTTTT	YILSKTPVKDKIAKIEK
AAATTCGTTACCTTCGATTTCCGAAATTTACACCCCTGGGACCTACCCCT	TATACCAA	ACGTT	SPFTLSAKTTETDID
AACCAATTCGAATATATCGTTAAAGACGACGACGACAAAAATTTACCGG	TCTTTTAA	TTATTT	TLKTNGLKTLGSFIGP
CCTATATTGCAAAAAACCGTTAATCCAAAAATTAATCGAATATTCGAAAT	CAAAAAAGC	CCCGTT	TELKKEFLQNKIQNFE
CCCAATCCAGAACACACCTCCACGCGCAGGCGGCTAAAAATTTTGCTG	GTAGCTTT	GTTTTT	SSINALKLPKQYGLLI
ACGCGCAAAATCGAATCGCAAAAAATTTTATACAAAGCCCAACGGAGAA	TTGTTGCC	CTATTT	LRKSTQLLRHLRLTL
AACATTATTTAAATTTACTTATATACCCCGCATATTCGGTATCGGGTTAAT	AATCTATT	ATTTGTA	NSQDLWELWEKTDK
AAACGGAAAAAGTAACTAAATATGCAAAACCTCCCATCCCAAAATACCC	AAAAAAA	CGACAA	LIADFVINLTVTKRKK
CCTATTCCAAAAATTTGATTTTCCATCCGAAAAAACCGATTCCGACCCGCT	GCGGAGCT	AACCTT	RPITDFVTPLITPIKD
TTTAAACGCCAAAAATTTTGGAAAAAGGGTATATTGGCCGCTGCGGCA	TTTTTAAC	AGCAA	GGFLLRHNGIAQDI

AAGGCTATTATCGATCCAAACCCCGTTGCCCCAGAAACCCCGGAATCGT TAAATATTTACGGGAAAAACACCCCTATTGGCCAAAATAACCCGTTTAAAT ACAAAATCCCAACCAATATCAGGCGGAAAGCTATACCGAAAAGCTATTT TATTAGCTATTTCTGCTATTGGCCGGGAAAGCTCCGGGCTTAGCGG GTGGACGAGATCGTTATTAGATGCAGCCATTAAATACCTACCCAAAC GACGTAATCCGGCTTACGACTCTTAACGGATATGATTGCGCAGGTA CCGACCGGTAGGGAATATTATGCGCTTCGGTTTAAATAGGGCTATC CAAACCGACGGCGGTACGCCAATAGCCGTTGGGACCTATTATAT AAATAGCCTTTAAAGCTATTTAAATACCTATGTCGCCAAACGTTT ATTACCTTACCAATTAGGTGTAATAGTAGTGCGTCGAACCCGCT ATTTTACCTCGAAGAGGCTATAATGGCCCTAATATTAAACGGTATAAA ATCGATTACCTCCCTCGATTAAAAAACGCGTTTAAATAGCGTATCCAGG CTGCAATAGCCTCGTCGGTAGCTAAATACGACCAACTTTTACCGTTCT ACCTGTGGGCTATAACCAACCTTCGATTTTAAATACGGAACCGGTTTC CGTCTGGCTAGTGCAACAGGTACGCCAAGGCGATCCGTTAGGCCCG TTGTTATTCAGCCTGCTTTTCGACCTACGTTGGAAACGATCCAAAATC GCTTCATATACGTATATAGCGGCTTATTGGACGACGTTTATATTTTATC CAAAACGCCCGTTAAAGATAAAATAGCCAAAATAATCGAAAAAGCCC GTTTACCTAAATTCGCCCAAAACGACAGAAACGGATATCGATACGTTA AAACCAATGTTTAAACCGCTCGGCTGTTTATTGGACCAACGGAAT TACGGAAGGAATTTTGCAAAATAAAATCAAAATTCGAATCGTCCATT AACGCCCTGAAAAAATCCCTAAACAATACGGATTGCTAATCTTCGTA AAAGTACACAATTAATCTTTACGCCATTTGCTCCGTACTTTAAATCCCAG GACCTGTGGGAATATGGGAAAAACAGATAAATTAATAGCGGATTC GTTATAAATTTAACTGTTACAAACGGAAAAACGGCCAAATACGGATT TCGTTACGCCGTTAATACGTTACCTATAAGGACGGAGGTTTGGATTA TTACGGCATAACGGAATAGCCCAAGATATTTATTTGCGGCCAAGGATT TAAACCCGAAATTCGGCACAATAATCCACGTATATCCACGATTTTCCA CAAAATCAAGCCCTACGCCACCGAGATTTTGCAATTTGTCATAACG GGTTTTAGCAGATTGCAAAAACGGGTAAACAAACGCCCAATTAACACG TTTAAACGAAAAACGCTAGTTATTTAGTCGAAAATGGCTTAACTTTTAC CTATCCAAAAATCAAAATCGATTAAACGGATTGGGAAATGGCTGAAGCCGT TCGATTAAGATTATTAGCCCCGGTTAAACCGTTAACCCACCCCTGCAACC ATTGCGGAAAATCGGACCAATATAAAACACGAGGACGTTTGCAAAAGTG CCGTACGCAAAATATACGGCCCGTCACGACCAATAAAACAGAAATTTTCGT CAATTGTTAAAAAGTCGACCAAGAAATCGACGTCGAAATCGAACCCGAT TTAAATACGAAAAATAACGTAATAACGCCCAATACCAACCCGAAAAATC CCACCCCTAGCCCCAACGGCCAAACGATACCGGATGCTTTTACAACC	TTTTCTTT TTTTTTTT TTTTCTTT TTTTTTTT TTTTCTTT TTTTTTTT TTTTTTTT TTTTTTTT TATATATA TTATTATTA TTATTATTA GCGGTGG GGCTATTT ATGCGCTT TAATTTGT GCGGGC TATTTATG CGCTTAA TTTGCGG GGGCTATT AATGCGCT TTAACTTT ACAAATTT TATTTATG CGCTTAA TTGCTGCG GGCCTGTT AATGCGCT TTAATTTA CAAAATTC ATTAATGC GCTTTAAC TTTTATATT TACTAATG CGTTATTT ATATAATT GCTATTAT TATCGTTG CTATTATT	YFAAKDLTEIRHKIQ RISNDFPQNQSPTAT EILHLLHNGVLADCK NGLTNAQLNALTEN ASYLGRKWLNILPIQK SNRLTDWEMAEAVR LRLAPVKPLTHPCN HCGNRTNINHEDVC KGAVRKYTARHDQIN RSFVNSLKSRIPEIDVEI EPDLNNENNANNAN TTTENPTSPNGQN DTGCLFTPIRSGRN GQNGLRADFAVING VSKYYVDVQIVAINKD SGNTNPLNLADAA NNKRRKYQFLDPFFH PIIISAGGLMEKDTAQ AYKQIQKLIGPVAAH WLDTSISLILRSRTTA AISIKNRPRA (SEQ ID NO: 1406)	TAAGCTT AGAATA TAATAA AGCGCG AATTAA AA (SEQ ID NO: 1284)

				CCTATTCGCTCCGGGACCCGTAACGGCCAAACGCGCTTAGGGCGGATT TTGCCGTTATTAAACGGCGTATCCAAATATTATTACGACGTGCAAAATCGTT GCAATTAATAAGGATTCGGTAATACAAATCCGTTAAATACGTTAGCAG ACGCAGCAAAATAACAAACGACGTAATAACCAATTTTGGATCCATTTTC CATCCAAATTATAAAGCGCCGGAGGCCATTATGGAAAGGATACAGCAC AGCGGTACAAACAAATCCAAATTAATAGGCCCGTTGGGCCCATTTG GTTGGATACGTCGATTTGTTAATTTGTTACGGTCCAGAACGACGGCA GCAATTTCTATTGCTAAACCGCCCTCGTGCCTAATAGGTAACGTCCT ATTTTGTCTTGGTTTGTATCTTTGTTTTGTTTTGTTTTGTTTTGTTTT TGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGCTT TGTTTTGTTTTATCTTAATTTGTTTTGTTTTGTTTTGTTTTGTTTTATTTGTT TTATATTACCTTTTGATTTTCTATTTTCCACCCCTTATTATTATAACCC CAACTACTAATAATTTTCTTTTCTTTTCTTTTCTTTTACGGTTTATTTTC CCGTTGTTTTTCTATTTTATTGTCGACAAACCCCTTAGCAATAAAGC TTAGAAATATAAAAGCGCAATTAATAA (SEQ ID NO: 1038)	ATTATTGC TATTATTAT CGTTATTA TTATTGCA ATTTTATTA TATAAACC CTCGTTTG TCCCTCGA TTATCCC GTTTCTTT CCATCCCA TCGCGCGT TTTCGTAA GCTTTGGT TTTCGTAG GATTTGCT TTCGTAGG CTTTGCTTT CGTAGGCT TTCGTCAG CTTTACCT GCTTTTAT TTTTCTTT TTCCTTTA TCCCCCC CCTTTTTT TACCTGGT TTATTAGC GGTTTACC TGCTTTTA TTACCTGG TCCCTTTT ACCTGTTT TATTAGCG GTTTACCT GCTTTTAT TACCTGGT TCCCTTTT	

HER O	HERO- 2_BF	.	Branchi ostoma floridae	TTTTCAGTCTGGCTCAGCCAGTGACCGCGGGGAAAGTCCGGCTGACTAC CACGAATAGGGTGGTGACAGCTGGATAGACAGACGACAGCTCGGAAA GACGGCATTGGGGCAGTATGGTTGGCACCCTAACTGCATCTCCCTA GGAGAGCATCCCGCAACAGCTACAAAGAACCAAGAGCAATACCC CCAGGGATGCGGAGAGGGGGAGGATGAGCATCCCATTCGGACGG TCCAATCGGTATTGACCCCAACGAAACGGAGAAATCGACAATGAATGCAGT CTGTGTGTGGCAAGGTATGTAAGAACCAAGAGGTTTGAGAAATCCA CCAACAAAGATGGCTGCTTAAGGAGGGTGAGGCGGAGCACCGCTC AGGGCTGTGGCAACCACTGTAGAACCAAGTGTGTGATCAGCATCAGCCCT GGTCAGACGAGGAGGATCAGGGCCCGGAAAGCTCCACAGTGCCCG GAACCTCCGCAACGCTGCCCTCCACAAAGGAGGAGTCAAGATCAT CACGAGTGAAGTGCCAGCGCCGAAACTCAAGGAGTGGTCGAGTTT GACGAGGACGTTGACATGATCTTGGAGTCGGTGTCAAGAGGTAGTACA GACCAAAAGCTTCAGTCCATGTGCACAGTATTATGTCATGGGGCAG AACGATTTGGCAGCATTGGGAGAGGAAACCGACAGACACAATGAAGC CAAATCGCGGGAAGTAAAGATCCGTCAACTGAGGAGGAGCTAAAGT CGTTGAGCGGAGCTTTAAGGCGAGTACGTCGGGAGAGGAGAGCT GCTCTGCAGAGCTCACACACCACCTTAGGGAGAAAGCTTAGGACCCCTCA	CTTTTATTC ACAGGGC ACCCCTGT TTTTACCA GCAGTTAA ATTACCT TTTTAAGG TTATTAC CTGCTTT ATTAACAA CCCTTTAT TTTCTCTAT TAACGGGT ATTATTTA CCTGTTTT ATTGGAAT TCACCCGT TGGACGGC (SEQ ID NO: 1161)	TGATTA AAGACC CGAAAC ACCCAA TGACCC CGGGT CATCACT GATGAT GTGTCC CTGTTT GCACTA CCAGAG TGTATTC TAGAG (SEQ ID NO: 1285)	MNAVVCCKVCKN QRGLRIHQTKMACLR RVQAEHRSGAVATT VEPVSASAPGQTEE DQGPEAPHSARNLR ATPAPPQGRKSDHH RVKWPAANSKEWS QFDEDDVMILESVR GSTDQKLQSMCTVI MSMGAERFGTIGQR KPTDTMKPNRREVKI RQLRQELKSLRRSFKA STSGEERAALAEELTH HLREKLRTLRAEWH KKKGERARKRSFAT NPFGFTKRLLGQKRS GNLTCPVEEINLHLSN TFSDASRDVDLGP

GAAGGGCAGAGTGGCAAGAAAGGGGTAAAGAAAGAGCCCGAA GCGCAGTGTCTTCATCACCAACCTTTCGGCTTCACCAAGCGACTCTAG GGCAGAAAGAGGAGTGGAACTGACCTGCCAGTCGAGGAGATCAACC TCCACCTCAGCAATACCTTCAGTGATGCTCGAGAGATGTGGATCTTG TCCTGGCCCTTGTGGTACCTTACCTGAGCCGGAAGTGCATTTGACA TCTCTGAACCACTCTGAAGAGCTCAGAGACAGACAGTCAAGCGGCGA GGTCAGTTCGGCCAGGTCCTCCAGTGGCTGTATACAAAGGTCTACA AACATTGCCACGGCTTGTGGTGGCTCTGGAGGATCCTAAAGGTGGT CTGGCGCAGAGGTAAGTGGCGGTGATTGGAGGCAAGCCGAGGGGG TTTGATCCCAAGGAAGAGGAGTCAAGTAAGGTAGACCAAGTTCGCT TAATTTCTGCTCAGTGTGAGGGAAAGATCTTCTTCAAGATTGTGGCC CAGCGTCTAATAAGTAGCTTCTGGACAAACAGATATATTGACACATCTGT GCAGAGGGGGAGTTCCTGGTGTCCAGGATGCTTGAACACACGGG CGTAGTGACCCAGCTCATCCGGAGGCTAAGGAGAACAGAGGGGACTT GGCAGTCTGTGGTGGATCTCGGAATGCGTATGTTGATCCTCCAC AAGCTTGTGAAACAGCACTGACACAGACCATGTTCCAGAGTCAATTC AGAACCTCATCTTAGATTACTACAGCAACTTCTGGCTAAGAGCTGGCTCC AGTACAGCAACTTCAGCATGGCAACGGTTAGAGAGGGCATCATACTG GATGTACGATTCAGTGCCCTCTTTCAGTACGATGAAATGATGTT AAAGGAGCGGAAGCAGGATGAGGGTCCCGTGTCTAGGTCTGGAACC AGGCAGCCCGGATCGAGCTTCATGGACGATCTGACGGTGTGACT GCAACAGTCCCGTGTAGATGGCTCTACAGGGATTAGAGCGTCTCA TTACATGGGCACGGATGAGTTCAAGCCGGCAAGTCAAGATCTCTGT CCTGAAGAAGGGGAGGTGGCTGAAAGTCCGTTTACCCTGGGAGG CACTCAGATCCACAGTGTAGAGAACCAAGTCAAGAGTCTGGGCAAG GTGTTCAACAGCTCTGAAGGACACCGCTTCAGTTCAGCAGACTAGGA GTGACTGACAACTGGCTGAGGGAAATTGAAGAAGAGGGCTACCTG GTAGCTTCAAGGCTGGATGTTCCAGCATGGAGTCTTGCCAAAGGTACT CTGGCTCTTCTGTGTACGAGTGGCATGACCATGGTGGAGCAACTG GAGAGAACCATCAGCAGGTTCTTCGAAATGTTGGGGTCCCGAGG TCCTTAAGCAACATTGCCCTGTACGGTAGATCCACCAAGTGCAGCTTCC CTTGAGTGGCCTGACTGAAGATTCAAGGTTACCCGTGCAAGAGAAAGT GTTGATGTACCGGACTCTCAGACTCCAAAGTCTCTTCAGCCGGCATC CATGTGAGGACTGGAAGAAATGGAAGGCACAGAAAGCAGTGGATCA GGCAGAGGCAAGTTGAGACACAGTGTCTCGTGGGTCCGTGGCAGT AGGACGGGCAGGACTGGCAGCTGCCAAAGCCTCGGTACGACAAAGT CAGCGGGAAGGAGAGCGTCTACTGATCCAGGATGAGATAAGGGCTG GGGAAGAGGAGGATCGCGGATGCAAGGATGTTAGGATCGGCAAGCAA	CTAGGAGA GCATCCCG CAACACGC TACAAAGA ACCACAAA GAGCAATA CCCCCAGG GATGCCCG AGAGGGG GGGAGGA TGAGCATC CCATTCCG ACGGTCCA ATCGGTAT TGACCCCA GCAAAACG GAGAAATCG ACA (SEQ ID NO: 1162)	LLVTSPEPEVHFDISE PTLKEVRETVKAARSS SAPGSGVVYKVYKH CPRLVRLWRILKV WRRGKVAADWRQA EGVWIPKEESSKVD QFRLISLVEGKIFFKI VAORLIYLLDNQYID TSVQKGGVPGVPGC LEHTGVVTQLIREAKE NRGDLAVLWLDLAN AYGSIPIHKLVTALTR HHVPESIQNLILDYYS NFWLRAGSSTATSA WQRLEKGIITGCTISV PLFALAMNMIVKGA EAGCRGPVSRSGTRQ PIRAFMDLTVMT ATVPVCRWLLQGLER LITWARMSEFKPAKSR SLVLKKGKVAERFRFT LGGTQIPTVSEKPVKS LGKVFNSSLKDTASV QQTRSDLTWLEGID KTGLPGSFKAWMFQ HGVLPRLVWPLLYE VPMTMVEQLERTISR FLRWLGLPRSLNIA LYGRSTKLQPLSGLT EEFKVTRAREVLMYR DSSDSKVSSAGIHVR TGRKWKAQEAVDQ AEARLRHSLVGSVA VGRAGLGSCPKPRYD KVSQGEKRLIIQDEIR AGEEEDRRRCRMVG MRKQGAWTRWEH ADSRKVTWPELCRAE

				GGTGGCTGGACTAGGTGGGAACATGCTGACTCCCGCAAGGTACATGG CCAGAGTTGTGCAGAGCTGAGCCTTCTCGGATCAAATTTCTCATCTTCTC AGGTACGACGCTTCCAAAGTCCAGCTAACTTGCAITGCTGGGGCTTG GCAGAGACCCCTCATGCCAACTCTGTCAGAGGAGAGGTACCTTGAAC ACATTCTAGTTGTTCCGAAAGCACTAGGGAAGGAGGTACCGCT GGCGGCATGACCAGGTTCTTAGGGTTGGCAGACACAGTTAGCAACG CCATCCAGAGTAGCAGGAGTCAGCAACCCCAAGAAAGTCAATTGTCTT TGTCAGGGCCGAGAGAAAACCCGACAAACCCACTTCGCAAGGTG GCTTCTCCACTGTAGAGATTGGCAGCTTCTAGTCGACCTTGGGAGA CAGCTCAAGTTCCAGAACACATTGTAGCCACGTCACCTCGCCTGACAT GGTACTGCTCAGAAATCCACAGACAAAGTGGTTCTGCTGGAGCTAACT GTTCCCTGGAGAGCGGATAAGCGAAGCCAAACGAGCGGAAGAGGCG GAAGTATGCCGAACGTGTAGTACAAAGCCAGAGTAATGGGTGGAGAGC CCGGTGTACCAAGTGAGGTTGGTTGCCGGGTTTCGCAGGGCAGTCT TTTGGCTTATGTGTTAAACTCCTTGAGTAAGAGGTTTCGTCCTCGGA AATCCATCAGGGATAATCTAGAGGCTGCGGAGAAAGCCTCACGTTGGTT GTGGTTCGTAGGGGGAAACCGTGAAGCCACACGGACACAGGTCGG GGAATGATCAACTCGGCTGGTGCCTCCGGCGAGGGTGTATGGTGAT TAAAGACCCGAAACACCCCAATGACCCCGGTTTCATCACTGATGATGTT CCCTGTTGCACTACCAAGTGTATTTAGAG (SEQ ID NO: 1039)			PSRIKFLISSVVDVLPS PANLHVWGLAETPS CQLCQRRGTLEHILSC CPKALGEGRYWRH DQVLRVLADTVSNAI QSSRSQQPPKKSIVF VRAGEKTRQQPTSA GGLSTARDWQLLV DLGRQLKFEHIVATS LRPDMVLVSESTRQV VLLELTPWEEERSEA NERKRAYAELVVQS QSNQWRARCVPEV GCRGFAGQSLAYVLK LLGVRGFRLRKSIRDIL EAAEKASRWLWFRFR GEPWKPHGHRSGN DQPRLRPGEGVW (SEQ ID NO: 1407)
HER O	HERO- 2_DR	.	Danio rerio	TTCAAGCC TGCGCGAG CCAGTGAC TCCTAGGA ATAGACTA GGTGGA ACCAAGAA TAGTTTG TCGACTAC TGGAGAG ACAGTTGA CGGCACG GAAAGAC GGCACTTG GGACAGTA TGGGTTAG CACCCAG CCTGTGTC	TGATCA ACCCCG GCTGGG TCACCTG GGTGAG AGTGTA TGATGTT GAGAGA CCCGAA ACACTC AATGAT CCCAGG ATACATC ACTGAT GATGTG TCCCAA ATGCAT CCATGA	MTHANEQTNKIWT CICGKLCNHWGLKI HQARMKCLEQESKV QRTGPEGETQEPPG PEATHRAKSLHVPEP QTPSEVVQRIKWP PASKGSEWLQFDED VSNIIQAIKGDADSR LKTMTTHIIFYALERFG CIEKGKTKPTPYTM NRRATQIHHLRQELR SLKKLYKKATDEEKQP LAELKNILRKKLMILR RAEWHRRRRGRERAR KRAAFITNPFQTKQL LGDKRSGRLECSIEEV NRFIETVSDPLREQE LEPNKALISPTTPARE	

GGGAGAGAGCCAGGAAGCGAGCTGCCTTCATCACCAATCCCTTTGG CTTCACAAAACAGCTGCTCGGGGACAAGCGGAGCGGTGCAGTTGAATG CTCAATAGAGGAAGTGAATCGTTTATTGAGGAAACAGTGAAGTATCA CTGAGAGAGAGGAGCTGGAGCCCAACAAAGCTTATCAGCCCCACCC CTCCAGCAAGAGAGTTCAAGTTGAGGGGGCCAAAGTCTGAAGGAGGTCA AGGAAATCATTAAGGCATCTCGTCAGCATCTACTCCAGGCCCTAGTGG CATACCTTACCTTGTCTAAGCGTGCAGGGCTTCTCCTGCATCTGT GGAAGATCTGAAGGTGATTTGGCAACGAGGAAGAGTTGCTGAGCAGT GGAGGTGCCGAGGAGTGGATTCTTAAGAGGAAAACTCGAAA AACATCAACAGTTTCGAATCATCTCTATTGAGTGTGAAGGGAAGG TGTTTTTCAGCATCGTCTCAGGAAGACTGACAGAGTTCTCCTCGAGAAC AATTATATTGACCTTCAGTGCAGAAAGGAGGATTCTCTGAGCTCCG GCTGCTTGGAACACACTGGAGTAGTTACACAACCTCATCAGAGAGGCCCA TGAGAACAGAGGGGACTTGGTTGCTTGTGGTTGGACTTGGCAAAATGC CTATGGGTCCATACCCCAAGCTGGTTGAGCTCGCTCTACACCGCCACC ACGTTCTAGTAAGATTAAAGGACCTAATTCGATTACTACAATAATTC AAGATGCGGGTCACATCTGGGTCAGAAACATCAAGCTGGCATCGCATC GGGAAAGGAATAAAGAGCTGCACCATCTCAGTTATCTTTTCGCTCT CGCCATGAACATGGTGGTCAAGTCAAGCCGAGTGAATGCAGAGGGCC CTTAAGTCAAGTCAAGTGGAGAGCCCTTATGAGCATATATGAT GACCTTACCATCAACAACACCGTCCAGGGAGAGGTGATCTTAC AAGGACTTGAGAGACTCATCGCTGGCTAGATGAGTTTAAAGCCCTC CAAGTCAAGTCCATGGTCTGAAGAGGGGAAAGTGGTGACAAAT CCATTTTCCATCTCAGGAAGTCTTGAATCCCAACCATCAGGAGCAACCTG TCAAGAGTTTGGGGAAGCTCTTGAATCCAGCTAAAGACTCTGCAGC CATCCAGAAGTCCAAAAGAACTGGAGCTGGCTGGCGAAGTTGA CAATCCGGCTGCCTGGTAGATTCAAAGCCTGGATCTATCAGCATCA ATTCTCCCCGAGTTTGTGGCCTCTGCTGATCTATGAGTCCCAATGTC AACAGTTGAGTCCCTAGAAAGGAAGATCAGTGGCTTCTCGAAAATGG TTGGCTCCACGAGTCTTACAGTGTGCACTATACGGGACAAGTA ACACTTGAGTACCATTCAGTGGCTCAGAGGAAATTCATGTTGT ACGCACAGAGAGCCCTACAGTACAGGGACTCTAGAGATGGCAAGT GTCATCAGCCTGCATCAGAGTGAGGACAGGAGGAAATGGAATGCAG GGAAGCAGTGGAGGTGGCAGAGTCAAGCTGCAACAAAAGGCTCTG GTGGCACTGTAGCGACAGGAGAGCGGGCTTGGGCTATTTTCCAAAG ACCTTAGTAAGCCAGGTCAAGGCAAGGAAAGAACACCACTACTCCAG GGAGAGGTTCCAGCAAGTGTGGAGGAAGAGAGAGTCAAGTGGTGGT AGGACTCCGGCAGCAGGAGCATGGACTAGGTGGAATACACTGCAACG	TTTCGTGA GAGAGAA CCCAACA AGCTACGG AAAGCCCC ACAGAGAT ATACCCCC AGGAGATC CCGAGAG GGGGGA GGATGAG ATCTCAA TCGGACGG ATCAAAGG TTA (SEQ ID NO: 1163)	GATGTTT CTTGCTAT AA (SEQ ID NO: 1286)	FSLRGP SLKEVKEIIKA SRSTPGSPGIPLV YKRCPLLLHLWKILK VIWQGRVAEQWR CAEGVWIPKEENSKN INQRIISLLSVEGKVF FSIVSRRLTEFLENNY IDPSVQKGGIPGAPG CLEHTGVVTLIREA HENRGDLVVLWLDL ANAYGSIPIHLVELAL HRHVPKIKDLILDY YNNFKMRVTSGETS SWHRIGKIITGCTIS VILFALAMNMMVVK AEVECRGLTKSGVR QPPIRAYMDDLTTT TVPGSRWILQGLERLI AWARMSFKPSKRS MVLKKGKVDKHF SISGVIPTITEQPVKS LGKLFDSLLKDSAAIQ KSKKELGAWLAKVDK SGLPGRFKAWIYQHS ILPRVLWPLLIYAVPM STVESLERKISGFLRK WLGPRSLTSAALYG TSNTLQLPFGLTEEF MVVTRREALQYRDS RDGKVSSACIEVRTG RKWNAGKAVEVAES RLQKALVGTVATG RAGLGYPKTLVSQV KGERHHLQGEVRA SVEEERVSrvvGLRQ QGAWTRWNTLQRRI TWNILQADFQVR FLVQAVVVDVLPSPN

HER	HERO-3_BF			TAGGATCACCTGGCGAACATCTTGAGGGGATTTCCACGTCGCTTCTAGTACAGCTGTACGATGTACTGCCAAGCCCATCAACCTCCA CGTTGGGAAAGAAATGAGACACCTTCTGCTTCTTGTCTGGAAGA GGCTCTAGAACATCTCTCAGCAGTTGCCCAAGGCTCTGGCTGATG GTGCTATCGTTGGCCATGACAGGTGCTTAAGGCAATGCTGCGAG CTTAGCTCAGCCATTACAGCAAGCAAGAACCTGCTCCAAGGAAG GCAGTCCACTTCATAAGCTGGAGAAAAACCCGGGCCCTCCACAT TAAACAAGGCTCTTCAAAAGCTCGGACTGGAGCTGGAGGTCGA CCTGGAAAAACAGTCAGGTTTCTCATCATCGCTGCAACACGCTCTC CGTCCAGACATTATAGCTATCTCAGAACTTCAAGACAGCTAATTATCT GGAGCTTACAGTCCGTGGGAGAGCGTATTGAAGAAAGCAAAATGAGA GGAAGCGGCTAAGTACCAGGAATTAGTGGAGAGTGCAGGGAGAGA GGCTGGAGAACTTACTATGAGCCCATAGAAATTGGATGCAGAGGCTTT GCAGGGCGTTCACTTTGCAAAAGTCTCAGTCGTTGGGCAATTACAGGCG TGGCGAAGAAAAAGGCCATTGATCCGCAAGCGAAGCCGAGAGGAAG GCCACAAGGTGGCTGTGATTAAAGAGGGCAGATCCGTGGACTGCTGT GGGACACAAAGTCGGGACTGATCAACCCCGCTGGTCACTGGGTGA GAGTGATGATTTGAGAGACCCGAAACACTCAATGATCCAGGATACA TCACTGATGATGTGCCAAATGCATCCATGAGATGTTTCTTGCAATAA (SEQ ID NO: 1040)	LHVWGNKNETPSCLLC SGRGSLEHLLSSCPKA LADGRYRWRHDQVL KAIAASLASAINTSKN HRAPRKAVHFIKAGE KPRALPQLTTGLLHK ASDWQLEVDLGKQL RFPHIAATRLRPDII AISEASRLIILELTV WEERIEANERKRAK YQELVEECRERGWRT YYPEIGICRGFAGRS LCKVLSRLGITGVAKK RAIRSASEAAEKATR WLWIKRADPWTAIV GTQVGT (SEQ ID NO: 1408)
HERO	HERO-3_BF	Branchiostoma floridae		CTGACCAGCAGCGGGAAGCCCGCAGCAACTAGTCTCCGCAATATTG CACACAGGGCGACCCTATGGAGCTGATTCAGTCAAAATTCCTCTGAGAT ATACCGATAACTATCTACAGAACTGCACAGTAGTTTGGAAAGAGCTT TTCTACTGAAAGACAGCAAAATCCGCCACTTAGACGAGCGTCAAGACT GCCCTCCCATAACCAATATGGCGTACCTGCTGACGTTCTGGACCAGC CAGCAGCTGGACACTGTTAATCAAGCTGGTTCATCGTCTGCTAAAGGT ACAGATGGTTTTATGTTTAACTGCCACTGCTGCTACTGATACCTG GTCTGGGTATAACAATGATGTGAAACCCCTGCTAGGCCCGCTCCACAC GAACTGGCCACAAATGAAATGTCCCCAACTAGCTGGGAGGGATTC AGTGACATCATGTGCGACTTTATGGCCAGTAAACCCAGAGTTTCAGCCACA CTACCGAAGAAAGTCACTCAGAAAGGCTATATAAGCCACGAAACACAGTC TCTCGCACAAGTAAACGCTTGAAGGCTGAAAGGCTTTCAGAGAGTCAATT AGAGCTGACGCAACACCTGAGGATCGAAAGGCTTTCAGAGAGTCAATT AAAACATACTCTTTCATGAAGCGCACACAGAAAGCAAGAAAGGAAACTACA AAATCGGCAGCACACCAAGAGAAAGAAATATCATAAGAACTTTTGAAG TTTCCCGGAAAATGTGCAAAAGGACAGCTCGATATCCCCCAGTAAAC CGGCATTCTCTGTTTATTATGCAAAATGAGTACTACAAAAACAAATACTCA CACCAACCCGTTGACTTCAACAAACTGCTCTGTTTCTCTCATTTGCC	MALPAVRSRGPASTW TLTLTLVVAAGTGDG FMSFKLPLLSTDTS GYNDVKTLGLPLHH ELATNEMSPKLAGEG FSDIMCDFMASKPEF SHTEESHSEGYISHE PQSLAQVKRLKNKLR KKAFRADATPEDRKA FRDAIKTYSFMKRQQ KRKETTKSAAHQKE YHKNFWKFAGKCAK GQLDIPPVKPAFSVY ANEYKNKYSHPTRV DFNKLLWFPHPVEE QLPANSEDMSPVRP KDIKAVLSKRCATSAP GPDGIMYGHKLHLP
				CTGACCAG CAGACGG GAGGCCG CGACCAAC TAGTCTCC GCAAAAT TGCACACA GGCGAC CCTATGGA GCTGATTC AGTCAAAAT TTCTCTG AGATATAC CGATAACT ATCTACAG AAACTGCA CAGTAGT TTGGAAG	TAGAAA CCCACA AGGCTG AGAAAT GTAGAG CATCTGT ATGGAC AATATT GATGAT TGAAAT GTTGTG ATTTTAG ATCAAA TTTAGA AATATG AAAACC GAACTA AACTAA

GGTGGAGGAACAACCTACCTCGAACTCTTTTGACATGTCACCTGTCAGG CCGAAAGACATTAAAGGAGTCTTATCCAAACGATGCGCTACATCTGCAC CTGGCCCGACGGGATCATGTATGCGCCACCTCAAGCACCTGCCAGCTTG TCACCTGTTCTTAGTACACTGTTCTCCAACTGCTTGAGTCCGGAGACC CACCAGCATCATGGTCATCTGGCAACGTGTCACCTTATACACAAGGATGG TAGTCCAGAAAGCTGCCGAAACTTTCGAATGATCTGCCTTACTCTCTGGG TCTCCAAGATTTCCACCAATACTCTCGAACGATGGGCAAAAGTACAT GACTTGCAATGATCTGATAGACCAGAAACACAAAAGGATCTCTGACC GGAATCAACGGCTGTGTGGAGCATGTCCAAGTTATGGGGAGATCTTA GCATGCCAAGAAAAACCGCCGAACAGTCCACATTACATGTTTGACC TCGCGGATGCCTTTGTTCTGTAGAACACAGTCCACTACTACTACAGAT GGAGAGAAACGGCTTCCGCCAATTATCACACGTACATTAAAAACCTG TATTCTCGCTGAAAGGGAAGTGAAGGTCAGGCTCGGAAAGTGAT CCGTTCCCGTTCCGAAAGAGAGTGTCCAAAGGAGACAACTTGTACCCCA TCATCTTCTTAACGGTGTCCAGCCTATTCTACAGCATCTCAAGGGAGTA GAGCAGCAACATGGCTACAACCTCAATGACAAGCATTATGTTACACTGC CTTTCGACAGACGACTTTGTCTATAACCAACAAAAACGACAGCATCAG AACTAAATTACTCAAAATTTCTTCCAAACAAAAAGTCAATGAACCTAAAGCT AAAAACGCAAGTGAAGTCTATGCTATGATGAGCGGAAAGCCATCG GACATCAGCTTCAATAGATGGGACCTGTCAAAACGACCAAGATG CACCAGAAAAATCCTAGGTGGCTACATCACCTTCTGAGTAAAAACAAA AGAGACCTATGACATCTAGCAAGACAAATAGAAAACGACTGTTGAAAAC ATAAAACAAATCAGCGATAGGAAGCAATACAAACTCAGGGTTTACATG GAGTACGCTTCCATCTGGAGGTACATGCTGATGGTACACGACCTGA CAGACACCCAGCTACAAAACTCGATTCATCCACACAAAGGCGATCAA AACATGGCTCAGAAATGCAACTAGTGCAACAAATGCAATCTGTACAAC ACAAGGGTCTCAACTTCAAAAGCATCTCAGACTTGACCTAGAAAGCCC ACGCTTGGCTACAGTAGTCACTCTCAAGCAGATGAGAAGGTAA AACAGGCTTTACAAGCCAACTGGACCGCAATCGCAATGGACTAGGA AAATGCAGAAATGGGTATTGGAAAGTGTCAACACCATCCACAGCAAG CCATCCATGTAGCAAAAGGACTCAGAAATGGACATCAGTACGCAAAACATGT CAAAACAAAGTCAAGATATGCGTATGACGTGGAAGTAAACATCAG GAAAACCTTCTACAGCAAGGCGAGATGCTACAACCTGCTTGAGGAAGAA AAATGGACCTGACATGGCGGTCCGCTATGTACAACCTGCCGAGGGGC ATCCTCAGTTTCGCTGTGCGTGCCTCCATCGACGCCCTCCCACTCTG TAACCTGACCACTGGGGAACGTAACACTGACAAATGTAAACTGTGT GGCAACCGGGAACACTCCACACGTTCTGAACCACTGCGGTGTCGCTC TCCAACAAAGGACGGTACACATTCCGACACAACTCGGTTTGAAGCACAT	AGCTTTTC TACTGAAA GACAGCAA AATCCGCC ACTTTAGA CGAGCGTC AAGACTGC CCTCCCCA TAACCAAT (SEQ ID NO: 1164)	ATATAAT GTTTTT TTAAAG TAATGA TAAGCA ATACCC ACATTGT GCAATA CTATCTA TGTTATG TCCTTG TCCCCC TGCATG TTTGGTC AATAAT GACCAT CGTGC CTGGC TCCGTG TACCTT CTTTACT ATGAAT AAAGAA TGATTT ACTAC (SEQ ID NO: 1287)	ACHLFLSTLFSKLESG DPPTSWSSGNVSLIH KDGSPAAENFRMIC LTSVSKIFHQILSER WAKYMTCNLDLPE TQKAFLTGNGCVHE VQVMREILAHAKN RRTVHITWFDLADF GSVEHELIYYQMERN GFPIITTYIKNLSRL KGKVKGPGWESDPF PFGRGVFQGDNLSPH FLTVFQPIQLHLKGVE QQHGYNLNDKHYVT LPFADDFCLITNKRQ HQKLITQJSSNTKSM NLKLPKCKSMSIVS GKPSDISFTIDGDPVK TTKDAPEKFLGGYTF LSKTETYDILAKTET TVENINKSAIRNEYKL RVYMEYAFPSWRYM LMVHDLTDTLQKL DSIHTKAIKTWLRMQ PSATNAILYNTRGLNF KSISDLYLEAHALAYS RSVLKADEKVKHALQ AKLDRESQWTRKMQ KWGIGKCHTHQQAI HVAKDESWTSVRKH VKQVDTMRHDVW TKHQENLLQQGQML QLLEEEKCDLTWRS MYNLPKILSFAVRA SIDALPTLCNLTTWG KRNTDKCKLCGNRET LHHVLNHCVALQQ GRYFRHNSVLKHIT

HER O	HERO Dr	.	Danio rerio	AACGGACACCATCATAGAGTCCATTGACACCTCTCGGATCAACGCCACC ATCTATGCGGACATACAAGGTTACACAACTAACGGAGGTACCATCCCGG TCCATACAATACCCACTACCCAGAAACAGACCTGATCATATATTACCA GAACAGAGACCCTCCACATCCATGAAGTACTGACCTTGTAAAGAA ACATCAAAAACAGTATGACCGAAAGGTCAACAAATACAGCACCTAGC GGCAGATTTAGAAACTGCTGGCATTCCGCTACACTAACCTGCTTTGAA GTCGGATCAAGGGGACTCGTACGCCAGAGAACAGACCGCTTAGA ACACTGTTCAAAATAGTTAAAGCCAAACCCAGAGACTCTGTTACTGA TATAAGCCGATTGGGATGTTATCGTCATATGCTATTTGAACTCACGCC ACGAACCGTATGGGAGTCTGTATGGACAATATTGATGATTGAAATGTTGT GAGAAATGTAGAGCATCTGTATGGACAATATTGATGATTGAAATGTTGT GATTTAGATCAAAATTTAGAAATATGAAACCCGAACCTAACTAAATATA ATGTTTTTTTAAAGTAATGATAAGCAATACCCACATTTGTGCAATACTAT CTATGTTATGTCCTTTGTCCTCCCTGTCATGTTGGTCAATAATGACCATCG TGTCTGGGCTCCGTGACCTTTCTTTACTATGATAAAGAAATGATTTTA CTAC (SEQ ID NO: 1041) AAAGCAGTAGAGATGACGACACATCGCGCAGAAAGTTACAACCTTCTGGT AAGACGAGGAGGAGCCAGGCCGAGGCAACCCACAGTGCCACAGAG CCTCCTAGTGTGCGCAACACCTGCTGCCGCCGCTCGCTGCTACTCAA GCTGCCCTCAAGTGACAGCAGCTCATAAAGTCCCAAGCCCAAG TCAGCAAGTGGCAGTTACAAGATCTGACTGTGTTCCCTTGGCACAGCCA AGAATCCAGTGGCCCCAATCTCAAGAAAGCTGAGTGGCTCCAGTTCTG ACAAAGGAGTGAATCAGATCTGGAAGTGACAGGCAAGGGGGGTGTG GACCAGCAGTGTCAACAATGACCAAGCTCATAGTGAACATTGCAGCTG AGCGATTCGGAAGTGTGACACCCAAACCCACTCCATCGACATATACTCCA AGCCACAGAGTAAGGAAATCAACCTCTCAGGAAAGAACTTAAGCTA CTAAAGAGGCGATCAAGGCGAGGCGGGAAGTAGAAAGAGCGGGCCT AGAAGATCTGAGAGGAATCTGAGGAAACAGCTGCTGTAACCTATGTAG GGCAGAGTATCACAGGAAGAGGCGGAGAGAGAGAGCAAGGAAAGG GCAGCATTTTGGCCAAACCTTTCAAGTTGACCAAGCAGCTCCTTGCCCA AAAGAGGAGTGGCAAACTCACCTGCTCCAAAGGAGGCTATCAACAATCAC CTCAAGGCCACTTATTCTGACCCGAATAGAGAAACACCCCTGGGGCCTT GCGGTGCACTGTGACACCACTGAGCCCACTCAGAGTTCAACATGAA GGAACCTGCGGAGTGAAGTAGAGGAAGTGTGAGGAGAGCAAGGT CAAGCTCAGCACCGGCCCAAGCGGAGTGCTTACAAGGTATATAAGA ACTGCCAAAGCTTCTACAGGCTCTGGAAGGCCCTGAAAGTCAATATG GAGAAGAGGGAAGATTGCCAGCCATGGAGGTATGCGGAGGGAGGT ACATCCCAAAAGAGGAGGAAGTCCGAGAACTCGACCAAGTTTCGAGTCA	AAAAGCAGT AGAG (SEQ ID NO: 1165)	TAGCAT GCCACTT GGACAC AGGCCG GGGTCT GATCAG CCTCGG TCGGGT CGCCTG GAGGAG GGTGTC TGTTGC AAGACC CGAAAC ACCCGTG GAGCCC AGGAAA CAACAC TGATGA TGTGTC CAAGGT TGTGCA	DTIIESIDTSRINATIYA DIQGYTTNGGTIPVH TIPTTQKPDLIYLPEQ KTLHIHELTVPFKNIK TSHDRKVNKYSTLAA DLETAGISATLTCFEV GSRGLVTPENKTRLR TLFKIVKAKPPKTLFT DISRIAMLSYAIWNS RHEPYWESETLL (SEQ ID NO: 1409) MTTHRAEVTTSKGT QEEPGPEATHSAQSL LVSPTPAAGRSPATQ SCPQVTAHNSPQSP QSQQVAVTRSDCVP LAQPRIQWPQSSKKA EWLQFDKDVNQILEV TGKGGVDQRLSTMT TLIVNIAAERFGTVTP KPTPSTYTPSHRVKEI KRLRKELKLLKRYKA AGEVERAGLEDLRI LRKQLVNL CRAEYHR KRRERARKRAAFIA NPFKLTQKLLGQKRT GKLTCSKEAINNHLK ATYSDPNREQPLGPC GALLTPPEPTSEFNM KEPCRSEVEEVVRA RSSAPGPSGVYKV YKNCPKLLHRLWKAL KVIWRRRGKIAQPWR

TCTCCTTGCTCAGTGTGGAGAGCAAAATATTCTTCAGCATGTGGCCAA AGACTCTCAACTTCTATTGAGCAATAAATACATCGACACGTCTATGCA GAAGGAGGCATACAGGAGTCCAGGCTGCCTGGAAACACACAGGCGT GGTAATCAGCTCATTAGGAGGCAAGAGAGGCGGGGGAGCTGG CTGTGTTGTTGGATCTACCAATGCCTATGGCTCAATACCCACAAAG CTGGTGAGGTCGCACTGGAGAAACATCATGTACCCAGAAAGGTGAAA GACCTCATCGACTATTACAGCAAGTTCAGCTTGAGAGTCTCCTCTGG CCAGTTAAACATCAGATTGGACCAAGTTCAGCTTGAGGAGAAATACTCTGG TGACCATCTCAGTGACCTCTTTCAGTGGCAATGAACATGATGGTCAA AGCAGCTGAGACAGAGTCAGAGGCCCCCTCAGCAAGTCCGGAGTAAG GCAACTCCCATCAGAGCTTCATGGAGCACTCACAGTGACAAACAG TCGGTACCAGGAGCAAGATGGATCTCTCAAGGGTTGGAGAGGCTCGTG GCATGGGCACGCATGAGCTTCAAACTGCAAAATCCAGATCCTTGGTG TTAGGAAAGGCAAAAGTCAGAGATGAGTTCGCTTCAGGCTGGGACAAC ACCAAATCCCATCAGTCACTGAGAGACCAAGTAAAGAGTCTCGGGAAG CCTTAACTGTAGCCTCAATGACAGAGACTCCATCAGGGAACACAGCAC TGCCATGGAGGCTTGTTGAAAGCAGTGGATAAATCAGGGCTCCCTGG AAGATTAAAGGCTGGTTTACCAACATGGAATCCTTCCAAGACTCCT GGCCTTGCTAATCTATGAGTCCCATGACTGTGTTGAAGGTTTTGA ACAAAAGGTGAGCAGTATCTACGAGATGGCTGGGATGCCACGAG CCTAAGTAACATCGCTCTGATGGAAACACCAAGCTCAAACCTCCT TTGGCTCAGTCAGGAGGAGTTCATGTGGCACGACACAGAGAACATC TGCACTCTGGATCCAGATGCGAAAGTGTCCGGGGCAGGGATTG TCATCAGGACAGGGAGAAAGTGGAGGCGAGCAGAGGCGAGTCGAACAA GCGGAAACCCGGCTGAAGCACAGGCCATCTGGGGGCGAGTAGCACA GGCAGAGCTGGACTGGAGCCTAGCAGCAACCCGATACGACTCGGCC AGTGGAGGGAGAGGCGAGGCTGGTGCAGGAGGAGTGCCTGCTTC AGTTGAGGAGGAGAGAACAGCAGAGCAGTGGCCATGCGGCAACAAAG GTGCTGGATGAAGTGGAGCAGGCGATGGAGCGGAATGTCACTGG AAGGACATCTGGACATGGAACCCCTGAGAAATCAGGTTCTTGATCCAA GGGTCTACGACGTTCTTCCAGCCCATCGAACCTGTACATATGGGGCAG AGTAGAGACACCTGCATGCCGCTGTGTTCCAAAGCCAGGGACACTAGA ACATATTTTGGAGCTGTTCCAAAGGCACTAGGTGAAGGTGCGGTATCGA TGGAGACACGATCAGGTCTTAAATCCATTGCTGAGGCAATCAGCAAG GGATCAAGGAGACAGTCGATACCGCAAGCCAGGCCAAGGTCAATTCAGT TCATCAAGGAAGGACAAAGGCCAGAGAGAACAGCAAAAGAACTGCTCT CTGGTTGCTCTCCACGGCCGAGACTGGGTGATGACAGTTGATCTTGA GAGGCAGCTAAAGATTCCACACACATCACCAGTCTACGTTGAGACCT					TCAGGA GATGTTT CTGTAA C (SEQ ID NO: 1288)	YAEGVIPKEEKENI DQFRVISLLESKIFF SIVAKRLSNFLLSNKYI DTSMQKGGIPGVP CLEHTGVVTLIREA REGRDLAVLWLDLT NAYGSIPIHKLVEALE KHVPQKVKDLIDYY SKFSLRVSSGQLTSD WHQLEVGITGCTISV TLFALAMNMIMVKA AETECRGP LSKSVR QPPIRAFMDLTVTT TSVPGARWILQGLER LVAWARMSFKPAKS RSLVLRKGKVRDEF FRLGQHQPSTVERP VKSIGKAFNCSLND DSIRETSTAMEAWLK AVDKSGLPGRFKAW VYQHGLPRLWPLLI YEVPMTVVEGEQK VSSYLRRWLGLPRSL NIALYGNNTNKLPLFG SVREEFIVARTREHLQ YSGRDAKVSAGIVI RTGRKWRAAEAVEQ AETRLKHAILGAVA QGRAGLSLAATRYD SASGRERQLVQEEV RASVEERTSRVAM RQQGAWMKWEQA MERNVTWKDIWWTW NPLRIFLIQGVYDVL PSPSNLYWGRVETP ACPLCSKPGTLEHLS SCSKALGEGRYRWR HDQVLKSI/AEAIKGI

HER O	HEROF r	.	Takifug u rubripe s	GACATAATCTTGGTCTCTGAGGCCACAAAGCAATTAATCTCTGCTGGAGC TGACGGTGCCCTGGGAGGAGAGGATGGAGGAGGCTCAGGAGAGAAA GAGGGGAAAATATCAGGAGCTAGTGGAGCAATGTAGGGCGAATGGAT GGAGGACCAGGTGCATGCCAGTGAAGTGGCAGTAGGGGATTGGC AGTACACCTGAGCAAGGCTATGTACACTGGGAATAACAGGCACA AACGAAAGAGAGCCCTAAGCAACACGTGAAGCAGCGGAAAAAGC ATCCAGATGGCTCTGGTTGAAGAGGGGGAACAGTGGGGCAGTAGC ATGCCACTTGGACACAGGCGGGTCTGATCAGCCTCGGTGGGTCGC CTGGAGGAGGTGTCTGTGCAAGACCCGAAACACCTGTGAGCCCG GAAACAACACTGATGTGTCCAAAGTTGTGCATCAGGAGATGTTTCT GTAAC (SEQ ID NO: 1042)	AGACTAGG TGACAAAC AAGAAAC GTTWGGT CGACTACT GGAAGA CAGTTGGC AGCTCGGA AAGACGG CACCCGGG ACAGTATG GGTTAGCA CCCCAGCC TGTATCTT TCGCGAGA AGGAACCC AAACAAGC TACGGA GCCCTACA GAGAAAC ACCCCCAG GAGATCCC GAGAGGG GGGGAGG ATGAGATC	TGATCA CCCCGG CTGGGT CGCCTG GGCGAG GGTGA TGATGT CGTGAG ACCCGA AACACC CTATGA ACCCAG GATACA TCCTGAC GATGTG TCCCAGT GCATCC AGGAGA TGTA TTAAGT (SEQ ID NO: 1289)	KDSRYRQATAKVIQFI KEGQRPERTAKNCSA GLLSTARDWVMTVD LERQLIPPHITQSTLR PDILVSEATKQILLEL TVPWEERMEEAQER KRGKYQELVEQCRA NGWRTRCMPVEVG SRGFASYTLISKAYGTL GITGTNRRRRALSNNV EAAEKASRWLWLR GEQWGG (SEQ ID NO: 1410) MTPAMEMTTVTCT CSKLCKNQRLKIHQ ARMKCLEREVEVQR TGPGPGETQEEPGQ EATHRSQSLHVPEPP NPNRVWQQQRIKWP PANRRSEWLQFDED VSNIIQATAKGDVDS RLQAISTIIVSYGSEF GRIEKGNTETTSYTM NRRFKIHQLRKELRT LKKQFKRAXDGDQKQ ALKELYNILRKKLTLR RAEWHRRRGRERAR KRAAFIANPFRFSKQL LGDKRSGRLECSREE VNRFLQNTMSDPLR GQDLGNRALISPAP PSAEFKLAEPSSLKEVE EVIKAARSASSPGPSG VPYLIVYKRCPEILRHL WKALKVIWRRGRVA DQWRCAEGLWIPKE EDSKNINQFRTISLLS VEGKVFESIVSRRLTE

CCAGTGGAGGTGTGTGAGGGACTTTGGATACCCAGGAGGAGGACTC GAAAAACATCAACCAGTTTCGGACTATCTCACTACTGAGTGTGGAAGG AAGGTGTTTTTAGCATCGTCTCCGAAGACTGACCGAGTTTCTCTCAA GAACAACTACATCGACACTTCAGTGCAGAAGGTGGGATCCCTGGAGT CCCCGGTGTCTAGAGCAATGGTGTAGTACACAGCTCATCAGAGAG GCCATGAGAGCAAGAGAACTAGCGGTTTTGGTTGGACCTGACT AACGCTACGGTCCATCCACACAAGCTAGTTGAGCTTGGCTACACC TACACCATGTTCCAGTAAGATCAAGGACCTGATCTGGATTACTATAAT AATTGAGGCTCAGGGTCACTCAGGGTCAGTAACTCAGACTGGCATC GCCTGAGAAAGGAATAATAACAGGCTGTACCATCTCCGTCGTTCTTC GTACTGGCGATGAATATGGTGTAAAGCGGCTGAGGTGGAGTGCAG AGGCGCTCTATCCAGATCAGGTGTTGACAGCCCCCATAAGAGCCTAC ATGGACGACCTTACCGTCAACAACAACATCAGTCCAGGGGTGAGGTGGA TCTTGACGGGTTTGGAGAGACTCATCCTATGGGCTAGGATGAGTTTAA GCCACCAAGTCAAGGTCCATGGTACTGAAGAAGGGGAAAGTGGTGA CAAATTCGATTTCTCAATCTCAGGAACCGTAAITCCATCGATCACGGAGC AACAGTCAAGAGCCTGGGAAAGCTTTGACTCCAGCCTGAAGGACAC TGCTGCTATCCAGAACTACGGAAGAGCTTGGAGGGTGGCTCACTAA GGTGACAAAGTCTGGCTGCTGGTAGATTTAAAGCCTGGATCTACAG TACTCCATCTTCCAGAGTCTGTGGCTCTCTCTGTGTGAGTCCC AGTAAACACAGTGAATCCTTTGAAAGGAAGATCAGCAGCTTTCTGCGC AGATGGTGGTCTTCTCGAGCCTCAACAGCGCTGCACTGACGGGA CAAGTAACACCTGCAGTACCTTCACTGGGCTCACTGAAGAAITTA GGTGCCACGCAAGAGAGCCCTACAGTACAGAGACTCCAGGGACTG CAAGGTGTCATCAGCCGGGATTGAGGTGAAGACAGGAAGGAAGTGA AGCAGAAAAGGCAGTGGAKTGGCTGAGTCACGCTAAGGCAAAAG GCACTAGTTGGGCGTGGCAACAGGAAGAACAGGCTTGGGCTACTTC CCAAAGACCAAGTACGCCATGCCCGGGCAAGAGAGAAACCACTA CTTCAGGAGGAGTCCGAGCAGCGGTGGAGGAAGAGCGAGTGGGTAG GGCAGTGGGACTCCGCGAGCAGGGGGCATGGACAAGGTGGGAGAGC GCGTTACAGCGCAAGTTACTGGTCAAACATCATGACGAGCAGACTTCC ACCGGTCCGGTTCCTTGTGGCGCAGTCTACGATGCCCTCCCGAGCCC AGCAAACTCCATCGTGGGGAAGAGTGAACACCCACCTGTTCCCTT TGCTCCGGAAGAGGCTCCCTGGAAACATCTCCTTAGCAGCTGCCAAAGT CCCTGGCTGATGGTCTATCGTGGCGCCACGACCCAGGTACTCAAAGC AGTGGCTGAGAGCATAGCTTGGCCATTAGCACCAACCAACCATCAT GCTCCGAAGAAGGCAATCTCTTCATAAAAGCTGGAGAGAGACCTCGTG CAGGCCACAGATAACAACGGGACTCTCTCCACACAGCTMCTGATTGGC	TCCAATCG GACGGACC TAACGTTA (SEQ ID NO: 1166)	FLKNNYIDTSVQKG GIPGVPGLEHNGVV TQLIREAHESKELAV LWLDLTNAYGSIPIK LVELALHLHHVPSKIK DLILDYNNFRLRVTS GSVSDWHRLKGLII TGCTISVVLFLAMN MVVKAEEVECRGPL SRSGVRQPPIRAYMD DLVTTSVPGCRWI LQGLERLILWARMF KPTKSRSMVLKKGKV VDKFRFSISGTVIPSIT EQPVKSLGKLFDSLK DTAAIQKSTEELGGW LTKVDKSGPLGRFKA WIYQYILPRVLWPLL VYAVPVTTVESFERKI SSFLLRWLGLPRSLNS AALYGTSTNTLQIPFS GLTEEFKAVARTREAL QYRDSRDCVKSSAGI EVTGRKWKAEKAV XVAESRLRQKALVGA VATGRTGLGYFPKTQ VSHARGKERNHLLQE EVRAGVEEERVGRAV GLRQQGAWTRWES ALQRKVTVSNIMQA DFHRVRLVAAYVDA LPSPANLHAWGKSET PTCSLCSGRGSLHLL SSCPKSLADGRYRWR HDQVLKAVAESIALAI STXKHHHAPKKAISFI KAGERPRAGQITTG LLHTAXDWQLHVDL

HER O	HEROT n			AACGACGTTGACCTGGGAAACAACTGATATTCCTCCAGCACATCGC AACACGCTCTACGGCCAGACATGATCATCATCTCAGAGGCTTCGAAA CACTGATCATCTGAGCTTACAGTGCCTGGGAAAGAGCGGATTGAG GAAGCCAAAGAAAGAACTGCTGCAAGTATCAGGAGCTGGTGAGGA GTGAGGGGCGGGCTGAGGACCTCTACGAGCCCATAGAAAGTTGG CTGTAGAGGCTTTCAGGAGCTCCTCTGCAAGCTTTGGCCGACTG GGAGTCAGGGACAGCCAAAGAGAGGGCCATTAAKCCGCGAGTGA AGCTGAGAGAGAGCCAGAGTGGTGGTGGCTKAAAAGGGCAGATCC GTGGTTGCTACTGGGACACAAAGCGGGTCTTGATCACCCCGCTGGG TCGCTGGCGAGGGTGTATGATGTGCTGAGACCCGAAACACCTATG AACCAGGATACATCTGACGATGTGTCCAGTGCATCCAGGAGATGTA KCTTAAAGT (SEQ ID NO: 1043)	AGATTGGT CTGGCTAA GCCAGTGA CGTCCAGG AACAGACT GGCTGACG ACCACGAA TAGAGTGG TGACAGCT TGGATAGA CAGCTGAC AGCAGGG AAAGACG GCAACCGG GGCAGGA AGGGCTA GCAACCCA GCTGCTAT CTTCCGTG AGGAAGA ACCCAAAA CTTGCTAC GAAGAGC CCGAAGCA AAGATACC CCCAGGG	TGATCA CTCCA GTCCGG TCGCCT GGGTGA GGGGT CTGATG TTGAAA GACCCG AAACCC CCGATG ACCCA GGTACT ATCACT GACGAT GTGTCC AAGACA TGATC AATAGG TGATTT AGAAAT C (SEQ ID NO: 1290)	MATTQASVKPTAVA TCVCGKICKNPRGLKI HQTKMGCLASVQPE QRARFSLSESREVPA RAEPYGPQQPHSPEA LGETQERGOESPHS AQNLAQVAQAPDN PQHRRVKWPPASK VSEWQQLDEDLEGIL ESTAKGGVDRKLQT MTTLVSFATERYGT MEKRAAPEKYTKNR RAEKISQLRQELRVLK KQFKGASEDQKPGLA ELRCTLKKLLTLRRA EWHRRRAKERAKKR AAFLANPFGFTKQLL GQKSAHLECAKEEV DSYLHDTFSDAEREN SLGECRVLISPPPEAC SFNTKAPTWEIQT VRAARNNSAPGNG VPYLKYKRCPKLLARL WKILRVWVRRGKVA HQWRWAEGVWVP KEEKSTLIEQFRTISLL
	Tetraodon nigroviridis			AGATTGGTCTGGCTAAGCCAGTGACGTCCAGGAACAGACTGGCTGACG ACCACGAATAGAGTGGTGACAGCTTGGATAGACAGCTGACAGCGGA AAGACGGCAACCGGGCAGGAAGGGGTAGCAACCCAGCCTGCATCTTC CGTGAGGAAGAACCCAACTTGTCTGCAAGAGCCCGAAGCAAGATA CCCCAGGGGAGCCGAGAGGGGGGAGAAATGAGCTCCCAACACGGA CGGATAACATGGCAACGACCCAGGCTAGCGTTAAACCCGACAGCGGTG CCACATGTGTGTGGCAAAATCTGCAAAACCCACGAGGTCTGAAGAT CCACGAGACCAAGATGGGTGCTTGGCAAGTGTGCAACCCAGAGCAGCG CGCAAGTTGAGCTCAGCCTAGCGAGTGGGGAGGTGCCAGCCAGGGCCGA GCCATGGCCCTCAGCAACCGCATCTCCTGAGGCCCTTGGTGAGACG CAGGAGGAGCGGGCCAGGAGTACCCACAGTGCCAGAACCTCCGT GCTCAGGTAGCAAGCGCCAGACACCCACCAACCCAGCGCTTGATGAG AAGTGGCCCCCAGCCAGCAAGTGGCGAGTGGCAGCAGCTTGATGAG GATTTGGAAGGTAATCTGGAGTCCACCGCAAAAGGTGGAGTAGACAGA AACTCCAAACAATGACACGCTGGTATCAGCTTGGCCAGCAGAGAT ATGGTACAATGGAGAAACGCGCTGCTCCAGAGAAAGTACCAAAAACC GCAGGGCAGAAAAGATCTCCAACTGCGGCGAGGAATTCGGGTCTGA AAAAGCAGTTCAAGGGCGCCAGCGAGGATCAAGAGCAGGATTGGCA GAGCTTCGTTGACCTTAGGAAAAAACTGCTTACCTTCGCGGAGCAG AGTGGCACCGGAGACGGGCCAAGGAAAGAGCCAAAGAAACGCGCTGCA TTTTAGCCAAACCTTTGGGTTCACATAAACAACTTTAGGCGCAGAAAGCG TAGGCCCCACTTGGAAATGTGCAAAAGAGGAGGTGATTCCTACCTCCAC GACACATTCAGTGACGACGAGACGGGAGAACAGCCTAGGCGAATGTAGA GTGCTGATCAGTCCACCTGAGCCAGCCTGCAGTTTCAACACCAAGGCTC CAACTGGAAAGAAATCCAACTGTGGTCAGGGCTGCAAGAAACAACT CAGCTCTGGACCCCAATGGAGTCCCATATCTGGTGTACAAAAGATGCC			

215

NeS L	LIN9_S M	.	Schmidt ea mediter ranea	GGCCGGTGAAACACCTCCTCCCAACCAAGGGCCCAACACGAGCCTCTT GCAACAGCTAGGACTGGCAACTAAGGGTTGACCTGGGAAACAATT AAGTCCCGGAAACATCGTAGAAACCAACCTGAGGCCAGACATTGTC TGCACTCACAGTCGTCAGCAAGTATTGCTGGAGCTGACTGTGCC TGGGAGGAGAGAAATGGAGGAAGCGTATGAAAGGAAGGAGGGAAGT ACGCTGAGCTGGTGGAGATTGCCGACAGCAGGGTGGCGCAGTAGA TGCTGCCATAGAGGTTGGGGTAGGGCTTTCAGGGAACTCACTC TGCAAGGCTTTAGCCTCTGGCATCACAGGCATGCGCAGGAGGAAA GCCATCTGCGGGCTCAGAGCTCAGAGAGGCGCTTCTGGACACAGGC TGGATCCAGCGGGAAGCCGTGGACGAGCGCTTCTGGACACAGGC GGAACTGATCACTCCAGTCGGGTGCGCTGGTGAGGGGTCTGATG TTGAAAGACCCGAAACCCCGATGACCCAGGTACTATCACTGACGATG TGTCGAAGACATGCATCAATAGGTGATTTAGAAATC (SEQ ID NO: 1044)	AAACGACA TCATGAAC GCTTGGCC GCAACAAT CCAGTTAT CCCTGCGG TAACATTG TGGAACTC ATAAGACA AGTACTAA AAGAAGA ATTAGAAA AATTAGAA GAAAAAAT TGAAAAATA ATTTATTTA TAAAAATTT AAAAAATTT AAAAAAT TAAAAAT TTAAAAAT TAAAAATTT AAATTTAA ATGAAGAT AAAAAATTT	TAAAT GGCAAA AAGATA TTTCAAG ATGAAT TGTGGA CTCATCT AAAAAA TGACCA CCTTGA GTCCAA ATATGC CTAGCT ATCATG GTTGCT GATGGA AACAGT AAGGCA CCTGAT AGCTAA CTTTTCA CTGTGA ATATCTT CAGATA	MMDSRQLNTPKIRK YQNPKMTNDIMKSY NYAVLSDVTPQETTQ TTTHLNVDDIDNETTQ PKQPLTSGKPKSKPI AVSYKFDATFIWDT TPQTNPPRDCTKLID KTRPKRTFKKSAFQS YLKELSNETFEVKT FLMATHKYRFKDENS RLLAYRIINRYVMETA NEFKETEFDMARFAK FFTIPENWLHLKPY TATETSPADRIKVQKL VDLTCRYPFKTQEEQ TSVANLHFFTQRSII GISRDYKFQKFIPIFMA RKNTRPETTSTMTVT SPTEQNRLPMVIIPL EETKSEHRRPEKRG SNDTIVLSDEEFLLK RRTLPTRKSKNPTGA GNVPTETECTDEVKFI LNNEYQIECKEKGK

TTCCCACTACTTAAAGGAGAACTCTTCCAAACCAGAAAAATCCAAAAATCC	ATTAAATC	TTCA	WENVRNGLNHLRQK
TACTGGTGAGGAAATGTACWACAGAAACCGAATGCACATGATGAAGT	CAATAAAT	GTGACA	HDFPNRTDVMVSCV
TAAATTCATCTCAACAATGAATACCAAAATAGAAATGTAAGAGTGTGGA	AATCAAGA	CGAAAG	RCEVPIKGAECVNHK
AAAGTGTGGAAAAACGTACGAAATGGATTAAACCACCTTCGTCAAAAAC	AAATCAAG	GACACC	NHKDDKEESEAGSL
ACGATTTCCAAACCGAACAGATGTTATGGTATCTTGGCTAAGATGTGA	AAA (SEQ	ACTAGT	VANTQDIPNESSLSQ
AGTACCGATCAAAGGAGAGAAATGTGTAATCACATTAAAAATCACAAA	ID NO:	AAAAAC	AAIEVYLRNLIKMKEN
AAAGATGACAAAGAAAGAGTGAAGCGGGAGTCTTGTGGCTAACACT	1168)	CACTAG	QERNIQYLEPSTANFL
CAAGACATCCCAATGAAAGTAGTGTGCTGTCACAAAGCCGCAATCGAAG		TTTTTC	INRNIRAFYQNVKIEK
TATATCTGAGGAATATTCTGAAATGAAAGAAACAGGAAAGGAATA		TGACAC	LIGWEQVWLHWN
TTCAATATCTTGAACCTAGTACTGCGAATTTCTCATAAATAGGAACCTC		CTCTTG	KCHWVYLANCDSKT
AGAGCATTTTATCAAAACGTCAAAATCGAAAGCTTATCGGATGGGAAC		TACAAA	SVILSDNQMTLQQ
AAGTCATCTGGCTTATACATTGGAAACAAATGTCAITGGATTGTATACCTA		CTCTGTA	RCNIKAKFDKFLGTF
GCTAATTGCGACTCAAAAACCTCTGTATCTTGGACTCTGACAACCAAAAT		AAAATC	EECTVLGTLERKVPQ
GACATTACAGCAAAAGATGTAACATAAAAACGCAAAATTTGACAAATTCCTA		AAAAAG	QPNNFDGCIWVIQIS
GAAGGTACCTTTGAAGAAAAACAGTGCTTGGAAACCCCTAGAAAGAAAA		ATCGAT	DFLKDQQRIDYHTPD
GTTCCTCAGCAACCAACCACTTCGATTGGGTATATATGTGATACAATA		AGGCCG	SKRIRKEIGELILEEMK
CATCAGCGACTTTCTAAAGACCCCAAAAGAAATAGATTATCATACACCCG		CGCTTTC	NPASKIKNPKNKEIQSL
ACTCAAAAGAAATTAGAAAAAGAAATAGGAGAAATTAATATTAGAAAGAA		ACGGTC	LQKFRLLQINVDVF
TGAAAAACCTGCCTCAAAAATCAAAAATCCAAACAAAGAAATACAATC		TGTATTC	HWFAAEYQKSLPKIR
TTTACTCAAAAATTCAGACTACTGCAATCAATGTAATGTAATGATGATTC		GTAATC	TKRDGKLNKLSQVQI
ATTGGTTGCGGTGAATACCAAAATCTCTACCGAAGATACGTACCAA		AAGATC	QRLFGAPKRAVKEIY
AAGAGATGGAAGAACTGAATAAATAAGCTGCTCCTATCAATCCAAAGA		AAGGAA	FQETSTADLETRVLN
TTATTTGGTCTAGCTCTAAAAGAGCAGTCAAGAAATATATTCCAAAG		GCTTTTC	EHFKDESTMKECKI
AACCTCTACAGCAGACTTGGAAACAAGAGTTCTAAATGAACATTTCAAA		CCCTTTT	KNGNHQDWTITKAQ
AAGGATGAATCAACGATGAAAGAAATGTAATAAAAAATGGAAACCAT		AGTCAA	IDNKEILEALKNSTDS
TACCAAGACTGGATAACAAAGGCCCAAATGATAATAAGAAATATTGG		CACCAG	APGEDNIPLRQWIIW
AAGCCCTAAAAACAGTACAGATTCTGCCCGGAGAGAAATAACATTC		GTTTCTG	NNDGVLDFDMFNYIK
TCTGAGGCAATGGATAATCTGGAACAACGACGGTGTCTCTTTTGATATG		TCCTAGT	RTHDIPDMWKNYTT
TTTAACATACATCAAAAGGACACACGATATCCAGATATGTGGAAAAACT		TGAGCT	TLLKPGKSQESNIPA
ACACCAACACTACTTATAAACCCGGAAGAAAGCCAAAGAAAGCAACAT		TCCCTTG	NWRPISILPYSYRIFM
CCCCGCTAATTGGAGGCCAATATCGATATTGCCAACCAAGCTATCGTATAT		GGACAT	KVLNKRVLWANRG
TTATGAAAGTCTCTAAATAAAGAGAGTACTAGAAATGGGCTAATAGAGGAG		CTGCGT	ELISKWQKAVDKAN
AACTGATATCAAAATGGCAGAAAGCCGTAGACAAAAGCTAATGGATGTG		TACCAT	GCDEHSYVIAALIEKA
ATGAGCACAGCTATGTACAAAGCGCTTATCGAAAAAGCAAAACAGAA		TGACAG	NRSYKNEQCHLAFL
GCTACTACAAAACGAGCAATGTCACTCGCCTCTTGGATTGGCAGA		ATGTAC	DLADAFGSPFQVIW
TGCTTTTGGAAAGCATCCCATTCCAAGTAATATGGCATACCCCTAAAAATA		CGCCCC	HTLKNMGMDEETIN
TGGGTATGGATGAGGAAACCATCACTGCTCAAGAAATCTACAAAGA		AGTCAA	LLKEIYKDCSTKYKCG
TTGCTCCACAAAATATAAATGTGGAAGAAATGAGTCAGAAAAAGATCAAA			KNESEKIKITKGVRRQ

ATTACGAAGGAGTCCGACAGGGATGCCATTGTCGATGACCCCTCTTCA GCCTCTGTATACAAATATCTTATACAAAGGCATAGCAGAAAAAGAAAAAGG AGCAACAATTGCAGGTCAAGAGTTTGCATATTGGCTTATGCGGACGAC CTAGTAATTGTTGCAAAACACAGCAAAAGACATGCAAAATGCTGTTAACAA CAATCGAAAATCTGGCAAAACAAAGCCGATCTCATATTCAAACCGGCAAA ATGTGGATATTACAGAGACCCCAAGAGATAAAAAGTCCATGATGAAGAT ATATGGCAAGAAAATCAGCATAGTAGACGAAAAGAAATGTTTACACCTAC CTAGGTGAAGAATCGGTGACACAAAGAAAAGAACCTAAATGTCAGA TTCGAAGAGGTCAAAAAGAAAACGACAGCAATCTTCAAAATCGAAATTGC GAAGTGACCAAAAATAGAGGCATACAAATCTTTTGCCAATCAAAATT TGTTGATCATCTACAAGGCGAAGATATCGCAAAAACCAAAATTGAAACT TAGACGAAGAAATCAAGAAAATGATAAAAAGAGATATATTTAAATTTAC AAGACAAAAGTCCGTTACAGACTTCGTTATCTACTCCCAAGAGAAAA AGGGGGTTAGGAATAACAAAGATAATAGATGAACAAACAAATTCAAAC TATTAATAGAACGGCAAAACTCTAAATAGTAGCCATAGAGCAATCCGG GCTATTATTTATGAAGAGCTAATACAAGTAGCTAACCTAAGAGGAGAAA AAGAAATCAACACCATTTGAAGAGCACTAAAATGGTTGGAAGGTACCA ACAAATACAAAAGAACTCCAACGCAAGACCACTGGGATAACAAGGG TTCGGGAGGCTTTCAAACCTAGAAAAGAAAACACAAAATCAAGGTTAG ATTTGTCCCAAGAAAATGCAATGGATATAAAATCAAATCGACACC CAAGAAAAGATAGTGAGCTTGATAACTCAAAAGAGTTATCAAAAAGC TTACACTGGATGATAAAAAGAGGCATATTATAAAGAAATGGAAGCCCTAA AATGCCAAGGATATATTAAAGCTAAAACCTCCGAATTTATGGAGTG GAAATGCCAGAGGCTTCGGACCTGATTGGAGATTCTTAACAAAA GTAAAGGCAAAATATGTTGGACGTAAACATGAACAAAGCCCAACGAGGA GGAAGTTGGGAAGCAGCAAAATGCCGAAAATGTGAAGATAAAGAAATC GGCAAGCCATGTTATAACCACTGTCCCTCAGGTAACTGGAGTAGAGTG GAAAGCAACACAGGTGCAAAATGAGCTAGCAAAAGAACTGACAAAG CGGAATATCAGCTTCGAAAAGGACAGCATCCCAAAAGAAAACAAAGAG AGCCTAAGACCAAGATTTGTTATAAGACTCAAAAGACAAGATAATGATAG TGGACATCAAAATGCCATTTGATGAGGAATCTGTATCGAGAGTGCCAG AAACAAGAACATAGACAAATATCGAGAACTGGCCAAAGAGATCCAAGC AAAAACTGGGTTACAAACCAACAGTCTCAACTTTCGTTGCTGTTCTTTGG GAACCTGGGATAAGAGGAACCAACGAGCTCCTACGGCAGATGGGAATAA GATATGAAGAAATCCAAAGAGATGAGGATCAATATGATCCAAAAGGCCA TCCACGGGCTAGAAAAACCTACGACCACCAAGAAAATTTTAACAATGG TTAAAAATGGCAAAAAGATATTTCAAGATGAATTTGTGGAATCATCTAAAA AATGACCACCTTGAGTCCAAATATGCCTAGCTATCATGTTGCTGATGG					ACTCCC ACCTGA CACTGTC CTCAAA ACAGTT CAATTG CATCCG AAGATC GCAATTT TTTCACT AAAAATA AATTAA CAAAAG TTAATTA TACTGCT TCATTGA GTAAGT AGAAAA ACAATC (SEQ ID NO: 1291)	CPLSMTLFLSLCIQYLI QGIAEKKKGATAGQ EVCILAYADDLVIVAN TAKDMQMLLTIIENL AKQADLIFKPAKCGY YRDPDRKKSMKIY GKEISVDEKNVYTYL GVRIGDTKKKDLNVR FEEVKKKTTAIFKSKLR SDQKLEAYNIFCQSKF VYILQGEDIAKTIETY DEEIKMKIKEDILKLQ DKSPFTDFVIYSPREK GGLGITKIIDEQTIQTI NRTAKLLNSSHRAIR AIIYEELIQVANLRGEK EINTIEALKWLEGTN KYKNSNAKTTWITR VREAFQTLKKHKIKV RFVPKENCIGYKICD TQEKIVELDNSKELSK SLHWMIKEAYYKEW KALKCQGYIISLKTSEF MEWKMPRGLPDPD WRFLTQVKANMLDV NMKANQGGRLGS TKCRKCEDKESASHVI NHCASGNWSRVEKH NQVQNELAKELTKR NISFEKDSIPKETKESL RPDLVIRLKDMMIVD IKCPFDEESAIESARN KNIDKYRELAKEIQAK TGLQTTTSTFVVCSL GTWDKRNNELLRQ MGIRYEESKEMRIN MIQKAIHGSRKTYDH

NeS L	NeSL- 1_CBr e	.	Caenor habditis brenner i	AAACAGTAAGGCACCTGATAGCTAACTTTTCACTGTGAATATCTTCAGAT ATTCACAGTGACACGAAAGGACACCACTAGTAAAAACCACTAGTTTTTC TGACACCTCTTGCTACAAACTCTGTAAAAATCAAAGGATCGATAGGCC GCGCTTACCGGTCTGTATTCGTACTGAAAAATCAAGATCAAGGAAGCTT TTCCCTTTTAGTCAACACAGGTTTCTGCTCTAGTTGAGCTTCCCTTGGG ACATCTGCGTTACCATTTGACAGATGACCGCCCGCCAGTCAAACTCCCCAC CTGACACTGCTCTCAAAACAGTTCAATTGCATCCGAAGATCGCAATTTT TCACTAAATAAATTAACAAAAGTTAATTACTGCTTCATTGAGTAAGT AGAAAAACAATC (SEQ ID NO: 1045)	AAGACGCTGGTTAAGGCCGAATTCGTCTGTTCTTTTCTGGCGGTCTT GCTTTGAGCTTGGTTCCGATCTATGCCCTTGWGCATCAGCGATTGCG TTCATCTGGTCTCGGCCGAAGGAGATACGATGAACGGGAGGTCCACTT GTGGCCATTGCTCGTTTCCTCTCTGTCGTAAAGTAGGTGAGGTCTTCC CCTTCCCCAGTGTCCCCCACCTTCCCCCAGTATCGGTCCAGATACA GGATTGCGGCTGGAATCATCGGCACATCGAGAGGATGTAGCCTTTGGT TGCCAGAGGTGGACAATGCCATTACAGTGGCTGAGAAAAAGGTTGG AACGAGACCATGAAGTTCTGGTTGTGGATTTGAGGAGCAAGCCACT GTCACTTCCAAAGCTAGACTCTAAAGAAAGACCCCAAGGAACACTGGT GTGGTTAGGCACATATTAGAAATTTGACGGAAGTTAGTTCTATCTAACT GTAAAGACCGAGTGTCTTCTACTTTGKACGTGAMGWGGCTG TCGAAGTAGTCAGGATATCTCAAATGTAAACCCCGTGAACCTGTGA ACCAAATGTGTTCTTCTATTTATGCAAGTAAAGATAGTCWGGATAT CATTGAATGTGAACWCCTGTAACCTTTCTTCGAGAGGATGTACCCGAA ATTGAGAACCTATTCTTCTGTTTCGAGAGGATGTACCCCAAGA GACCTATTCCTTCAATTTATTTGTTGGATGTATGACCTGAGACAAAGGAT GACCTGGACSTTGCAAGWAATCCCTTTCTCMAGAGAACAAAGA CAGCTGGGACGCCAGCTGTCAAAAGAAATGAGATAAGATCTATGAG ATCTTATCTGAAGAGAAATGTGAAGGAAACCCGCTGAACCTTTTGA AGTTAAGAGGTGGTGTGAAGGAAAGAGATGATCAGAAAGTTGGT TGCAAAAAGAAAGCGACACAGAGGCTGTCTGCAGGATCTGTACCC ACTTGATGATCGTTATGAGTGTGTTGTTGATGTTGTGAGACAAACATCAA CGATGGGATACGGGTCTAGTACCTGAAATACATGACCAACACATATAA GAAAGAGCATGGTGTGAAGTCCAATGGACATATGAGTGTCTCCCTGTG TAATAAGCAAGCTCTTTCATGGGGGAGCTGCTGCAAGTGGGTACA GCGCATGGAACAAAGCATACCGAAACCGTGAAGTTGAAGCTCAAA CCAAGCATCTCGACTACTGCCAAGTTGCTGCGAAGCTAGATGAGATCG CCGTGTCGTACCCAAACCGAGACAAGTACGTGTTATTGAGAGACCCAGA TGAAAGTGAAGAGAGGTTGCAAAACCAACACTTGCTTCCACGAGAGA	AAGGACG CTGGTTTA AGGCCGA ATTCGTT GTTCTTTT CTGGCGGT CTTGCTTT GAGCTTGG TTTCCGAT CCT (SEQ ID NO: 1169)	TAAACC CACACG AGAMCT ACGACG CCATAA GATCAG GCATGT ACGGAT GTGAAT GAGACT GATGAA CGGAAT GAGCAC GTGCCC ATAAGA TCGGGT ATKAAA GAWCA GAGACG ATCCCTA MCATCG GGAAAA CACGAG TTATACT GCTTCAC TGAMCT CGCTAA GCTCTCA TAATGA	MPLXISDCVHLVSAE GDTMNGRSTCGPLS RSSSVVSRSSSPSPS VPPHPSPSIGPDTGLS AGIIGTSRGCSLWLE VDNALSQWLKLE RDHEVLVCGFEAAKP LSLSKARLLRKTPRNT GVVRHILEFDGRLVH TNCNETECLVSLXSH XAVEVVRISLKCEPRE PCEPKCVLSILCSKIV XISFECETREPFPPFXD RKREPIPVFERMY DPRDPIPSFICWMYD LRQRTPGTLPXNPL SXENKDSWGRPAVVK NEIRSMRSYLEENVK ENRNLRLRRGGGE GKKMIRKLVAEKKSD TEAVCRILYPLDDRYE CFVDGCTTSTMGY GSSDLKYMTHHIKE HGVKVQWYECSLC NKQAPFMGGAASK WVTAHMAKHTETV KLKLPISISTAKVAA KLDEIAVSLPKPRQVR VLRDPDEVKEKVAKP	HRNFNNG (SEQ ID NO: 1413)

220

CAAGACACCGCTCCGGGTGAGCGGACTACGGTATCACCATTGAAATGTTTTGATCCAGAGTAAAGATGTTGACACTTCTCTACAATGAATGTAAGAACATCGAAAGATCCAAAGTCATTGAAAGAGGCGAGAGACGATTCTCTTTATAAGGAGGTGATGAAACGAGGCCGACAACTGGAGACCTATAAGTTGATGCCACGATCTACAACTGATTCTAGTCTTTGGAACCGAAAGATTAGATCAGTTGGTGTGATGAGCAATGTCAACGAGGTTTCCAAGAGAGAAGGTGTAATGAAAGCATAGGAATCCTTAGAACGGCTATCGATGCTGCTAAGGAAAGAGAGAACTGTGTCAGTTGCATGGTTAGAAGTTACGAATGCGTTTGGTTCAGTACCCCATGAAGTGAATAAAAGTACTCTGGAATCGTATGGATTCAGCAAAATGGTGACAGAGATTGTCTATGGATATGTACAGAGGTGCATCAATCCGAATCAAGAGCAAGAAATGAGAAAAGTGAACAGATTGTTATCAAACTCTGAGTGAAAGCAGGAGACCCCATCTCCCCACGCTGTTCAACATGTGTTAGAGAAATGTGATAAGAAAGACATCTGGATAAGTATGCGGAGAGGTTCTGAGAGGGAAGCCCAATCAGAAACCTAGATGAGAATGGTTCTATAAGTACCTGGGAGTACATACAGGAATCGCAACAAGAGTTTCAACAATGCAATGTTGAAAGTGTACGAAAGAAATGGATCTAGTGAATCAGAGCGGATGGCTCGTTTCAAAAAGTACTGCTGCTAAAGACGTTGTTTCCGAACTAACGTACATGTATGCGAACGCAATACCTAAGTTAACGGAGCTTAAGGCTTTGCAAACTTGACGATGAGAATGGTGAAAGATTGATGAAATCCCATCAAGGATCTCCGTAGAGTATGTACAGCTACCTCAAGTCAAGCGGATTAGAGTGGCTTGTCCTCAAGATAACAGCGTTGATTACCTTCTGGTCAACGTATGAAAAGCTATGGTCTCTGACAGTTACATCAGAAAACATACAGGACTACCTGGATGAAGTCGCAGAGACGAGACAGGTATGGAAGAGATGACGAAAGAGATATTGCAAAATATCTGAGTGGTATGCCGATCGCAAGAGGCGTTCGGTTACAAACAMTTCAAAAGAGTAAGAGATGTGTGCAACAGCTCACTAASATAGKGGGAGCTCCACTGCACAAGTTAAAGATTGTGGAGAGACGGTGACTTTGCCATCTAGTGCAAGCCCAAGAAAGAAATGGAGAAAATCTTCCACTGTGCTCAGGAGAAGAACTCCAACAGCTTCTGAAAGCAGAAAGTMAACACGGCTAGCGCACCGTTTCTTCAACGAGAAACCMGTGAAAAGTGCAAGTATGAGTGAATGAGACAGATATCCACAGAGCAATGCCTTTGTGAAGAAATGGAAGAATGTGAGCATTTGCTCCACTCGTGGATACACAAAGCAAGGTTGATGCGCTGCAATGCAACTTCAACACGTACGGTGAAAACCAAGTCAAAAGTGTCCGACGTTGCGGCAAGACGTTGGAACCCAACTGCACATCTCTGCA					GGGAAA GGAGAA AGTTGTT TGTTGG GCTGGC AAGAGT GAAGTT TGAATG TGAACC ACCGTC ATGCAA CCACTA AACCAG TGCGCA TGCGGG TGGAGT CATCAC AGGAAA ATGTTTC TGTTGCT TGACTTA TCAGTG TTTGATA TCGCCCT CAGGCA CAAGTA TGAAGG CCCCAC CCACAT AAACTC CCTAGC AACTGG TAGTCC AGCAAG CGCTGG TWCTTG CTACTAT TGCGCC	FEGETAETNPAEVL KEMGSRLPKLEALD WMEANFIESEVSDA MKTKDTAPGVDDL RYHHLKWFDPVKM LTLYNECKNHRKIPS HWKEAETILLYKGGD ETRPDNRPISLMPT IYKLYSSLWNRIRSV GGVMSKQCRGFQER EGCNESIGILRTAIDV AKGKRRLNSVAWLD LTNAFGSPHELKST LESYGFPEMVTIVM DMYRGASIRIKSNE KSEQIVIKSGVKQGD PISPTLFNMCLNVIR RHLDASGHRICKTK VKVLAFAADDMAILAE NRDQLQTELKLDKE CESLNLIFKPVKASLI IERGMVNKNAEVVL RGKPIRNLDENGSKY YLGVHTGIATRVSTM QLLESVTKEMDLVN QSGMAPFQKLDCLK TFVLPKLTMYANAIP KLTELKVFANLTMRM VKEIHEIPKGSPLYV QLPPSQGGGLGVACPK ITALITFLVNVMMKKL WSSDSYIRKLYRDYLD EVAETETGMEEMTK EDIAKYLSDGVPIDKK AFGYNTFRVRDVCN SLTXIXGAPLHKLIVE RDGDFAILVQATKEG MEKIFTCAQEKLLQQ

NeS L	NeSL- 1_Clap	.	Caenor habditis japonic a	GWCATCGGAGTACGGGTTACCAAGCTAATCAACGAAAGACATGATGC GGTGTACATGTGGTGAAGAACTCATCCGCAAGGCTCAAGAAAGA CTGGAAGCTAAGATAGATGAACTGTGCAAGTTGTAATCAACTTCGT CCAGACATCTATGTGTAGCCAGATGGGAAAGAGGTCAATAGGCA GATGTAACCTGTCTTATGAATCAGGAATGCAAGCTATGCAAGAGATT GGAACGAAAGGTCASAAATACGAAGAGGCTTAGCCACTTCCAWA AGATGGGAAAGAAATCACAGTGTGCCAATAGTGGTTGGATCACTGG GAACGTGTGGAAACCCACACGAACAGCTTAGTTCAACTAGGCATAG AGAAAGSACGATAAGAAAGATGATCCCGAGCTGTGCTCAATGACCA TGGAATACAGTAAGGATGTCTACTGGAACCATATATTCGGGGACACCTT CAGGAAGCCACCAATGAGATTTGGTGTAGAGAAGCCAAAGGGTAATAG TTGGAAGAAAGGAGCGAGCGAGCCGAAAGGTGCTGCTTCTCCGACTA AACCACACGAGAMCTACGACGCCATAAGATCAGGCATGTACGGATGT GAATGAGACTGATGAACGGAATGAGCACGTGCCCATAGATCGGGTAT KAAAGAWCAGAGACGATCCCTAMICATCGGGAAACACGAGTTATACT GCTTCACTGAMCTCGTAAGCTCTATAATGACCGAACTGTTTCGCAAC TGCCTCTAACCGGGGGTGTGAGAAAGGAGGTCGCTTGAGGCGGA CGCAATGAGGGATGTGCGAGTTCCCTCTTGAGATCCGAAAGTCTA AAAGTACTAGACCGAAAGATCGAGGACGGAGCGGGATGGCCGAGGC ACACGGCGGTAACACAGCCAGATAACCTAGTAGATCTTCGGATCTCGT CGGCTGGAGATATGTGAAACCTGGGAAAGGAGAAAGTTGTTGTG GGCTGGCAAGAGTGAAGTTGAATGTGAACCCAGTGTGCAACCACTA AACAGTGGCGATCGGGTGGAGTCATCAGGAAATGTTTCTGTG CTTGACTTATCAGTGTGATATCGCCCTCAGGCACAAAGTATGAAGGCC CCACCCACATAAATCCCTAGCAACTGGTAGTCCAGCAAGCGTGGTWC TTGCTACTATTGCGCCCCAGGCTCGCCC (SEQ ID NO: 1046)	CAGGC TCGCCC (SEQ ID NO: 1292)	LLKAEVNTALAHFFT EKPVSAMVMSVMRQ YPQSNAFVKNKGV SIAVHSWIHKARLNA LHCNFTYGENKSKV CRRGKDVTQLHIL QXCEYGLPKLINERH DAVLHVVRNLIRKGS KKDWWKLKIDETVSSC NQLRPDIYMCSPDGK EVIMADVTCPYESG MQAMQESWNRKVT KYEGGFHFVKMGK KFTVLPVVGSLGTW WKPTTNSLVQLGIEK XTIRRVPELCSMTME YSKDVVWNHIFGDTF RKPPMRFGVEKPKG NSWKKEGSEPKGAA SSD (SEQ ID NO: 1414)
				GCGCCCCG GGTTACAT TGTCGGG GCACCTT TCTCTGG AGTAGAGT ACAGTCTA CTAATTTT TGATAAGC TAGTCGGG TCCGAACC ACTAGAGT	TAAAG CCAAAA GCCACG GAGCAT CGGGAA AGAAAA ATGGAA AAGGAC TGAAAA CGAGAC TGAAAA ATCCCA	WRRPAPKQTKNSL HHLGHEVKRIARLP GIFFEFAKPKNSLHH LGHGVKRXARLPKI FEFHAKXKNSLHHL GHEVKRIARLPKIFE FHAKPKNSLHHLGH EVRRNSRLKPGIFGY QKSKNSLHHLGHEV RRIARLPKGFLEHAK NRIKSLKVTFLSDLX AHAGALACSRFLAST

GGGATTTGGCAAGAAATGTTGGAAATAACGAAATTTCTGTTATTTTCAGC ACAATTGTCAACCGGCAAGAAAGCTGGATGGACAAGACACACAATTTA CCGGAAATTTGCTTGTACGTGCAATTTCCCAATTTGAAAAATTCCT CGTTCACCTGGTGGGACGAGTGCAGACGATCTGCAGTCTGAAACC CAAAATCTCGGATTTATGACGATGCGCGCGCGCTCCAAACA AACAAAAATCTCTCGTGCACCATTTGGGCACGAGTCAAACGAATT GCAGTTGAAACCTGGAAATTTTGAATTCATGCAAAACCAAAAAATTC CTCGTTGACCAATTTGGGCACGGGTCAAACGAWTTGCAGTTTGAA ACCGGAATCTTCGAATTCATGCAAAACMAAAAAATTCCTCGTTGCAC CATTGGGACGAGTCAAACGAATTCACGTTTGAACCCGGAATTT TCGAATTCATGCAAAACCAAAAAATTCCTCGTTGCACATCTGGGCA CGAGTGCAGACGAAATTCACGTCTGAAACCCGGAATCTTCGGATTCAT CAAAAAATCAAAAAATTCCTCGTTGCACCATTTGGGCACGAGTGCAGAC GAATTGCACGTTTGAAACCCGGAAATTCGAAATTCATGCAAAAAACCG GATAAAATCCGGTCTAAAAAGTACGTTTTTGTGACATCTTWTGCTCAC GCTGGTGGTGGCGTAGCAGATTTCTGGCATCGACACTTAAACCG AGCACTGCCGACAAAAAGCTTCAAGCCAGTCGTTTTCTGTTCTATTC CTAAAAAATCTCTCGATCAACGAGGTGGCGTCTCTCCGCAACGTCAAAA AAWKTTTCTTGAATTTCTCAGGAAAAACCTATCGGTGGGATGGCTTC TTTCAGTAGAACCAAAATAACTTTTCAAACCTTTGTTGAAAGAAATTTGT TTTGTCTGTGAAAAATCCCGGATAATCCGCCAAAAAACAACCAAAAC AAAGCGAGCKTGTCCAAATCGAAGAGGAGTCACTGTGACAGACTGC TKCCGTCCCAAAAGATGGCAGGAGTCTCGGACGACTATTTCTCGGT CAGAGCACTCTCGCACCCCTTTGACACGTTATTGGGAGCGATG ACGACAAAAGAGGAGGAAAAACGATCAAACTCCAGTTTTTATTAAAGA AAATCTGGTCACACCTGGTTGCTAGGACGTGAAAAATTTAAAMAAA CAACAAAAAACAATTTCTGGAACTCTGATAGCAACAGTGAACAAAA AACTTCTAAAAATAAAGTTTCAGACAAAAATGGCCCATTTGATTGGG GCGGTAAACCAAAAAATCGGAGAAAAACCAATCACACGCACTTTGG AAATTGAATCCAAAGCGATGACAAACAAAGTTTGTCTCCGAATACT GTACCCAACTAATGATTGGTCAAGTGTACTCCCAATGGTGCCAAACAC AAATCCCTTGTGGATATGGCGTCAAGATTTAAAAATACTTGACAGACCA CATAAAGTCCACTCATTTAAAAAGGTTGAGTGTCTTATCAGTGTAGTA TTTGTGACGCAAAAGCCGAAGGTACCGGTACAAAAGCWGCTAGATGG ATTACAGCCCAATGCAAAAAATGACCGGTTTGAAGCAACACACAGAA TTAAACAAAAATCTGAAAAAACAACAAATGTTAAAACTCGCAACAGTCT CCAGGAAATGGCGTGTCTGCTCAAAAAACCAAAAAATGGTCCGAAAAA AGTTGTAATGGCAACTAGTACGACCCCAAGAAAAAGAAATCTCTGAACTG	TTGCTTGA AAATGCGT CAAAACGAG CATTTAG AACTCGCC CAAAAGTT CGGCCCG ACCCCAA ACAAATGG GACTTCT TGACGATT TTCCTGA AAATCGGA GGATGGA ATGGTCCC CTATTCT GTAAATAG KACTGTGC AATACCCC TTCGTAT CTGTGG GAACAGAT GACACGTG ACGTATC CGTGTAGA CGTCAGT TTTCCGT GCTGCGG GAGCCCC AATCGAGC AATTTTG CTCTTTG AGTGTCTG GAACGCT GAAACCCC AGACAAAT CAGGCCA GTCGTCG	AACAAA ACAAAT CCAAA CAAACT GAAAA AAAAA AAAAA AAAAA AAAAA AAACTG GACAGA CACTGG AAACAG TGTCAG GCAAAG TCGCG ATTATAC TGTTCCA CGCTTA AAAGTC CCGAAA TGGCG AAAAA ACCTGA ATCTATC TGAAAG TGCTCCA AACCAC GCACAA CTCGGA GAAAT CAGGGA CAAGTT CAAGTT GCTTAC GCAACG GGCTGG GACAGG TACCCC TCCTGA	LKTEHCRQKSFKPVG FLHFLKNSSINEVAS LRNVKKXFEFFSGKP IGGMASFRTKITFFK LCLKNFVLSAENPPiR QKTNQNKASXVQIA RGHLSDCPLPSQKM AGVLGRFLSVQSTLS HRPFDLLRSDDDKR GRKTIKLQFFIKENLV TPXVARDVKILXKQT KNSGNSDSNSETK NFSKNKVSQRNGPLI GGGNHKKIGENQITR TLEIESKDDDNKVLVL RILYPTNDWYKCSQ WCQHKSLVGYGAHD LKYLTDHIKSTHKKV EWSYQCSICDAKAEG TGTKAARWITAHMP KVHGIEATHRIKQNS EKTNNVKTANSLQE MALSLQPKNGPKK VVMATSTTPEKKISEL ESKIQTREVAKQLSAL KESAQKNQQGNKTK NVKSSLKTIAENTNET KKISARKSLINYLKPED VLNHIKPEKPASAKX GLQELTGAORLQETR RRFMAGNRRDSiAR RESLSGKISNSFKIEL KNAPEKTTLKKPAVT QKQNTSQNVSSSTV VKENKTGNDVITIDD TETVKKRINTWCLDH ESTENAWMADDIIF WYIQKQIEISLDNKKF

GAATCAAAATCCAAACCCAGAGAAAGTGCCCAAAACAAATTGAGCGCTCTG	AAAAATTC	AACCGC	KVIDPLIWTTTYRIYGV
AAGGAGTCAGCTCAAAAAAATCAGCAAGGAAACAAAAACAAAAATGTT	TTTTGAA	GAGGT	ECVQDELVGFEKYFF
AAATCAAGCTTAAAAACAATTGCTGAAACACAAATGAAACMAAAAAAG	ATTTTTG	GAGGAT	PICENGHWVLLIIDDK
ATWAGCGCTCGAAAGAGCCTTATAACTATCTGAAACCTGAGGATGTG	GCGCTGC	GACGG	RVWYSDSLADKPIEVI
CTCAATCACATTCGAAAGGAGCCAAACCCAGCTTCGCGAAATGCGC	GAAAAA	GAAGC	EDLINKLNRTQGFEN
TKCAAGAACTGACTGGTCTCAAAGACTGCAGGAAACMAGAAGAGG	ATTTTTAA	CGCGAG	QTVPKQKDGFCNV
TTTATGGCWGAAACAGAGAGATTCAATTGCAAGAGAGAAAAGTCTG	CGCCACA	GCTTAT	HVCLVAKSVITENFW
TCTCTGGCAAAATCTCAAATCTATTTAAATTTGAGCTGAAAAATGCTCC	AACCCCG	GGCGG	YTEKDNDFRKTVKL
GGAAAAACAACCTTAAAAACCGGCTGTCACTCAGAAAAACAAACACG	GGAGCG	TAACTC	WLFSEGFELYSEPYK
AGTCAGAAATGATCTAGTTCTACGGTTGTAAAGAGAAACAAAAACAGAA	CGGWTAG	GGTTGG	QIQNKINISVNSEKNQ
ATGACGTGATCACAATTGATGACACAGAAACTGTAAAAAGAAAAATAAA	GGATATCG	TGTGCT	ISDNEKNWGDKTQT
CACCTGGTGTCTCGACCACGAATCCACAGAAAAATGCGTGGATGGCTGAC	ATGTCATC	AGTAGA	VNESTLKERDEDEIFLL
GACATCATATTTCTGGTACATCCAGAAACAGATTGAAATCAGTTTGGACA	GACTCGTC	TGATTTA	RPHISVGVALKTEDEK
ATAAAAAAGTTCAAAAGTGATTGATCCACTCATCTGGACCACATATCGAATT	GGTGATCT	TATCGG	NQKAENLKAPQKLIK
TATGGTGTGGAATGTCCAAGATGAATGAGTTGGTTTGGATTGAAAAATACT	TTGATTT	ACAGCC	AIRRLKILKTKCLKLTA
TTTTTCCAATCTGTGAAAAATGGTCAATGGGTTTGGCTGATTATCGATGAC	CTCTCTGC	CCAACT	VKGKPEETERAAIPNL
AAAAGAGTCTGTACAGTATTCCCTGGCCGATAAACCAATTGAGGTTA	GTCTCCTA	AAGAGG	MAIKLTPPKVPEVR
TTGAGGACCTCATAACAAAACTAAATCGAACCAAGGTAAATTTAACCA	TTTTGGAA	AATCCT	RNPEKGENYKKSQP
AACGGTTCCAAAAACAAAAAGACGGCTTTAAATGTGAGTTTATGATGT	CAGTCTCG	GGGAAA	NKKRQIPTGKDELV
CTGGTGCCAAATCCGTTATCACTGAGAACTTTTGGTACACAGAAAAAG	ACCAAAAA	GGAAAA	KKVREWFELQFQAYF
ACGTTAATGACTTCAGAAAAACTGTCAAGCTTTGGCTTTTCAAGTGAAG	ACCGGGCC	CTTGAA	EDGKSFQRLWXTGL
GTTTGAACCTCTATTCAGAGCCGTACAAAACAAAATCCAAAAACAAAAATTT	TGGCAACC	AAAGTT	LTAAIHKASAGDEQA
CCGTTAATTCGGAAAAAATCAAATCAGTGATAATGAAAAAATTTGGG	CACCGAAT	TTTACAG	VGKIIRCPPLEIEGE
TGATAAAACTCAAACCTGTAATGAGAGTACTCTGAAAGAAAGAGATGA	CCGGATGT	GGCTGG	MATQTETKQPKNQ
AGACATCTTTTGTCTCAGACCACACATCAGTGTGGAGTTGCTCTCAAGA	CGGAGG	TAATAG	KSTKGANSSSIREAY
CAGAAGCAGAAAAAATCAAAAGCTGAAACTTGAAGCCCAACAAA	ATTTGGCA	TTCAGC	AENRARTFNKIGKD
AACTCTGAACACCGGAAGAAATCCAAAGTGGAACAAACGAGAAACCCAG	AGAAATGT	ACAATT	DKCEIPIEKIEKFFENT
AATCTCAAATGCCAGGAACTCCAAAAACGAGCCAAAAATGGTTCC	TGGAATA	GATGTC	TSNTNVPETETLARITS
AAGTCCGAAGAAATCGGAAAAAGAAATTTCTACTGAGCTTCTGGATGCT	ACGAAATT	TACTGTC	DLPKLEIGSWIEEFER
CAAGAAGCCGGAAGAGCTGAAAGCGATCCGAAGGCTGAAAAATCCT	TCGTTATT	TTGCAA	EKEVAEALKTKKDTA
GAAAACTGTCTCAAAAAGCTGACGGCAGTGAAGGGAACCGGAAGA	TTCAGCAC	CCACAA	PGVDGLRVHHLWSWF
GACGGAAGAGCCGCATTCCAAACCTCATGGCAATCAAGCTCAAGAC	AATTGTCA	CAAAAC	DPKXLLTKLYNECRE
GCCTCCAAAAGTTGAACCTGTAAGAAGAAACCCCTGAAAAAGGTGAAAA	AACCGGCA	AGTGGT	HKKIPGHWKEAETVL
TTACMAAAAAAGTCAGCCAAACAAAAAGAGACAAATACCAACCCGAAA	AGAAAACT	TCTCGG	LYKGGDETQAENWR
ACCGGATGAATTGGTTAAAAAGTCCGAGAAATGGTTTGAATTCATTT	GGATGGA	GGTAGA	PISLMPTICKLYSSLW
CAAGCATATTTTGAAGGACGGAATAATCCTCCAGAGGTTAGAGTGGWTG	CAAGACAC	TCAAAC	NKRIKSVTGVLSKQ
ACAGGTTTGTCTACGGCTGCAATTCAAAAAGCTTCGGCTGGAGATGAGC	ACAAATTA	ATAATT	RGFQEREGCNEJAIL
AAGCTGTGGMAAAATCATCAAACGTTGTCCACCTCTGGAAATTGAAG	CCGGAAAT	GTGTGT	RTAIEAAKGTKKLSLI

AAGGGAAATGGTACCCAACTGAACAAACAAACAAACAAACAAACAAAC	TGTGCTTG	TTTCTT	AWLDTNAGFVPH
AAAAGAGCACAAAAGGAGCAATAGTTCAGCTCAATTCGGGAGCCT	TTACGTCG	TACTTGA	ESIEATLIAYGPPGMV
ATGCTGAAAACCGAGCGAGAACCTTCAACAAAATTATTGGAAAAGACG	AATTTCC	CCCGG	TEVIKDMYNGASIRV
ACAAATAGTGAAATTCCAATTGAAAAATTGAAAAGTTCTTCGAGAA	AATTTGA	CAACAC	KTKNEKSKQILKSGV
CACAACTCAAAATACCAATGTTCCAAACAGAAACACTAGCGAGGATCACT	AAAAATTC	ATTATAC	KQGDPIPTLFNICLE
TCTGATCTTCCAAACTCGAGATTGGTAGTTGGATTGAAGAAGAGTTCA	CTCGTTCC	CACGTC	SVIXRHLKSADGHKCI
GGGAGAAAGAGTAGCGGAGCTCTTAAAAAACAAAGGATACTGCC	ACTGGTCG	CACAAG	XSNIKLLAFADDMAIL
CAGGTGATAGTGGATTACGGTACCATCATCTGAGCTGTTTGATCCAAA	GGACGCG	GACGAA	SDSKTLQQLQELQKM
AAKGAAAGTCTCAGAAAGTGTACCAACGAATGCAGGAGCACAAGAA	AGGTCAGA	TTCATAA	DDDCPLNLIFPAKC
AATCCAGGTCACTGGAAGAGGCGAGAAATGTACTCTCTACAAAGG	CGATCTGC	TGCCCC	ASLIEWGKVQKDQK
GGGGACGAGACGCGCGGAGAAATGGCGACCAATCAGTCTCATGCC	ACGTCTGA	CTCCCTA	IKLKGQFIRSLAEQDT
AACCATCTGCAAGCTATACTAGCTGTGTGGAACAAAGAAATAAATCC	AAACCCAAA	AATAAA	YKYLGVQQTGIEIVSA
GTGACAGGTGTTCTGAGCAAAATGCCAAAGGGGTTTTCAGAAAGAGAG	ATCTTCGG	CTCCCTA	MQLMKKTVSELDKI
GGTTGTAATGAAAGCAATTGCAATTCTCAGAACCGCTATTGAAGCGGCAA	ATTTTATG	GCAACT	NCSALAXWQKLDV
AAGGAACAAAAAGAGCCTGTCAATTGCTTGGTTGGACCTTACCAATGC	CAGTAG	GGTGGT	KTFVLPMKTYMYAN
ATTTGGCTCAGTTCCACACGAATCGATCGAGGCCACACTAATTGCTTACG	(SEQ ID NO: 1170)	CCGGCG	TVPKLSELKEFANITM
GTTTTCCGGGAATGGTAACCGAGGTAAATAAAGACATGTATAATGGCG		AAGCCG	RAIKVMQNIPIVKGP
CATCGATTCTGTAAAAACAAAAACGAAAGAGTAAACAAATCCTGAT		GTTCTTG	LEVQQLPIGKGGLV
TAAATCGGGGTAAACACAGGGTGATCCAAATCTCACTACTCTTTCAACA		CCACTAT	ACPKTTALITYLVSTM
TTTGCCTTGAAGTGTCATTMTGCGCCACCTAAAAGCGGGATGGTCA		TGCGCC	KKLWSTDDYIRKLHT
CAATGTCATCGAMTCAACATCAAAATTATTGGCGTTTCCGATGACATG		CCAGGC	DYLMVAIKETKKE
GCAATCTGTGAGATTCAAAAACAAACTCCAACAAGAGTTACAAAAA		TCGCCC	VTLEDASLYLSDDKTV
TGGATGATGACTGTACACCGCTCAACCTTATCTTCAACCCGCCAATGT		(SEQ ID NO: 1293)	CKKAVGYNSTFRVREI
GCAAGTCTGATAATTGAGTGGGAAAAAGTACAAAAAGATCAAAAAATA			CKTLSKNKGALLSQLK
AACTAAAAGGTCAATTATCATCAGAAAGTTTGGCCGAAACAAGACACCTACA			IIAKDGKLAILVQAXK
AATATCTGGGGTGCAACTGGCATCGAAACGCGCGTTTCTGCAATGCA			DGKTKIFTHDHVKTL
ACTGATGAAAAAAACKGTGAGGAGCTTGACAAAATAAATTGCTCTGCA			QKXLLKEINEALLHRF
CTGGCTCMWTGGCAAAACTGGACGCGAGTAAAAACTTTTGTGCTCCA			TTEKRVKSEVVRVVQ
AAAATGACGTACATGTATGCAAAATACTGTACCGAAACTCTCCGAGCTCA			EYPQCNSTFRVDDGGK
AAGAGTTGCAAAATATTACAATGAGAGCAATAAAGTAATGCAAAACAT			VSIGAHRFVHKARLN
TCCAGTAAAAGGTTACCATGGAGTATGTACAGTTACCCATTGGAAAA			LLACNVTWQDAAT
GGTGGACTAGGAGTGGCATGTCCAAAAACAACTGCGTTGATAACCTATC			KQCRRCGYEKETQW
TGTTTCAACAATGAAAAAATTGTGGTCCACTGATGACTATATCAGGAA			HILSSCPKSMGGKITE
ACTACACACAGACTACCTGAAAATGGTGGCCATAAAAAGAAACGAAAAAC			RHDSVLKTVKEMIQT
AAAAGAGGTCACTAGAGGACCTTGCCTCCTACCTAAGTGATGATAAA			GSLKNWKLKLDHELP
ACCGTCTGAAAAAAGCGGTTGGTTATAATTCAATTCACAAGGGTACGAG			GSTRLRPDYLRSPNG
AAATCTGAAAAACGCTATCAAAAAACAAAGGAGCAGCTGTTAAGCCAAT			SEIILGDVTIPYEHGIE
AAAAATCATTGCAAAAGATGGAAAAGTTGGCTATTCTGTTACAGGCTSTG			AMQTAWQKKIEKYE

NeS L	NeSL- 1_CRe m	.	Caenor habditis remane i	AAAGATGGCAAAACAAAGATTTTCACGCATGACCACGTGAAACCTTGC AAAAAATCTTAAAAAGAAATAAATGAAGCCCTTCTGCACAGATTCACA ACTGAAAAAGAGTGAAGCGAAGTGGTGGAGTGGTCCAAAGAGTA CCCCAGTGCAACTCTTTGTGACAGATGGAGGAAAAGTAGCATTTGGA GCGCATCGCTTTGTGCAAAAGCCAGTTGAACCTGCTCGCTGTAAAT ACAACAGTGGCAGGATGCAGCCACAAACAAATGCAGAAGGTGTGGAT ATGAAAAAGAAACCAATGGCAGATCTCTCATCTTGCCTCAAAAGATAT GGGAGGAAAAATAACTGAAAGACAGATTTCTGTGTTAAAAACAGTAAA AGAGATGATTCAAACTGGATCTCTCAAAACTGGAACTAAACCTTGAT CATGAATTGCCAGGATCAACAGACTTCGCCCGGATATCTATTTGAGAA GCCAAATGGATCCGAAATAATCTTGGCGATGTGCAATCCCGTATGA ACAGGAATTGAAGCTATGCAAAACAGCATGGCAGAAAAAAATTTGAAAA ATATGAAGAGGGCTTCAAAATACCTTCGTTCTACCGGCAAAAAACTCACA ATTGTCCAAATTGTGTCGGAGCACTAGGAAGTTGGTGGAAGCCACACA ACAGACAGTCTGTGAGTCTGGGAATCGACAAAAATACTGTAAAAAGAG CTATTCAGAAATTTGCTCTACAGTACTCGAATACAGTAAAAACATTTAC TGGAAACATATATTCGGGGATTCTACCAAAAGTACCCATGTTTTTCGG CGGTGAAAAACCAAGGGGCAAGTTGGAAGAAAGTGAAAGCCTCTGA AGGCAAACTGCTTAAACATGAGCCTCAGGTTAAAGCCCAAAAGCC ACGGAGCATCGGAAAGAAAAATGAAAAGGACTGAAAACGAGACTG AAAAAATCCCAACAAACAAATCCAAACAAACTGAAAAAAACAAAAA AACAAAAACAAACTGGACAGACACTGGAACAGTGTGAGGCAAAAGTC GCCGATTACTGTTCCAGCCTTAAAGTCCCGAAATGGCGCAAAACA ACCTGAATCTATCTGAAAGTGTCCAAACCAACGACAACTCGGAGAAAA TCAGGACAAAGTGTCTTACGCAACGGGCTGGGACAGGTACCCCTCTCT GAAACCGGAGGTGAGGATGGACGGGAAGGCCGCGAGGCTTATGGC GGGTAACCTGGTGGTGTGCTAGTAGATGATTTATATCCGACAGCCCA ACTAAGAGGAATCCTGGGAAAGGAAACCTTGAAGAAATTTTACAGGG CTGGTAATAGTTCAGCAATGTAGTCTACTGTCTTGCACCAACACAA ACCAGTGGTCTCGGGGTAGATCAAACTATAATTTGTGTCTTTCTTTA CTTGACCCGGGCAACACATTATACCACTCCACAGGACGAAATTCATAA TGGCCCTCCCTAAATAAACTCCCTAGCAACTGGTGGTCCGGCGAAGCC GGTCTTGCCACTATTGCGCCCCAGGCTCGCCC (SEQ ID NO: 1047)	CGGGAACC AGTCAT (SEQ ID NO: 1171)	TAGCCG ATCGTA AAAGAA ACCGAG CCGTAA	MTVFIDRIGERGQ MAVCSLHRYFSFPF SPIPPYNNNGSFGEN GCGTDKSLLPVIEVV REVKNWSENILVVE

ATGTTAAAGAGCGGAGAAAGAGTCGTTGTAAGAGACAAAAATCTGGAA AAAGTTATTAGAAATTTGGCAAGAAATCAACTCAACTCTATTTTCCAATCT AGGAAATCAGATATTTTCGCTAGTACCCAGATAAAAGACAGTACCAAT AAAGAGCAGGGATACAGGAAAGAGAAAGCAAWTGAATTCATGTATC ATTCCGAAGTATAAAATCCCAAGTCCACCATATTTGAGAGGTGGGGA GATGTAATGGAAGATACAGAGATAAGAGGTATCAGAAAGTTGGAGCCA GAGGTCAGTTAGACAGCTCAAAACCGCTGATCTGCAGAGTTCTTACC CAACGCAAGTTATATGTATAAATGTTTTATCCAAAGTGAAGGACA TAGTAATGGATCAACAGATCTGAGAAGTCTGAAGAAACACATGGTGGA TAAGCATTTACGAATATTGAATTTGCATATAAATGTGCTACGTGTATGT TTTTACGACTGGGAAATCGGCCACAGCGTTAAATCAATAAAGGCACA TATGGCAAGTCAACCACAAGGTAACGATGGAACCCGGTAAAAAGAGTCT CGTGCAAAAGTTGAATGCCAGACTCGAAGAGCTGCTCCATCAGTTCCA ATGCCGAGAAATCGATCAAAAGTCTACAGTTGACCCCGAGAAATCGA TATCGGAAATGGAGAAAAAGAAAGCAAACTCGTTCTGTGGCAAAACAGC TTAGCACACTGAAAGAGTCGGCACAGAAAAAGGAAGAGGAGGTGAAG ATAGCGGAGGTCAAAAAGAGAGAACCCCGTCTATCAATAATCCAGAG TCGAATGTCAGGCGAAGTCTGGCGGCAAGGACTCGAACAATGTATAAAC CCTGAGCAATCGGTAGCTCAGAGGATAAGAGAAAAAGAGAAATAC GCCAAAGCTTAGGGAGGCGCGGCAAAAGAAAGATCGAGTTGGCA ATGAAGCCAGCTAGATTACAGACAAAGAAACGAGATTACACTCCAG GAAACGAAAAAGATCGATGATCCAATCGTTATAGACCTGGAAAAAGAA TGTTACTACTACAGTACTTCAAGTCCCAAGAAACAGTCAACTCGTG GTGCTAGAGCATGAGACAACGATTGACGCTGGTTAACGGATGAGGT AATACATATGTACATGTGCACAATAACGAGAAATCGAAAAATATTTATG GCAATCGATCCGGTCTGTGGCCAGTCTATGTGAGAAATGGAGCAGAG GATCTACTGAGGCGTACTAGTGGCCAGGAACATCTCTTCCAAATTTG TGAAAGTAATCATTTGGTTCTATTAGTAGAGAACACGATGTGTATGG TATCTGGATCCGAAAGGCGAGGAACCAAAAGGAAATGTAGAGATTCTT TTAGAGTCCATGAAAGGAAAGGAGGAGTACTATGAATTCACACACCT CACAGAGATAAATGTGAATTTGGAGTGCATGTCTGTCTTATGGCAAA ATCAATAGTAGATGAATGTGTTATAAATTTGTTATCTGAAGAGGACGTA AGGTCAATCAGAACCAATATGAAGGACATCTGAAAAAGTAAGGGATAT GAGTTATGCTCTGAGCCTTATAATAGGCAAAATTTATTAAGAAACAGAAA AACAAAAGGAAGTTATCTGGAAGAAATGATCGATTTCATTCGTTGTAGA AGACGATATGACGTTACAGTGATCGGGATCTCTGATCATGGTGATGAT GAAGTTGAACATCTGAAGACCATGAGCAGGAACCTGAAAAATGAAATA AGTGAAATTGAGAAATGAGAGGGATCTGTAGACTCAGTCAATTCCAAAGT	CAACAA GCAAG TAAACA AAAGAA AAATCA ATAAAA AGGAAG GTTGAC CTCAGA CCCCGA GGAGGG AAGAGA GACACC SAGAAA AAGAGA GACGCA GAGAAA AGGAGA GACACC TCTCATA AGGAGA GGTAGG TCAATCC AAATGT AAACAG AAAAAA CCAGTG GGGAGG AAAGAA AGACTG ATTTTAC CCACTA AAATGA ATTTGG AAACAG AATTTG GAAGAG AAAAGA	CLIMVKSGERVVVKR QNLEKVIQNLARINST LFSNLGNQIFCVVPRI KDSNKEQGYRKEKQ XKFHVSRFSIKSQVPP YLRGGDVMEDEIR GIRKLEPEAQLDSSKP LICRVLYPTQGYMYK CFYPKCKGHSNGSTD LRSLLKHHMVDKHFTN IEFAYKCATCMFLTTG KSATALKSIKAHMAS HHKVTMEPGKKSIV QKLNARLEEEAAPSIP MPNRNSKVIQLTPEK SISELEKKKQTRSVAK QLSTLIKESAQKKEEEV KIAEVKKREPRLSIPE SNVRRSLAAGLEQCI NPEQSVAQRIREKRE EYAKASREAAAKRRS SLAMKPARLPDKENE ITLQETKKIDDPVIDL EKECILTVLVQVPRNQ FNSWCLEHETTIDA WLTDEVIHMYMCTI TENRYFMAIDPVL WPVYVRNGAEDLLR RTSCPGTFFFPICESN HWVLLVIEHDVYVY LDPKGEPKGNVEILL ESMKRKRQYVEFPFP SQRDNVNCGVHVCL MAKSIVDECGYNWY SEEDVRSFRFTNMKDI LKSQGYELCPPEYNR QNLLKTEKQKEVILEE MIDSFVVEDDMTFT

TGATGAAATGAGAGTGAGACACCTCCAGTGATCAATGAAAAAGAG GTAAAAAGCGAGTATCGCCAAAGAGAAACCGAGAAAAAGCAAAAGGAA AAAGAGCAAAAAGTGCCAAACAGGAAACCCAGATGAGCTGGTTAAAAGA GTAAGAGTATGGTTGAGAAAGAAATCAAAATCGTATGTGGAAGATGGA AAAAGTTCCAAAGGTTGGAATGGMTAACAGATGTTCTCACTGCAGCA ATTGAAAGGCGTCAGCCGAGATGAGAAAGCAGTAGAACTGATTGAG AAAAGATGTCACCTTTGGAAWKCGAGGAGGTGAAATGTGTACCCAG ACTGAAAGAAAAAGAAACCAAAAGTGTAAGGGAATGGCGGTCA AGAAAGTATGAAGTCTTTGATGGCCTCATACAGTGAGAACCGAGCCAA AACCTACAATAGAAATAATTGTAAGCATTCAAAGCAGTGTGAGATCCCA ATAGCCAAAGTACAAAAGTTCTTGAAGGACCACTGCCGAGACAAATG TGCCAAAGGAAACACTTAAGGAAATGTTCACGCCTCCCGAAAAGTTGA AGTGGGAACGTGGATTGAAGGTGAATTCAGTGAAAGTGAAGTGACTGA AGCATTGAAGAAAGACAAAGGACACAGCACAGGGGTAGATGGATTAA GGTACCATCACCTGAAATGGTTGATCCCGAGTTGAAAATGCTGCACA GATCTATAATGAGTGTAGAGAAACACAGAAAAATTCCAAAGCATTTGGAA AGAGGCAGAGACAATCTCTCTATAAGGGAGGAGATGAGTCMAAAM CGGATAATTGGAGGCTATCAGTCTGATGCCAACCATCTATAAACTGTA TTCTAGTCTCTGGAACAGAAAGGATTAGAGCAGTGAAGGGGTGATGAG CAAGTGTGAGAGAGGTTTCCAAGAAAGAGAAAGGATGTAATGAAAGTAT CGGAATATTGAGAACAGCCATAGATGTGCCAAGGGCAAAAGAGAGAAA CATAGCCGTAGCATGGTTAGATCTCACGAATGCCCTTGGATCAGTACCA CATGAGCTGATAAAAGAACTCTGGAATCTTACGGATTTCCAGAAATAG TAGTAGACGTCGTAGAAGACATGTATCGAGATGCATCGATCCGTGTGAC GACGGAACGGAGAAAAAGTATCAGATTATGATCAAGTCAGGAGTGAA GCAGGGAGATCCAATCTCGCTACTCTCTTCAACATGTGTCTCGAGAGT GTCATCAGAAAGGCATCTCAGACAGATCAGTCGGCCATCGGTGCCTGAAAA CAAAAATAAAGTATTAGCCTTTCAGACGATATGGCAGTATTAGCAGA AAGTAGTGAACAGTGTGCAAAAGGAGTTGACAGCTATGGATGCTGACTG CTCAGCACTGAATTTGCTATTCAAAACCGCTAAATGTGCAAGTCTGATAT TGAAAAAGGAAATAGTAAACAGGTTAAATGAGGTAGTTTGTGAGAGGGA AACCGATCAGAAACCTCATGGAAAAATGAGACCTACAAAGTACTTAGGTGT TCAGACAGGTACGGAAACCAAGGGTTTCATAATGGATCATATAACCGGA AGTGTCAAGGGAGATAGATCTAGTGAATATGAGTCAACTGGCAATGCA CCAGAAACTAGATATACTCAAAGCCTTCATACCTCCAAAGATGACCTATA TGATCAGAAACAGCACCTAACTGTGAGAACTGAAAGTGTGTTGCCAA TTTGGTAATGAGGTGAGTGAAGGAAATTCACAACTCCCTTAAAGGG TCACCGTTGGAGTATGTCCAACTTCCCGTAGGAAAAAGGAGGATTAGGA	GAAGG GAAAC TAAAGA AAATAG TTCTCT GCCAAA ATTCTGT AGAGGA ATACTTT GTCAAA ACATGA TAGAAA CCAGTA ATCTGG TACGAA AGACAA GTAAGA CCTGAA CTGACA AGAAGG AAGTCA GAAAGA AATACC GCTCAC AAAGCC TGTGAT CGATTCT CTTACT ACTGAA CTTGTC TCTTGG CCTCGTA ACCGGC TAAAGG GAGAAG GAATGT TAATTG GAGATA	VHRSDHGDDEVEH LKTIEQEPENEISEIEN VEGSVDSVIPKLMEM RVQTPPVINEKRGKK RVSAKEKPRQKEKE QKVPTGKPDDELKRV RVWFEKEFSYVEDG KSFQRLWXTDVLTA AIKASAGDEKAVELI EKRCPPLEXEEGEMC TQTEKKKKPKSGKN GGQESMKSLMASYS ENRAKTYNRIIGKHSK QCEIPIAKVQKFFEGT TAETNVPKETLKEMC SRLPKVEVGTWIEGE FSESEVTEALKTKDT APGVVDGLRYHHLKW FDPKLMLSQIYNEC REHRKIPKHWEAET ILLYKGGDESKXDNW RPISLMPTIYKLYSSL WNRRIRAVKGVMSK CQRGFQEREGCNEI GILRTAIDVAKGKKR NIAVAWDLTNAFGS VPHELIKETLESYGFP EIVVDVVEDMYRDAS IRVTRTEKSDQIMIK SGVKQGDPISTLNF MCLESVIRRHLDVRSV GHRCLTKIKVLAFAD DMAVLAESSEQLQKE LTAMDADCSALNLLF KPAKCAJILEKGIVN RLNEVVLIRGKPIRNL MENETYKYLGVQTG TETRVSIMDHTEVSR

229

NeS L	NeSL- 1_TV	.	Trichom onas vaginali s	GGTGAGTAGTCTAGTGGTATGATTCTCTGTTTGGGTACAGGAGGTCCC GAGAAAGCTTCCACTGCAATCGTACGTGTAAGTGTGGCAACACAGCTATAA CAGACAGTTTCACGCCAACCGCGAAACGATTTGAAGCCTGAGGAACA AAATTTAGATATCGTTTGAAGAAATATTGACGCGTTGAATCCAGAAAAT ACTCCGACTTAATCAGGAGCCTATCGAAGATGGAGTTGAGATTAGATT CCGAAAGAAATAGAGAAATTAAGGATTTCCGAAAAATTTAGCCAA TCCATCGCATCATGCCCCATCAGTTGTTAGTCGCATCCATGTTCTCACCT GAAGACCGTGACTTGAGTACAGAACCGTTCCACTGTAAACGCTGATGGCT	GGTGAG TAGTCTAG TGGT (SEQ ID NO: 1172)	TAAGAA GAGATA AGACGA GTGAGA AGAACA GAAGCA TAGTAG GATTGG	CTGCGG GTGAAT CACTTTC GAAAAG AAGTGA ATGGAC GTGCTG ATGTCT GACTTTA AAGAAG TCTGAA ATTAAA AAAACA GATATA AAGGCC CCTCACT ATAAAC TCCACA GCAACA GGTGGT CCGGCG AGGCCG GTTCTTG CCACCAT TGCACC CCAGGC TCGTC (SEQ ID NO: 1294)	MIPVLGTGGPEKPLP QSYVYCGNTAITDSF TPTAKTILKPEEQNLD IVLKNIAALNPENYSD LIRLSKMEFRLDYPK EENYWISEKLFQSIA SLPISLLVASMFPED RDLSTPEPFHCNADGC

231

TAGCCACAAGATACGCAGCTACAGGAAATACAGAAAGCAGACAAAATCA AGGTAGATTACTTGACACTAAATGCAATGCCCTTGAGAGAAATGAT CTACACCAATAACACAGCGAATCAAAGTTTATGACAGCAGAAAAATGGA GAAGACACAGACCACCGCCAGGATATCGGAAGACACAAAGAGATCGC ATTCAAAAAGCAGCCAAATGAAATAAAGGAACTCTCATCAAAGTAGTCA AACACATAAGTCACGGAGATGCTCAAAGACAGCAGCAGAGAGATG AACACATAAATTCGTGAAATGATTGCAAAAATCAAAAACGATCTCAG AGATAACAAATTCGAACAATATAACATTGAAGAAATATTTCAAGGACCG ATCTCGACACAGAGATTCTCAACATCGTCAACACCGGAGGACAAACAG AATTATCAAGAAAATGGATTACATTAAACCGAATTCTCGGAACACACCA GGATGCATCACCATATGCAAGGAAGAAATTACAAGCATGTTTCGCCGAT AACCCAAACAAAGACTCTCAGAAAACATAATCTTAGCCGACAAAAGTTCCAC AACAAATCATTGAAGCCAAAGCGAATACCTTGATTACTACGGACCTCAATG GGCAACGAAAGCTGAAGGCTACGAAAACCTTCTGTCATCATGACTACGCG TTACCGGAGAGATATGGCAAGTTTTCGCAACGACTTCTCGACTTCAT GACAAACGAAATCGAAGATCATCGAAGTAATCCGCAACAAAGAAATCATTTA TCGGCACACGGCTCGATGGAATTCGAACTCAGTTTACATGCTATTCCC AGTCAGCGCCGCAAAATTCCTCAGTATATTTACAGATCAATCATCATAT CAGGTACATCCAGACTGCTGGAAGCTCTCAAAGACAGTATGCTTTT TAAGAAAGGACACCCATCGTTAGCAAGAACTGGAGACCAATCGGCAT CACGTATGCATTCAGAAATCTTCATGACTTTAGTCAACAAAGCGTTAC AGATGATCCCAATGTTCCAGCAATGCAAAAGGTTTCGTTCCGGAGC AACACTGAGTGAGCACATTGCAGTCGGAACGAACTCTTGCCAAATCA ACCAGAACACAGTCTGAAATGTTCCAAACAGCAATCGATTTCACGAACG CTTTCGGCACAGTTCCTCATCAATGATCTTTGATTCTCTCGAAGCGAAG AAAGTTCCGATTGATCATCAATCTGCTCAAGGACCTCTACAAAGGAG CAAGAACGGTATCTACAAGACATGCACACTCCGAGATAGTTCCGGT TCGAGAGGTGTCATCCAAGGCTGCCACTCAGTCCAATCTCTTCAACT GCTGCTAGATCCTTTATATATGAGTCCAGAGGAGACACTTTGAGGA CGTTACAGATTCCAAGACAAAGCAGGACAGATTTCAATTGCCATTCAA GCTTACGCTGACGAGTTCAGTATCTCTCCAACACATGAAGGAATGC AAAGAATCTTAAACACAGTAGATGAATTCAGAAAATTCGGAACCTCAA AGTTGCAACACAGAAATGCGTCACACTTGCCAAAACATCCACTGCAATC CAACCTTTCGGCATTTGTCAGACGAAATCCCAATCAAGACGAGCATGG ACACATCACATATCTTGAAATACCAATCTCTGGAACAAAGACATCAAG ATTTGCAGCTGCAACTGGCATCTGGAAGGTTCAAGGACATGATCAGATCAG GTCGCTTCGCGTCACATCTCGCTCTCTCAGAAAGATTATCGCTCTCAG AGTCTTCATCTTGCCACAACCTTGACTTTTACATGTTCCACAACGATTTCAG					GCCCCA CCTCCAC CCAAAGT AGCTCG CTTCGCT CGCTCA CCTAAA ACTTTGC TCGCTC GCTTCG CTCGCTC GTCTTAA CCCTTTC CGAATA AACACTT ACAAATC CCGGCT CGCCCC ATTTTTT (SEQ ID NO: 1295)	GCVTPKMTSLVRLKL LKGGELENKKLTPL ATRYAATGNTADKI KVDYLTLCNAALRE MIYNNHSESKFMT AENGEDTAPPRISE DTRDRIQKAANEIKG TLIKVVKHISHARCLK DSTRDDEHNKFVEMI AKINDLRDNKFEQY NIEEFQGPISDQSILN IVNTEDNNEFIKKMD YINRILGTPQDASPYA RKKLQACFADNPTKT LRNIIADKVPQQSLK PSEYLDYVGPQWAN EAEGYENFLHHDYAL PERYQGVFANDFLDF MTNESKIEVIRKNKH LSAHGLDGPNSVVM LPVSAAKFLSLFRSII ISGHIPDCWKLSKTV MLFKKDDPSLAKNW RPIGITSCTYRIFMTLV NKALQMIPMFHAM QKGFVRGATLSEHIA VANEVLCQSTRQSE MFQTAIDFTNAFGTV PHQLIFDSLEAKKVPD SIINLLKDYKGARTAI YTRHAHSEIVPVRRG VIQGCPLSPILFNCCL DPLLYAVQRRHFEDG YRFQDKAGQYSIAIQ AYADDVLVISPTHEG MQRILNTVDEFQKIA KLKVAPOKCVTLAKT STAIQPFRIQPDPIPIK

AGTCAATGACTTGAAAGCGACAGATCAGATGATCCGAGGCGCTGATCGA CAAGAAAGCGCGACGTCAACATTCGGTTTCATTTTCTACATGCCGA AGAACAAGGCGGCTTGACTCGTTAAATGGAACTTCGCCAGCCTCA GCTCGTTCTACTAAATTTGCGAGGTTATGTTAAGTCAACAAGCAGAA ACCAAGCCTCTTTCACACAATGGCTCAAGAAGAGAGTCAATCCGCA AGTCTGGAAGACCAAGAAATGTTCTTAGGCATCAAGATGAAA ACGGCAAAATGTCAGAAGAACGAAAGATCCAAACGCACAAATGTTT CATCACAGGCGCTAAAGCAGCAGACAACTGGAAGTCAGATTCAA AGAATGGACAAGGAGGATACAGTCAAGTCAAGGTTAGGAGAAAATG CAACAGACTGTACCGCTCGAAACACATCGGCCAAATCTCACCTTAATC GGTCGCTCATCCAACAGAGGCGAGTACGAGGAGTTCAAGAAAGACGAA ACACACTCACACACTTCTCGGAACAGCAGCGCTAGCGGAGTCACAG ACATCATGAAGAGACCAACAGCTGTTCCAAACAACTCTACTCAGCGGC TATTGCTCTCGGTACAAACACAGCTCAACCCACGCAAAACATGCACCTCC ACAACCCAGAAGTTTGGCTAATTGTCAATTGTGCGGATGCCAATCCTGC ACTCTTCCACACATGAACATGTGCAGAAACCGTTTCAGTCTATACAA ATGGCGCACAAATATCATATCGATGACATTTACCAATTCATTCACGATC ACTATCCAGGAGTAACCATCAATGCTCGGCGAGAAATACAAGTGACGG CTACCAACAACAGGCCAGAGCTCGACGACACAGTTAAAGATCTCCTC CCAGACCTTGTGTACGATGAAGCGAACAAAGATGATCAAGATCATTTG AAGTCACATGCCCTTACGGCAGGACAACTGTTGGCACTCTCTGA CGGGCATACGACAAAAGGTTAACAAAGTATAAGAGCCTTGTCTGAACA AACAGAGAGATTATTAAGTGGACCACGCGCTCTCAATTCATCGTAGTCT CATCACTAGGAGTATCCCTCTCGGTACAAAACCTCGACGCGATTGAGAATC TCACCTGCAGATCACATACAGTACTCAAGAGACTTTCGATGCACGCGA TAGCTGCGAGTGCTTGCAATGTTTTGAAAAGTGCCAGAAATCTTCGGT ATGCGTGCGTCCCTCCAGGACGAGTCACAGTCCCAATGCAGCGA TCCACCAACAACAATGAAAACAATAACGACACAGATCATGGTCAGGA GAACCAACAGGCAACCTCTGAAGAGCAACCAACCAACAAATGGAATGC TCAAGAAGACAATGGCCAGGCGAACAAATAAATAATTCACCCGAACA AACTATCTCTGTTGATCAATCATCGAAGAGATGCTGAGAACAAACGCG ATAGAAACAAGCCTTAGACCAACCCGATGAGGACGAAATTCCTTAACCTAAG AAGAGATAAGACGAGTGAGAAAGAACAGAGCATAGTAGGATTGGCAG AGCTTAAGCGATGTCACTCGGTACGAAACGTTACCAAAACACCGGATTC CGTGTAGGAATCACAAAGCCAAAATAAAGAGACACACCGAAAAATTAAT CACCTCCCTCAAAACAGATAATAATTAACCTCCCATCCATCAGTCCGT ATGGTCTGATAACAGACTAGCACCATCCATGATACACTCATTTGGAGT GAAAACCAACAACAAATCCACCTAGACCAAAATCCTGCCCCACCTCCA					TSMDNITYLGIPIST KTSRFAAATGILEKVK AQIRVVFAASHLALSQ KIALRVFILPQLDFYMI FHNVRVNDLKATD QMIRGLIDKEAPTSNI PVSEFYMPKNKGGF GLVKLELRQPLVLTIK FARWLSQQAETKAF FHTMAQEEKSRKVV EDQENGFLGIKMEN GKIVQKNERSKRTNC FITQAAKAADKLEVR FKEWDKGGIQVRGV GENATDWYRSKHIG QISPLIGRVIQQRQYE EFKKDETHSHTFCEP AALAESHDIMKRPQA VPNNLYSAAIALRTN TAPTANMHFNPE VLANCLPGCQSCSL FHTLNMCRNRFSLYK WRHNIIICDDIYQFIH DHYPGVTIKCSARIS DGYQTTGPELDDTVK DLLPDLVVYDEANK MIKIEVTCPYGTDDN VGNSLDAAAYDKKVN KYKSLAEQTERLFNW TTTLSIIVVSSLGVIPLR TKILDALRISPADHIQL LKRLSMHAIASACIV FEKVPFEFFGMRCRPL PGRVTAPNAAIIPPNN NENNDTDHGOEN QQATSEEQPTNNGN AQEDNGQGEQINNS TEQTISVDQIIEEDAE

NeS L	NeSL- 2_CBr e	.	Caenor habditis brenner i	CCCAAGTAGCTCGCTTCGCTCGCTCACCTAAACCTTTGCTCGCTCGCTTC GCTCGCTCGTCTTAACCTTTCCGAATAAACACTTACAATCCCGGCTCG CCCCATTTTT (SEQ ID NO: 1049)	GACTCGCC TTGGGGAA GGTWTTC AGGGKSA ATTGCCG MAGGCAA GGCAGCCC CCSMMTA GCTTACAA AGTAAGTA CMCATTTT CATTTCTT GTGAATTC TTTAAACA TATTTTCT TGTTTTT GATTTCTT TTTTCTTA CCTTCCCC CAATTCTT CCCTCAT CTTGTA TACATCCC CCTCCTCC AACCAATC AATACTT GACTCTC TCTTCTGT CAAAAAAT CAATACTA GTATATTG TCCCTTGT ATAGTATT ATTGACG TCGTCTTT	TAAACC GGCTCC TCTGGG AGGAGG TATGTCA GAGGAC ATTCTCC GTGGGC GGATGG GAGGAG TAGGGT AACGAC CCGTCTAT TCTGGA TGCCTA AACCAC CACAAAT CTGTCA AGGCAA AGTGCC CCAAA GCACAC GCGTGG ATCGGT TTGGAT GCCGAC TGAGCC AGAGGG CAAAAT CGAAGG CCGTA GGCTCC CGGCGG GTTGTC CGTCAT	NNAIEQALDQDEDE FLN (SEQ ID NO: 1417) MRVHSNXPAXRTS DNXWRSXKDVRRP DPSTIEKSRYNRSIGI PDSLKXRSSAVRSXS XPPSGPDVRLXNSP SLDRRLVDCETTL GSYREWTDKPMIMG KMTYAAVTKRPPRP QTGGARLSTNLLADE MEIKYRDTNDIRLVID LPNPHLIKPLCKSCIS ARGGANALKYMKR HIADAHHLNADFVYK CSRCQHEPENVCV AKWIVNHLKRVHGY TLEDAVSTAKPSTRQ QJANAFNDSAPFIDA RKTSVPPEKKSREAG LEKFLAPTKSEDTREK TPPSTRKSESSEASI QSTIQETLSESDTLT VQEINIISSEDEMDEE PKRRVNVWALIHE NGKDAWIDSDLMI FLESARGYESCIIDP LNFICTDMISYLTIVR RRMEEGYKKIIFPLCA NDHWLTVITGSTAT FYDPMGNEPTETVKK MIDELDLEMQLAPS NSPRQRDSWNCGVF VMKMAEAYIKDTQ WDLTDVDTDVKTFR RSLLTAKAKFNIFAE DIQTYRPPSRKALTRN

AGATGATCGATGAGCTCGACCTTGAATGCAATTAGTACCCCGTCAAACTC TCCTAGACAGAGAGACTCGTGGAACCTCGGCGTTTTTCGTCAAGAAATG GCGGAAGCGTACATCAAGGATACGCAATGGATCTCACGGACGTAGAC ACGGACGTCAAAACGTTAGAGGAGCTCTTACAGAGCTCAAGCA AAGTTCAACATCTTTCCGAGGATATCAGACCTACCGGCCACCTCAA GGAAGCCTTAACGAGGAACAGCCATCGCTGTGTTGTGTCACAA GTGCTCGGCCAGCCACACCGATCCAGGATGTAGCAGAAATGGAAGT GGAAGAGCGCCAGTGTGTAACGACTCTGAAGAGCCTCACAGGA ATGGACCTTCGTCGGAATAAAGCAAGCGTGTGACAAAGCGGAAC CCCGAACACGTCGCCGGAAGCAAGCGACCGCTTTCCACACAGTACCC CTCAAAACATCAGCCAAAGATGGACCTTTCCAGAAAGAGAACTGAAA AGATGGAGGTCTCAAGTGCCGACGAGGTGAAGAACTCTACCCCTCCAA AACCACCAAGATACCTAATCTTCGCGATGAAATCGCAGTCCCGTA CCTCTGAAGAGAGGAAACCCGTCAAAGAAAGCACGGTAAAGGACACATG ATGAATACAGCGAGAAAGGGTCCGACAAAGAAAGAGATGCCCAAAGG GGAACCAAGCAACCTTATAGTTAAATCAGAAGCTGTTGTAGCAAA CTGAAGATGTACAAGGATGAAGGGTCCAATCTACAAAGACTCACATGGT TATCGGACTCTCGACCGCGCCATCGGAAAGCATTCAATGGCAACAA ATACATAGTGACCAATCATCAAGAGAAACCCACCACTTGTGAA AAGGCGCAATGTCAACACAGACAGCCGAAAGAGAGACGAGTCAA GCCAAGGGAGAGAAATGCCCAAGAGCCCAATGAGCCGCTCGTATTCA ATATGCCAAGAAATAGGCAAAAAACGTTCTCAAGATCATTTGGGAAACAG TCTGAACAGTGCACCATCAACATTGAGACTGTGCAACAGCACTCCGAA AAACACTCAAGGCTCCTGTAGTCTCAGAGAAATGCAATTAACCGGTCTG CGGAAGCATCAAGAAATGATGATGCCAAAGACCATAGAAAGACCCAAT CTCCTCGTAGAAGTCAAAATCCATCTTGACGAAAGTGAAAGACAGTCA CCAGGAACAGATGGTGCAAGTATAGCAATCTACGCTGGTTCGACCCAG AAGGGGAACGCTCGCCAACTGTTGCAAGAAATGTCGCAAGCATAGAG AAATACCCAGCCATTGGAAGGAGGGGAGACGATCTTATTGCCAAAGG ACTGTTAGATGAAGAGAGAAAGGCGGAAATTTGGCGTCCCATCG CCCTCATGGCAACAATCTATAAGTTGTTATTCAGCAGTGTGGAGCAGAAG AATCTCCGGTGTTCAAAGGGTAATTAGCCCGTGCCTCAAGAGAGGCTTTCA TCCCTCGACGGATGCAATGATCGATCGGAATATTGCGCATGTGTATTG ACACCGCTTCGCTACTCAATAGGAATCTCTTGTCTCATGGCTTGATCTC ACCAACGCTTCGGGAGCGTTCCTACGAGCTGATAAGGAGATCCCTAG AATCATTCGGATATCCACAAATCAGTTATCCAAATCGTGAAGTACATGAC AAGGGAGCAACGATGAAAGTCAAAACGGCAGATCAAAAGACGCAAG CATCAAAATAGAAAGCGGGGTGAACAAAGGAGACCCCATTTCTCCAAC	GTATTAGG AGTAGGTA ACAAWCT GTGTATGG CTTCAAAA AGCATGCA CAAAACWC CTGTCAAA AAGTAWT TCCCATCM TGTGAATA GCTCAACG ACWKGA GMCCAAT GAT (SEQ ID NO: 1173)	AGTCAG TGGTGC GCCTAC ACCCAA CTGCTAT GACACA CAAGGA CAACCC AAAATA AATAAG CAAAGG CGCGCT TAGCTTC GAGCTA ACAAGC TCCCGG AGAGGA TGGTGG CCACAG GGCACC ATCCTG GGGAAC GACCCG ATCTTC GGATGC CCAACC ACCGCC AATCTGT CAGGCA ACGTGC CCCAAA AGCACA CGTGCG GAGCGG TTGGAT GCCGAC TGAGCC AGAGGG	SQSPVVVCHKCSRPA TPIQDVSRMEVEEAP VLVPTPEPPQEWTF VGKNRKRGVTSRTP NTSPEAKRPAFPVP LKPSANRWHFPEET EKMEVSSADEVKNST PPKPKIPNLLAMKIA SPVPLKRGNPSSKHG KGHMMNTARKGPT KKEMPKGEPANLIVKI RSWFDEQLKMYKDE GSNLQRLTWLSDSLT AAIGKAFNGNKYIVD QIIKRNPPPLVEKGA MSTQTSRKRDEFKPR ERMAQEPNEPLRIQY AKNRQKTFKIIGKQS EQCTINIEVEQHFRK TLKAPVVSENAIVTC GSIKKVLMPTIEDPI SSVEVKSILTKVKDTS PGTDGVKYSNLRWF DPEGERLAKLFECCRK HREIPSHWKEAETILL PKDCSDEEKKPEN WRPIALMATYKLYS AVWSRRISGVQGVIS PCQRFQSLDGCNES IGILRMCIDTASVLNR NLSCSWLDTNFAFGS VPHELIRRSLESFVGP QSVIQIVTDMYKCAT MKVKTADQKQTSIKI EAGVKQGDPISTLF NICLEGIIRMHQMRE KGYDCVGHKVRCLAF ADDLAILTNKDEM

CCTATTCAATATTTGCCTTGAAGGCATCATCAGGATGCATCAGATGAGA GAGAAAGGGTACGATTGTGTCGGGCATAAAAGTTCGCTGCCTAGCGTTC GCCGACGACCTTGCATTCTAACGAACAACAAGATGAAATGCAGGAA GTTATCGACAAGTTGGATGCAGACTGTAGAAGCGTGTGTTGATCTTTA AACCTAGGAAGTGTGCATCTTGACTATCGTGAGAGGTGAGTTGATAA GTATGCAAGATCAGAATAAATGGAGACGCGATCAGAACAAATGGCGGA TAGAGACACTATAGATATCTGGGTGTAAGACCGGAGTTGGTGAAG AGCATCGGAAACGGAAGCTTTAATCAGGTGGTCAAGGAGCTCCAAA GGTCACGAAACCGACTGGCTCCACATCAAAAACCTGACATCTCTGAAG ACGTTCTTACTGCCAAGACTGCAGCATCTACAGAAATGCCACTCTCAA ACTGTCAGAGTTGAGAGATTGAGAACGTTGTTATGAAATCAGTGAA ACGGTATCATAACATAACCAATAAAGGGCTCGCTGTGGAATATGTCAA ATCCCTGTCAAGAGGGTGGACTAGGAGTTCTATCTCTCGACTCACAT GCCTGATCATTCTTACCTCGACCTCTGCAAGCTATGTCGCGATGAT CCATTATATCTTCTATCCACAAAGACGCACTAAGCAGAAATCACAGTGAA AGCGATGGGACTCACCACTCAAAAGTCCACAAATAAAAGAGACATGTGA GTACTTAAACACAAGGAAAGCTGTACGAAAGGAGGATATAGTCTATTC TGCCGATGAATGAATCTCTCCGACGCTGTCTGTCATCCAAGGTGCTCC ACTGAAATCAATGGAATTCATCCCGGTGAACAATGAAATCGGTATAGCG GTACAAAGCCAGGAGGATCCGAGATCAAGTCTTCAAAAGGCTGACA GCCTGAAGCTAATGAGTAAGCTAAAAGATCTGGTCAGATCTGCTATGCT CAACGGTCTCTCGAAGAGAGAGTGTAAAAGCAGAGTCACCCCAAGT ACTCCAAACCCACCAATCCAATAGATTCGTACGGGACGGCCGAAAC TGACGATAGCAGCWACAGATTCGTGACCCCKGCCCTGTGAACCTCC TCTCTGCAACGCCAACACATACGATGTTAACCATCCAAAGGCTGCAG AAGTGCCAGGCTGACTTCGAGTCTCAGCAGCACATCTGCAGAACTGT CACTACAGTTTGGCAGGGGAATAACCCAGAGACATGACAGAGTCATG AACAGGATCTGCAGGAAATTTGAAACGGGAGAGAAAGCTCACTACAAG ATAATGGTGATATGAAACCGGGCCCAAGAGAAAGACCGGATATC ATCATGGAGGAAAGGACGGTCCAGAGTGTACTAGCCGACGTGACA GTGCCCTACGAGAAATGGGGTTCAAGCGGTTGAGAGGGCGTGGGATAA GAAGATAGAAAAATACAAACGACTTCTAGATTACTACCGCAAAATCGGA AAGAAGGCTACGATTCTTCCCTAGTAGTCGGTAGCCTAGGAACCTACT GGCCGACACAAAGCCACTCACTGAAGATGCTTGGCCTTCCGACGGTCA AATAAGGAATGTTATACCTGAAATCTGCCAAATGCACTGGAATCTCTCA AAAATATCTATTGGAAACACATCTCGGTGATAGCTACAAAACGGTGA GGGACTATTTGTGAGAGGAAATAACAAAGAAAGTCCGATTCGAAGGAAA AGGTGAAAAACACCACGTTGTCAAAAGATTCCAACCTCTGAAATGTGAA	CAAAGT CGTAGG CCGGTA GGCTCC CGCGGG GTTCTCC GTCGTA GTCAGT GGTGTT CCTACAC CTAACT GCTATG ACAAGC GTATAG GAGGCC CGGAAA AACAAG CCAAGG CGGCGT TAGCTG AGAGCT AACAAG CTTCTCG TGGATG GGTGCC AGAGGG CACCATC CTGGTG GGTGGA TGGGGG GAGCTT GGGAAC GTCCCG ATCCTTC GGATGC CCAGAC CACCGC AATCTGT	QEVDKLDADCRSVS LIFKPRKASLTIVRGA VDKYAKIRINGDAIRT MADRDYRYLGVKT GVGGRASETEALIQV VKELOKVHETDLAPH QKLDILKTELLPRLQH LYRNATPKLSELEFE NVVMKSVKRYHNIPI KGSPVEYVQIPVKKG GLGVLSPLRLTCLITFLT STLCKLWSDDPFISSI HKDALSRITVKAMGL TTQSATIKETCEYLNT RKAVTKGGYSLFCRM NESLRTLSVIQGAPLK SMIEFIPVNNIEIGAV QATKDSEIKVFTKADS LKLMKSLKDLVRSAM LKRFELEKSVKSRVTQ VLQHHPQSNRFRVD GRNCIAAQRFFVHPA RLNLLSCNANTYDVN HPKGCRRCCQADFES QQHILQNCHYSLAG GITQRHDRVMMNRIL QEIGNGRKAHYKIMV DMETGATRERPDIIIM EERDGPVLLADTVV PYENGVAVERAWD KKIEKYKHFLLDYRKI GKKATILPLVVGSLGT YWPDTSHSLKMLGLS DGQIRNVPIECIJAL ESSKNIYWKHILGDSY KTVEGLFCQRNKEV RFEGKGEKHHVSQRF QPLKCEKVRMTMKSTK

AAGGTGCGTACAATGAAAAGCACAAAAGAGAGGGGTAGAAGTAGATC GAATGCCAAGAAAGGTCCGAACTGGCGAAGATCAAAAAGCGAATCGGA CGGAAGGAGTGTAGTAAAGGAGATACTGGCGAGATCCGTGCAACAA GCCGCCACACTCGAAGATGACCCAGTCGGCTTTAGCTAAGCGCTAAACC GGTCTCTGGGAGGAGGTATGTAGAGGACATTTCTCCGTGGCGGAT GGGAGGAGTAGGTAACGACCCGTCATTCTGGATGCCTAAACACCCAC AATCTGTCAAGGCAAGTGCCCAAAAGCACACGCGTGGATCGGTTTG GATGCCGACTGAGCCAGAGGGCAAAGTCGAAGGCCGTTAGGCTCCCG GCGGTTGTCCGTATAGTCAGTGGTGCCTACACCCAACTGCTATGA CACAAAGGACAAACCAAAATAAATAAGCCAAGGCGGCTTAGCTTCG AGCTAACAGTCCCCGAGAGGATGGTTGCCACAGGGCACCATCCTGG GGAACGACCCGATCTTCGGATGCCAACCAACCGCCAATCTGTAGGCA ACGTGCCCAAAAGCACACGTGCGGAGCGGTTGGATGCCGACTGAGCC AGAGGGCAAAGTCGTAGCCCGTAGGCTCCCGGGGGTTCTCCGTCGT AGTCAGTGGTGCCTACACCTAACTGCTATGACAAGCGTATAGGAGGC CCGAAAAACAAGCCAAGCGGCGTTAGCTGAGAGCTAACAAAGCTTCT CGTGGATGGTGCCAGAGGCGACCATCCTGTGGGTGGATGGGGGA GCTTGGGAACGTCCCGATCCTTCGGATGCCAGACCAACCGCAATCTGTC AAGCACCGTGCTCCAAAAGCACACGCGCGGGTTGGTTGGATGCCGA CTGAGCCAGAGGGCAAAGTCGTAGGCCGTTAGGCTCCCGCGGGCTCT CCGTATAGTCAGTGGTGCCTTACCCAACTGCTATGACATGCGTACA GGAGGCCCGAAAAATAAGCCAAGGCGGCTTAGCATAGGGCTAACAA AGCTTCTGTGGATGGTGCCAGAGGGCACCATCCTGGTGGGTGGATG GGGGAGCTTGGGAACGTCCCGATCGTTCCGATGCCCAACCAACGCA TCTGCCAGGCAACGTGCTTCGAWGGTCAITGGTCTAGACTTGTAAATA GACCATTTGCCGGAAGAGCACACGCGCGGTTGGTTGGATGCCGACCGA GCCTAGAGGGTGCAAACTGAAGGGCGAGGTCGAAGGCCGTGAGGCT CCCGCGGGAAACTCCGTATAGTTAGTGGTGTGCCTACACCCGACGAC TATGACACATAGGAGGAATCCTGATCTGATATGATCATGTATATAGGA GGCGAAGGTAATAGTCAGKGTCAAAGTCCACGTGGCAGCTACTCCC CAGCATAGTAGTGATCGGAGTGGAWCCAACTTTGACACTGATGTTCCCT GAGCCTGACCCATCTGCACAAATCCAACAGTGTATGATGGCCACACAC TGAGGACGAGTATCACTTGTGATACTCAGAGGTGTCCCCCATGATCAAC CAATATCACAGTAGCGGACCTACCGTAGGTTAGACCCCGCGGCTGTA GCAGGCTCGCCTC (SEQ ID NO: 1050)	CAAGGC ACCGTG CTCCAA AAGCAC ACGCGC GGGTG GTTTGG ATGCCG ACTGAG CCAGAG GGCAAA GTCGTA GGCCGG TAGGCT CCCGGC GGGCTC TCCGTCA TAGTCA GTGGTG TGCCTTC ACCCAA CTGCTAT GACATG CGTACA GGAGGC CCGGAA AAATAA GCCAAG GCGGCG TTAGCAT AGGGCT AACAAAG CTTCTCG TGGATG GGTGCC AGAGGG CACCATC CTGGTG	EGRSRNNAKKGPN WRRSKSESDGRSVSK GRYWRDPSNKPPhS KMTQSALAKR (SEQ ID NO: 1418)

238

239

NeS L	NeSL- 2_Cre m	.	Caenor habditis remane i	CCAACTCTCATCGTATTAACTACGGTATTCACTCTCTAGTGAGTGAATA AAGGTTAATTACGTTTCTCTTTCMAGAGAAAAAGAAATTCGAATCCT TTTTGTGTAACCTCACAACACTGACAGAGACCTATCGAATTTCTTTGTTT GTATATAGGAATAGTCACTCTGGACCACGAGTGGACAGTTGTCGGCG GACTTCCAGAGTGGAGAGAAAGGTGTGAAGAGAGAGAGGTCTAGAAA CACTTCGGCTGTCTAGGACCAGTTCCTGAGTGGAAGAGGAAGGTCTA GAAACACTTCGGCTGTCTAGGACCAGTTCGTGAGATCTCTCTGGAGAG TTGAAAACAGTCAGTCAGGCTACTGTATTTCTTGATAGCCCCGCCCA ATCCCCCTCCCCCTCGACAGATTTTCTGTTTGACCTCTCTGGAAT TGCGAGAGTGCAGAGAAATTTTCGAATCTTCGCGCGTTTCTCGAAA TTTTCCAGAGATTCGAGCGGAGAAATCTTCGAGAAAGTGAGCTGAATTT CGCGGAAATTTCCGCGATTTTCAAATATCGATTTTGTGCGGAAATTT ATTTCTGGCAAAATTTGATTGAGTTACGCGGGAGAGAGAAATTTGTT GGAAGGGTATTGATTTTGTGGCGGAGGAACTCCCACTGAATCAAT AACTCTCAAAGGAGAACTCATCGAACAACTCGGGTGAACCTGAATCTTG GGCGAAATTTTCGATTGACACAAGATAAACAATTTACTGTGAAAAT AAATCAGAAACAACTGTCAAAAAGAGAGACAAAAGTATTGATTAAAC ACATCATGACAAATGTATATCTTAAGCCTGTGAATGATAACAGACTAAC AAAACCGGTGATAATCTAGAAAATCTATGTCAAAATAGTCAATGTGAAA TGACGTGGAACCTGTAGCCAGAACATATGCTCAGGCGAGCAGTACTAA CCCGGCCGACGACAAAACGGTGACTGCTCTGGGTGCAAAATACAATCTG	CCAACTCT CATCGTAT TAACCTAC GGTATTCA CTCCTAGT GAGTGTA TAAAGTT AATTACGT TTTCTCTG CMAGAGA AAAAAGAA AATTCGAA TCCTTTTG TGTAACCT ACAAAATG ACAGAGAC CTATCGAA TTTCTTTG TTTCTTTG TTTCTGAT ATAGGAAT AGTCACTC	TGAATA CCGTCA GATAAG CCCCCA ACATAA AAATAA AAGTCG GCGTTA GCTAAC CACTAA ACCGGC TCCTCAT TGGGGG AGAGTA TCATTCC GGTGCT CTCCGTT TGGCGG GTAGGG AGGAGT TGGGTA	ACTCAG AGGTGT CCCCCAT GATCAA CCAATAT CACAGC TAGCGG ACCTACC GTGAGG TAGACC CCCGCC GCTGTA GCAGGC TCGCCTC (SEQ ID NO: 1296)	MTNVVLPVNDNQ NKTGDSRNTMSNS QCEMTWKPARTYA QAASTNPADDKTVT VLGCKYNLLKGNTP QTSKRSPKPSRGA RISSVYLTDELETHR EEGKITFAIDLPNKNN ILCPLRECTQTRGRG SSFTKHMKLHVKEKH QLDATFIYKSCMCNE YEPEKKCGTKWQTH LQKVHNYKYDESAIV VPVPPNTRQQIANEL NNAAPFVDIRKPKAA AVEEKKKTENGALLKFL TKSNKDDQVKSPSXD IPDAESPEKETQALTI DPKGNNSPSKSIRSS QSSASSVCQEIQEITL SEDEDPKGARPKPGI

CTAAACTGGGAAATACTCTCAGACGTCGAAAGGTCGCGCTCCAAAC	TGGACCAC	GCGACC	NVWSLINETGKDAYI
CATCGAGAGGAGGAGCTCGAATCAGCAGTGTGTATCTCTGACTGATG	GAAGTGG	CGGAAG	DTDIMMAFLKMRVE
AGCTGGAGATTACGCACAGAGAAGGTAAGATCACATTCGCGATAG	ACAGTTGT	TATGGA	NCDSVNIIDPLNYQF
ACCTTCAAACAAGATAACATCTTGCCCGCTGTGTCGCGAGTGCAC	CGGCGGA	TGCCCA	PARVDLVP LIQRNLE
CCAAACCCGTGGGAGAGGTCAGTTTACCAAGCATATGAAACTCCAC	CTTCCAGA	ACCACC	DGKKRVFPICADEH
GTGAAAGAGAACCAACTTGATGCCACGTTTACATCAAGTGTAGTA	GTGGAGA	GCAATC	WTLTISNGIAAFYDP
TGTCAACGAGTACGAACCGGAAAAAATGCGGTACGAAGTGGATCC	GAAAGG	TGATCT	TGSRMSSYIEELVNEL
AGCCACCTTCAAAAAAGTGCACAACTACAAGTATGACGAGTCTGCAAT	TGTGAAGA	GGCATT	GLIIPKEQDEQPRQR
AGTTGTCCAGTACCAACCAACACAGACAGCAAAATAGCTAATGAGTTG	GAGGAGG	GTGTTTC	DSYNCGVFVMKMAE
AACAATGCTGCCCATTTCTTGACATCAGAAAAACCGAAAGCTGCTGTG	TCTAGAAA	GGATGG	AFIQDTEWEMEEVE
TTGAGGAGAAAGACTGAAAATGGTCTCTGTTAAAAATTCCTGACCAA	CACCTCGG	TCTCTGT	EDVKNFRNLLLEELK
GTCAAATAAGGACGATCAGGTAAAAATCCCATCGGAWGATATTCCAGA	CTGTCTAG	CTCTAG	PNYEIFAEEKIKYVNSP
TGCGGAAAGCCCTGAAAAAGAAACTCAGGCGCTCACTATCGATCCGAA	GACCAGTT	ATCTGA	GKSFARQSRPTSSSQ
AGGGAACAACCTCACCATCAAAAAAGCTCAATAAGATCGAGCCAGTCTCTCA	CCTGAGTG	AATAGA	CAVCPTCSRSATPM
GCTTCTCCGTTTGTCAAGAAATCCAGGAAATCATCACGTTGAGTGAGG	GAAAGAG	GCTCTG	MDVGNMEVDVPVQ
ATGAAGACCCAAAGGGGCTGCTCCAAAACCCAGGAATCAACGTGTGGA	GAAAGTCT	GCCTGA	QQETPKSREPEQDEG
GCTTGATAAATGAAACGGGAAAGGATGCATACATTTGATACAGATATCAT	AGAAAACAC	AGAACA	WKVVGKARKRGWVT
GATGGCGTTCTGAAGATGAGAGTGGAAAACTGACTCCGTGGAACAT	TTCGGCTG	CACGGG	ERSPNISPEAKRQFTG
AATTGATCCACTCAATTACCAAGTTCGCCGAGAGTGGACCTAGTCCAC	TCTAGGAC	CGGACC	PEIKVSPGKFHPLVG
TTATCCAGAGGAATCTGGAAGACGGAAGAAAGAGTCGTGTTCCGA	CAGTTCGT	GGTTGG	ETEEEMEVTCDSPTK
TCTGTGCAGACGAACACTGGACGCTCTTGACCATCTCGAATGGAATTGC	GAGATCTC	ATGCCG	EPTEPKVTPSLPAM
TGCATTCTATGATCCGACTGGATCGGAATGAGTAGTTATATTGAAGAG	TCGTGGAG	ACTCGA	KIASPEVTKKQTSKKK
TTGGTGAACGAACTTGGACTGATTATCCAAAAGGAACAGGATGAACAG	AGTTGAAA	TCTGGA	GKYGKKQXTKKAQ
CCAAGACAAAGAGACAGCTACAACCTGGGGTATTTGTGATGAAAATG	ACAGTCAG	GGGTGC	PPKGEPTKKAQPKGE
GCGGAAGCCTTCATCCAAGATACCGAATGGGAAATGGAGGAAGTAGAG	CTGAGGCT	AAACCT	PAKLIEQVRTWFDKQ
GAAGACGTGAAAAAATCCGAAGAAATCTCCTTGAAAGAACTGAAACCCA	ACTGTATT	GAAAGG	MKSYPEQGSNIQTLT
ACTACGAGATATTTGCTGAAAAAATCAAATATTATACTCTCCGGGAAA	TCTTGATA	GAAAGT	WIADSLTAAIFKANS
AAGTTTCGCCCAAAGTCGACCCACAAAGTCGAAGCAGCCAGTGCCGTC	GCCCCGCC	TGAAGG	GNKYLVDKITARCPP
TGTCCGACGTGCTCTGTTCACTACACCGATGATGGATGTAGGAAACA	CCCAATCC	CCGTGA	PLNEGEMATQTSRR
TGGAAGTGGATCCCGTTCCACAGCAACAAGAGACACCCGAAGAGTCGCG	CCCTCCCC	GGCTCC	TEAVKPKDRFVKESN
AGCCAGAAACAAGATGAAGGCTGGAAAGTGGTGGGAAAGGCTAGAAAAG	CCCCCCTC	TGCGGG	EPLRIQYAKNRAKTF
AGGGGAGTGGTAACCTGAACGATCGCCAAACATCTCTCAGAAAGCCCAAG	GACAGATT	GAAACT	NVIIGKHSARCEIDIN
AGACAAATCACTGGTCCAGAGATCAAAAGTCGTCTCACCTGGGAAAGTTTC	TTTCTGTTT	CCGTCAT	VVENHFRQTLKAQP
ACCCACTGTGGGCGAAACTGAGGAGATGGAAGTGACGTGTGACAGCC	GACCTCTT	AGTCAG	VTEEALNTVCSGIKKA
CACCAACGAAAGAGCCCACTACGGAACCGAAAGTGACTCCAAAGCTGTC	GGAATTTG	TGGTGT	KVDPSIEGPISSGEVK
CAGCAATGAAAAATTGCTAGCCCAAGAGTGACGAAAAAGCAACGTCAA	CGAGGAG	GCCAAC	AILAKIKDTSPTGDGV
AGAAGAAGGGGAAAGTATGGCAAAAAGAAACAGSAGACAAAAGAAAGCT	TGCGCGAG	ACCCGA	KYSDLKWFDPGERL
CAGCCCGGAAAGGGGAGCCCAACAAAGAAAGCTCAGCCAAAGGAGA	AATTTTCG	CGACTA	ALLFDECQRHGKIPS

ACCGCAAGCTCATTGAGCAAGTGAGAACTTGTTGTTGATAAACAGATG	AAATCTTC	TGACAT	HWKEAETVLLPKDCT
AAATCGTACCAAGAGCAAGTTCTAAACATCCAGACACTGACCTGGATTG	GCAGCTTT	AGTTGG	EEERKKPENWRPISL
CCGACTCACTCACTGCCGCCATCTTCAAGGCAACACAGTGGAAACAAGTA	TCTCGAAA	AGGAAT	MATVYKLYSSVWNR
TCTGGTAGATAAGATAACTGCAAGATGCCACCACCATTTGCTGAATGAA	TTTTCCAG	CCTGATC	RISSVKGVISDCQRGF
GGTGAGATGGCGACGACAGCAGAGCAGAGGACAGAGCGGTGAAACC	AAGATTGG	TGATAA	QAIDGCNESIGILRM
AAAAGATCGATTTGTAAAAGATCTAACGAGCCGCTCAGAAATCCAGTAT	AGCGGAG	TAATCAT	CIDTATVLRNLSCS
GCAAGAACCGAGCAAGACCTTCAATGTGATAATTTGGGAAACACTCC	AAATCTCG	TGTTTAT	WDLTNAFGSVPHL
GCACGATGTGAAATTGATAITTAACGTCGTGGAAACCACTTCAGGCCAA	AGAAAGT	ATAAGG	IRRLAFAFGYPESVINI
CCCTGAAAGCAACCACTGATCAAGCAATTTGAATGCTGTGTGCA	GAGCTGAA	GAGGGG	ISDMYNGSSMRVKT
GTGGAATCAAAAAGGCGAAGTTGATCCAAAGCATTTGAAGGTCGATCT	TTTCGCGC	GATGGT	AEQKTQNMIEAGVK
CGTCAGGAGAAAGTGAAGCGATTTCTTGCAAAAGATCAAAGATACCTCTCC	GAATTTTC	AAATAC	QGDPISTLNFNICLEGI
CGGAATGATGGAGTGAAGTACAGTATCTGAAATGGTTGATCCGGA	CGCGATTT	CCAGGG	IRRHQTRKTGYNCVG
AGGTGAACGTTTGGCGTTGTTTTCGATGAATGTCGACAGCACGCGGAA	TCAAATTA	TCCGAA	NDVRCLAFADDLAIL
GATTCGAGCCACTGGAAAGAAAGCAAGAACTGTTCTGCTACCAAAAGAT	TCGATTTT	ACCATC	TNNQDEMQRDLNQ
TGCACTGAAGAGGAAAGAAAGCAAGCAAGAAATTTGGAGACCCATCTCT	TGTCGGAA	AAAGCA	LDKDCRSVALIFKPKK
CTAATGGCTACTGTATACAACTCTACTCTCAGTCTGGAACAGGAGGAA	AAATTTATTT	GCTACT	CASLTIKGSDVDQYA
TCTCCTCAGTKAAAGGAGTCAATGATCGGAAATCTGCGGATGTGCATAGA	TCTGGCAA	GACCAG	RIKIHGMPIRTMSDG
GATCGATGGATGCAATGATCAATCGGAAATCTGCGGATGTGCATAGA	AAATTTGAT	CATAGT	DTYKYLGVQTNNGG
CACAGCCACAGTTCTCAACCGAAACCTGTCGTTCATGGTTAGACTTGA	TGAGTTCA	AGTGAT	RASESLTQIAAELQ
CGAACGCTTTTGAAGCGTGCCCCACGAGTTGATCAGAAGATCACTAGC	CGCGGA	GAACAC	MVHDTDLAPNQKLD
CGCATTCGGGTATCCTGAATCAGTCAATATAATCAATGACATGTATA	GAGAAGG	ATAGAC	VLKAFILPRLQHMVR
ATGGATCGTCAATGAGAGTCAAGACAGCGGAGCAGAAACTCAGAACA	AAATTTTG	CCTGGG	NATPKLTKEFENTV
TCATGATGAAGCTGGAGTTAAGCAAGGTGATCCCATCTCGCCAACTCT	GAAAAGG	GTTCCCT	MKSVKMYHNIPIKGS
ATTCAACATCTGCTTGAAGGCATAATCCGAAGGCATCAGACGAGGAAAG	GTATTGAT	GAATC	PLEYVQIPVKNGGGLG
ACAGGTTACAACTGCGTTGAAACGACGTACGTTGCTGGCATTTGCTG	TTTTGTGG	GACCCA	VMSPRFTCLITFLAST
ACGATCTTGCTATCCTTACCAACACAGGATGAGATGCAAGATGTGCT	CGGAGGA	TCTGCAC	LFLKWSDDDEYISSIHK
CAATCAGCTGGACAAGGACTGCTGAGTGTGCTGCTGATATTTAAGCCA	AACTCCCA	AAACCC	KALSRITAKVMGLKT
AAGAAGTGTGCTTCACTGACGATCAAAAAGGAAAGTGTGATCAGTATG	CTGAATCA	ACTTTGT	QKATLQEQCEYLNTK
CAAGAATCAAGATTCATGGAATGCCCATTCGGAATATGCGGATGGGG	ATAACTCT	ACAAAT	KAITKGGYSLFSRMIN
ATACCTACAAGTATCTCGGAGTCCAAACCGAAACGGTGGTAGAGCCTC	CAAAAGG	GAACCA	EAIRTL SVN LGAPLKS
GGAATCAGAAATCCCTGACTCAGATTTGCCGCGGAATCCAAATGGTCCAT	GAACTCAT	AACTGA	MQFIPENGEIALEVQ
GACACAGACCTGGCGCCGAAACCAAGAACTTGATGTGCTGAAGGCATTC	CGAAACAAC	TGAAGA	ASENSQIKVFSKADS
ATCCTGCCGAGACTGCAACATATGTACAGAAACGCCACTCCAAAGCTGA	CTCGGGTG	GTTTAAAT	MKLVTKLKDLVSKA
CGGAGTTAAAGGAGTTTGAGAACACAGTCAATGAAAAGTGTGAAGATGT	ACCTGAAT	GATTTCT	MLKNFLENKKVKSKV
ATCACAACATCCCGATTAAAGGATCACCACTCGAATATGTCCAAATTC	CTTGGCGG	TACATCA	VQVLQHHPQSNKFV
GTAAGGAATGGAGGACTGGGAGTTATGTCTCCCGATTACAGTGTCTCA	AAATTTTC	CAGCTA	NDGKNXSISQKFVH
TAAAGTTCCTGGCGTCCACACTGTTCAAACTGTGGTCAGACGACGAAATA	GCAATTGAC	GCGGAC	PARLSQLVCNGNSYS
CATCTCGTCCATCCAAAAAAGGCGTTGAGTAGAATCAGGCGAAAGGTG	ACAAGATA	CTACCGT	KDLPKNCRWCGYEC

ATGGGACTGAAGACCCCAAAAGCCACGCTCCAAGAGCAGTGCAGTAC CTGAACACCAAGAAAGCAATCACGAAAGGAGGTTACAGCCTCTTCTCGC GAATGAACGAAGCTATTGAAACGCTCAGTGTCAACCTTGGAGCACCGCT CAAATCAATGCAATTCATCCGGAAATGGCGAAATGCTTTAGAAGTG CAAGCATCAGAAAACCTCAGATCAAAGTATTTCTGAAAGCTGACAGTA TGAAACTGGTGACAAAGCTGAAAGATCTGGTGAATCGGCGATGCTCA AGAACTTTGGAAAACAAGAGGTCAAAGCAAGTTGTGAGGTTGC TTCAACACCAACCAATCAAACAAATTCGTAATGATGGAAGAACWK CAGCATWTCCTCCAAAGTTCTGACACCCAGCAGGCTGAGCCAGCTG GTCTGCAACGGGAACAGCTACAGTAAAGACCTTCCGAAAACTGCAGAT GGTGGGCTACGAATGCGAGTCTCAGGCTCACATCTCCAGCATTTGCAC ATACAGCCTTTTCATCTGGAATCACCCAGAGGCATGACCGTGTCTGTAAC AGGATCTTGASGAGGTGATAAAAGGCGAGAAAAACAACGACTACTAT GACATAATGGTGGATACGGAGCCCGGACCAACAGAGAGCGTCCAGAT ATCATCATGATACAGAAAGATGGTCCGGAAGTCTTACTGGCGGATGTTA CGGTACCATACGAGAAATGGAGTTGTTGCGATCGAAGCCGCTGGGATT GGAAGATGGAGAAAGTACAGTCACTTTATTGATTACTTCGCAAGACTGG AAAGAGAGCAGTAATCTTCCACTAGTGGTTGGAAGTCTTCTGACCTAC TGGCCAGACACCTCAAACCTCTGAGAAATGCTGGGTCTCTGATGGC AAATCAGAAACCTGATCCAGACATCTCCATGATTGCTAGAGTCTCC AAACAAATCTACTGGAGGCATATCTCGGAGATAGTACAGAAATGTA GTGATCTATCTGCAGAAAAACCCAGCAGGAGATCAGATTCGGAGATG AACCATGGAATGTTCAAGTCTCAGATCGATTCAGCCTTTTAAACA AGAGAGCGTGAGAAATCCGAGGAAGAGAAAAAGAGAGATCAAA GTCCAAAAAGGCAAACTTGGCAGGATCCAAAAACAACCTGATTC CCGGCAATCCGGCAAAAGCAATCAGAAATCAGGGCTTCCAAAGAACGCT TGACAAAGCGGTATCACGGTGAATACCGTCAGATAAGCCCCAACATAA AAATAAAAGTCGGCTTAGCTAACCACTAACCGGCTCTCATTGGGGG AGAGTATCATTCGGTCTCTCCGTTGGCGGTAGGGAGGAGTTGGG TAGCGACCGGAAGTATGGATGCCAACCCAGCAGTCTGATCTGGCAT TGTTTTCCGGATGTTCTCTCTAGATCTGAAATAGAGTCTGGCCTG AAGAACACACGCGGACCGTTGGATGCCACTCGATCTGGAGGGTG CAAACTGAAAGGGAAAGTTGAAGGCCGTTGAGGCTCTGGCGGGAAC TCCGTATAGTCAGTGGTGTGCCAACACCCGACGACTATGACATAGTTG GAGGAATCCTGATCTGATAAATCAATTTTCATATAAGGGAGGGGAT GGTAAATACCCAGGGTCCGAAACCATCAAAGCAGCTACTGACCCAGCATA GTAGTGATGAACACATAGACCTGGGGTCCCTGAACTCGACCCCATCTG CACAAACCCACTTTGTACAAATGAACCAAACTGATGAAGAGTTTAAATGA	AMACAAA TTACTGK GAAAAATA ATCAGAAC AAACTGTC AAAAAGA GAGACAA AAAGTATT GATTAAACA ACATC (SEQ ID NO: 1174)	GAGGTA GACTCC CGCCGC TGTAGC AGGCTC GCCATT G (SEQ ID NO: 1297)	ESQAHLQHCTYSLS GITQRHDRVLRILXE VIKGRKNNDYYDIMV DTEPGPTRERPDIMI QKDGPEVLLADVTVP YENGVAIEAAWWDW KMEKYSHFIDYFARL GKRAVILPLVVGSLGT YWPDTNSLRMLGL SDQIRNLIPDISMIA LESSKQIYWRHIFGDS YRIVSDLYCRKDQOEI RFGDEPMENVQVSD RFQPFKTREREKSEE EKKRRSKSKGKTWR GSKKQTDTSRQSGKSN QNQGFQSVGGQVS R (SEQ ID NO: 1419)

NeS L	NeSL- 4_Cre m	chrUn	Caenor habditis remane i	TTTCTTACATCAGCTAGCGGACCTACCGTAGAGTAGACTCCCGCCGCT GTAGCAGGCTGCCATTG (SEQ ID NO: 1051)	CCCTTTTCT ATCGTATT AACTACGA TAACCGCT CATTTGAG TGTAAGAA AGGTTCCC CCCTCCTC GCTGCTCT TACCCACG CATCTCTG CCTCTGGG AAGGCGG AGGGTCAA CTTGCGGG TCTGTGGA TTTCTTTT CTATCCAC CGCCCAT TTCTCTGT CGAAAGCC TACCTAGA TCAGCCGG GAGTTTTT CCTATCCC ATTCAGGC GATCGCTC AAGGCTGT TTTATCGA CACTCCTT CTTGACAA GTATTTAT TTCTTGAC AAATTCTA TTTTCTCT TTATCGAT	TAACAT GCCTTG GAAGGC ACCACG CCAAA GTCCTG GCAACT GATTTG ATAAT GTATA AAGTAA CTGGAA CCAAAT GCCCCA TAGGTA GGGCGG GAGAAA ATGACC TAGAAA ACACAA AGTCCC AAGCCC CCGGAT TCGAAA GACCTA TAGGAA GTCAGT GAATAG AGAGAA ATATCA AACAAA TCTCAC CATTCAC AAGGAC TACTG GTCGAG	MTKTEWSWRHRSR RSVGIVVKIDTSDYAN VRVHVAADLSNEDG HTSHNNGIILPIMKP SVDFCQIQYPPRGY YVPHQSQKGHDAK PSRHWNEEAQPPYV HNNHGRGRSAKP SGRRPPRPILQEESL AAHPQIPGDTASAVP LYSDVWNNENKSQG KPPQGSRRSRPGT KPSVPVGEAEQETNS RPIAPEPIVKFKHDKH GWTTVQGSSSHGRP VPKPSVPVSEANRF QLLQEGDFPLTTSES SQEEIKVPNYQIRVSP IPLPSEEDSKLPTKSNY RAPKGRKSRNYKKPQ QQNPKYQQLPVQ PKVNNAPTDRMAPE QLKGGGGKTAHNDI EEMEIEEDTDEKIQV KRIKVNKLTPHFVC MMTYPTDNVYRCFV KGCTATSQGGWGAE DLKYLTVHIREHKIK VEWTEYECIGCDLSG GAGKHSKWIKPHM RKKHNRDAPTNFKM GSRSGKPKITELLEES APSCSNPRRKTINQK KTAIITQVTPKLTGT YQTRSVTKALSVLKES RQKELEVLREEKAN

TTGCTCGAATCCAAGGAAGAAAAACCTCAACCAAGAAAGAGACTGCTATA ATCACGGAAGTCACTCCGGAGAAATGAAAAACGGGCTATCAACAGAGA AGTGTCAGGAAGGCTCTCAGCGTCTGAAAGAGTACGACACAAAAAGAG CTGGAAGTGTGAGAGAAGAAAGGCTAACGCTAACACAAAAGTCT AACTTCATCTTTCTTCCACCAAGCCCTCATATAGATGTTGTAACCC AACGTACGGAGAGAACTATCAAAAATGATTACTCCGGAGGAGAACAA TAAGGGAACAAAGATACCAATGGTCCACACCAAGCGCGTCTCATCCAA AAGATAACAGAAAAGCTAAAAGGCTAAACCAATGCATCTTGACGAA AGTACCATATAGAAGCGTCAAGCTCGACGTCATCATATTGACGACG ACGACGAAGACGACAAATGACACCAATGCGAAGAAGATTCAACACTT GGTGTCTTGACCCAGACGAGACTCAAGAAAGCATGTTAACTGACGACG TAATCAATTGGTACTTGAAAGACCTATGCTTTGGTAACGAACAATACAT GCTCGTAGACCCACTAGTATGGTGATATACAAGATGGGAGGAATGGC AGGCGTCGAAACAAAGGTTCAAAAGCAAGAAAGACGTGCCTATTCCCAATC TGCGAAGCTGACCACTGGATTCTTCTTGATTTCGATGAGACCAACTTGTG CTACGCGAATAGTCTTGGATCCCAACCAACGGAAGTAAAGAACTTC ATTCAACAACTCAACCGAAAGCTCTGCAGCTTTGAGAAAGAAAGTTCCAC TTCAGAAAGATAGTGTAACTCGGAGTACATGCTGCTGATAGCAAAA GTCAATAGTCAATGGACAATTTGGTACGATGATTCAGACGTTGCAACG TTTAGAACCAACGCCAAGCGGCTCTGAAAGCCAGGGCTACGAGCTCT TCTCGAAGCACCAACAAATCGAAACCCAGACTCCAGCCACAGAG AAGACATCAAGGAGAACAGTATGAAATGTTCGGAATCTTTGATGAT CGTTGCGACTCCAGAGGAGTGAAGCACCTATGGAACCTAGTCGACACT GAGCCTAGTGATCTGGAATCGCCAAAGTCAGACAGAGTAGTCTACGAA GACTGCATCAGACTCTATCTGATGTTTCGGAGCCAAAGTACTCCAG AAAAGCGGAAACTCCAGAGTGCCAGTGGTGGAAAGAAAGAGATCTG GATTGGCCAAACTGGAATCGCCAAAGTCAGACAGAGTAGTCTATGAA GACTGCATCAGAGTCTGCTGATGTTTCGGAGCAAGAAATGACTCCAG AAAAGTGGAAACCCAGAAAGCGCCATTGGTTGTAAGTGTGTTGAGTT GGAAGGCTACTAAGGATCTGCCAGTCACAGACAGGTCAACTGTCGT GGCAATCCCTGAAGCAGTAAACTGGAGGAAAGTCAGAAGTGGTAAT TCCACGGCTCATGGAGTTATCATACACCTCCCTCCAGAACCCCTCTCCAG TGGTTGAATACACCCAAACCATACACTCACACTCACACTAAACCAAGGTC AAAGCTACATGCCAGATGGAAAGAAAGAAAGGTAACCAACTGGGAA ACGAGACGAAGTATTGAGATTGTGAGACAATGGTTGAGAAAGAAATC AACGATTATGTTACGGAAGGACAAACTTCAACGACTGGAATGGCTTA CGAACTTACTACCGCCGCAATACAGAAGGCATCAGCTGGTATGAGG AAACAATCGAAAGATTGAAAGAGATGCCACCTCCAGAAAGTTAGAG	TTTCTCTTA TTTATCGA TTCTTGTG AAAAAT (SEQ ID NO: 1175)	TAGAAA ACAAGC CAAAAC ATCAAG CAGCAC GCAAAA AGGGGT AACTTTG GGCAAC TAATTA CGGATA CCTCGT GTATCA GGCAAA GCCGCC ACCAAC AGCAAA TACTGC CCGATA GGTAGG GCGTGA GAAAT GACCTA CAACCTC CAAGAC CCGAGC CCACGG AATCGA AAGACC TATAGG AAGTCA GTGAAT TGATGG AAATAC AAAAAC AAATTC TTCCATT CACAAG	AKQSKLHPFFTKAP HIDGVKPTVRRELSK MITPGGEHKGTKIPM VHTKRGLIQKINRAK KAKPMHDESTIEAS QLDVITDDDDDDN MTPMRRRFTWCL DHETTQEAWLTDV NWLKDLFCGNEQY MLVDPLVWLIYKMG GMAGVEQRFKSKT CLFPICEADHWILLVF DETNLVANSLSQSP NGQVKNFIQQLNRKL CSFEKEVPLQKDSVN CGVHVCLIAKSVNG QFWYDDSDVTRFT NAKAALKAQGYELFS EAPKQIENPDSSHRE DIKENSMECMCSL MIVATPQRSEAPMEL VDTEPSDLESKSDR VVYEDCITALSDVSEP RMTPEKSETPEVPV EERDLDPKLESPKS DRVVEDCITDLSVVS EQRMTEKCECTPEAP LVVECVLELRLPKDLP VTRSTVVAIPEAVKL EEKSEVWIPRLMELSY TVPPSPVVEYTOP YTHHTKPKVKATCQ MGKKRKVPTGKDE LIQVIRQWFEKEFND YVTEGRNFRLEWLT NLLTAAIQKASAGDE ETIEKIRKRCPPPEVRE NEMSTQTSQRQKPT

AAACGAAATGTCCACTCAGACATCTCAACGTCAAAAAGCCTACCAAC	GACTTA	TTNQKRSRNTTQSD
GAATCAGAAAGAACGCTCTAGAAACACTACTCAATCGGATACACAAGCC	CTGGTC	TQANTYWRNRAKTY
AACACATACTGGCGAAATCGAGCAAGACATATAATCAATCATAGGTC	GAGTAG	NQIIGQDFKQCDIPIA
AAGATTTCAACACAGTGTACATACCGATCGGATACAGAAATCTA	AGCACA	ILEEFYKTTSVTNVP
TAAAAAGACTACCTCAGTGACCAATGTCCCTCAGGAAACCTTGTGAAA	AGCCAA	QETLVKTSRLPRLDI
GTCACTCAAGACTACCAAGTTAGACATTGGAAAGTGATCGAGGATC	AATATC	GKWIEDPTEQEVFG
CGTTACGGAAACAAGAGTATTTGGTGCCTCAAAAAGACAAAAGACA	AAGTAT	ALKTKDTAPGTDGL
CTGCGCAGGAACAGATGGGCTCAGATACATACCTCCAATGGTTTGA	GACGCA	RYHLQWFDPDCKMI
TCCCGACTGTAAATGTTGAGTAGCATTTACAATGAATGCCAGACCAT	AAAATG	LSSYNECQHHLKIPAI
CTGAAAAATCTGCCCAATGAAAGAAAGCTGAAACAAATCTCCTCTTCAA	GGTAAC	QWKEAETILLFKSGD
AAGTGGCGAGCAATCCAAACAGACAACTGGCGGCTATAAGTCTCATG	CTTGGG	ESKPDNWRPISLMPT
CCACCATCTACAAGCTATACTCAAGTCTCTGGAATAGGAGAAATACGGA	CATCCA	YKLYSSLWNRRTIV
CGGTGAAGGGGATTATGAGCAAGTGCCAACGAGGATTCCAAAGAGAGA	ATCAAC	KGIMSKCQRFQERE
GAAGGTTGCAATGAGAGTATCGGAATACTCGGAGTGCTATTGATGTG	GGATAC	GCNESIGILRSAIDVA
GCTAAAGGAAAGATCCACCTGTCCGTTGCATGGCTGGACCTCACCA	CTCTGC	KGKRSHLSVAWLDLT
ATGCCTTCGGTTCCGGAGATGTTGTACACATTGTCAAGGACATGTATAAA	GTATCA	NAFGSVPHIELSTLS
GACGTTCCATAAGAGTCAAGAAATAGAACGGAGAAAGTAGAGCAGATT	GGCAAA	AYGFPEMVVHVIVKD
ATGATAAAATCTGGGTAAACAAAGGCGACCTCTATCTCACCAACACTAT	GTCGCC	MYKDASIRVKNRTEK
TCAACATGTGCTCGAAACGGTGATTAGACGACATCTGAAAGAAATCATC	ACCAAA	SEQIMIKSGVKQGD
AGGTACAAATGCATTGACACCAAGATCAAGCTTCTTGCAATTTGCAGAT	CTGTACT	ISPTLFNMCLETVIRR
GATATGCCGTTCTAGCAGAAATCAAAAGAGCAGCTACAAAAGGAGCTT	ACTCCG	HLKESGHKCIDTRIK
ACAGAAATGGATGAAGACTGTACACCTCTCAACCTAATTTTCAAGCCGG	AAAAAA	LLAFADDMAVLAESK
CGAAGTGTCAAGTCTCATATAGAGTTCGGGAAAGTGAGGACCCCATG	CCAAGA	EQLQKELTEMEDDCT
AGCAGATCATGTTGAAGCGAGAGCCGATCCGAAACCTCAATGATGACG	AACATG	PLNLFIPAKCASLIEF
GAACATACAAGTATCTGGGAGTGATACGGGAGCAGATGCAAGGACAT	ATTTTC	GKVRTHEQIMLKREP
CAGAAGAGGAGCTGATCTTCTGTAAACAAAGAGGTAGACCTTGTC	CATCC	IRNLNDDGTYYKLV
TCGTCGGCGCTTACGCCACCCAGAACTGGACTGCTTAAAGAGCTTC	GTTAAA	HTGADARTSEELIIS
ACACTCCAAAGATGACTACATGTATGCCAACGCCATACCAAACTTAC	GCATCTC	VTKEVDLVNRSALTP
CGAACTTTCAGCGTTGCTAACATGTCATGCGAGGAGTCAAGATAATC	AACCAA	PQKLDCLKTFTLPKM
CACTATATCCAGTTAGAGGATCTCTCTTGAATATATTTCAAAATTCGAC	GCTAAA	TYMYANAIPKLTELSA
CGGCAAAAGGAGGACTTGAGTTCATGCCCTAGAAATCACGGCATTGATT	GCGGTA	FANMVMVRGVKIIHYI
ACCTTCCTTGTCTCAACCATGAAGAAACTGTGGTCTGATGATGAATACAT	AGGTTA	PVRGSPLEYIQIPTGK
TCGTAAGCTCTACAACTCTTATCTGAAGAAAGTTGTGGAGGCGGAAACG	TCATGTC	GGLGVPCPRITALITF
GGAATAGTGGAGGTCTCCACAAAGGATCTAGCAGAGTACCTCAGCAAC	AAAAGG	LVSTMKKLWSDDEVI
AAGGTACCATCCAGAAAGCAGCAATTCGGGTATAACTGCTACTCGAGGA	TGTAGC	RKLYNSYLKKVVEAET
TTCCGGAAGTTTGTAAATGGCTAGCTCTCAACCAAGCTGCCCTCTCTAC	TACAGC	GIVEVSTKDLAEYLSN
AAACTTGAAATTCATCGAAACAAGACAATGAGTTAGCAGTTGTTGCCAGC	AACCTA	KVPSRKHEFGYNCYS
	AAGCCC	RIREVCNGLALNQAA

				CGACTGAGGAGAGCAAGGAAAGGATTTTCACTAAAGATCATGTGAAAA AGTCCAGTCGCTACTGAAAGCCAGCGTGAATGACGCACTGTACACAG ATTCTTGACAAACAAACCCGTCAAAAGTGAAAGTGGTACAAAGTTCTCCAG CAGCACCTCAAAGCAACACAGCTTCGTCGAAATGGAGGTAAAGTAAGT ATATCGGTACATGATGATCCACAGGTACAGGTACCGTTAAACCAACTAACGT GCAATTATAACATCTTTGATCCAAAGCAACCGAAAACTGCCGAGGTG TGTTATAAGAACGAGACTCAATGGACATCTCTGCAAGACTGCACATAT GGCTGGGTAAACTTATACGAGAAAGACACGATGCCGTACATCACAAAG GTAGTCACAATGATTTGCGTGGGCAAGAAAGAACTGGGGCCGGAAA ATCGACCAAGAACTGCCGGTTTCACTTCACTCCGTCCAGACATTTGTCT GACGAGTCCGGATGGCAAGAGGTTATCTTTGCGGATGTTGTGTCCT TACTCAAGGACAAGGAACATCGAATTCGCGTGGAAGAGAAAAATCCGA AAGTATACAGAAAGGATACAGTCATCTTGTGCAAGAGGAATCAAAGTGA CAGTCCTCCGATAGCCATAGGATCACTCGGAACCTTGGTGACGCCAAC CAACGAAAGTCTCTATCAACTGGGTATCAGCAAGAGCGATATTCGCACT GCCATTCCATTACTATGCTCTACTGTGATGGAGTATAGTAAGAACGCCTA CTGGAATCACATATACGGAACCTATACCTCGGTCCCACTGAGATAC GGACACCAGAAAGCCGATGGAGACGATTGGAAAGAAAGAACTGAGTTGC GAACCAAGTTAGTCTCAACAAATAACATGCCTTGGAAAGGACCAACGC CAAAAGTCTGGCAACTGATTGAATAATGTATAAAAGTAACGTGAACC AAATGCCCGATAGGTAGGGCGGAGAAAAATGACCTAGAAAAACACAAA GTCCCAAGCCCCGGATTGCAAGACCTATAGGAAAGTCAGTGAATAGA GAGAAATATCAACAAATCTCACCATTCAAGGACTTACTGGTGCAG TAGAAAAACAGCCAAACATCAAGCACGACGCAAAAAGGGGTAACTTT GGCAACTAATTACGGATACCTCCGTGTATCAGGCAAGCCGCCACCA ACAGCAAACTACTGCCGATAGGTAGGCGTGAGAAATGACCTACAA CCTCAAGACCCGAGCCCGGAATCGAAAGACCTATAGGAAAGTCAGT GAATTGATGGAATACAAACCAAAATTTCTTCCATTCAAGGACTTACT GGTCAGTAGAGCAAGCCAAAATATCAAGTATGACGCAAAAATGGG TAACCTGGGCATCCAATCAACGGATACCTCTGCTATCAGGCAAAATGGC GCCACCAAACTGTACTCTCGAAAAAACCAAGAAACATGATTTTCCCAC TCCGTTAAAGCATCTCAACCAAGCTAAAGCGGTAAAGGTTATCATGTCAA AAGGTGTAGTACAGCAACCTAAAGCCGAAAGGTAGGGCCGTATAAA AAGACCTACACCCCTCAAGACCTAAACCCACGAACTCGAACGACCTACA GGAAGTCGTTGAATGGAGAGAAATATCTCACAAATCTCTTCCATTAC AAAGGCTAACTGGTCAAGTAGAGCAAGCTAAGCCTCCCAAGCACGAA GTGATATGGGTAAATTTAGGCAACCAATCAACGGATACCTCCGTGTATCA GGCAAGTCCGCACAAACACTGTACTCTCCGTTACTCCCAAAACACATG	PLYLEFIEQDNELAV VVQPTESKERIFTKD HVKKLQSLLKASVND ALLHRFLTTKPKVSEV VQVLQHPQNSFV RMGGKVSISVHWI HRSRLNQLTCNYNIF DPQPKNCRRCGYK NETQWHLQDCTYG WAKLIRERHDAVHH KVVTMICAGAKKNW GRKIDQELPGFTSLRP DICTSPDGKEVIFAD VCVPYSRTRNIEFAW KEKIRKYTEGYSHLVA QGKIKVTVLPPIAGSLG TWWTPTNESLYQLGI SKSDIRSAIPLLCSTV MEYSKNAYWNHIYG NSYTSVPLRYGHQKP DGDDWKKELSCEPV LALQQ (SEQ ID NO: 1420)	GAAGG TAGGC CGTATA AAAAGA CCTACAC CCTCAA GACCTA AACCCA CGAACT CGAACG ACCTAC AGGAAG TCCGTG AATGGA GAGAAA TATCTCA CCAAAT CTCTCC ATTACA AAGGT AACTGG TCAAAT AGAGCA CAAGCT AAGCCT CCAAGC ACGAAG TGATAT GGGTAA TTTAGG CAACCA ATCAAC GGATAC CTCCGT GTATCA GGCAAA GTCGCC ACAAAC

	GATCTCTTCTCTACCAAAAAGCTTTATAACCAAGCTAACGGTGGAAG GACATCATGTACGAGGAGTAGCTACAGTAACCTCTCTCTTGAGACTGC AAAGTCGAGGATGGATTGGGAAGCCGCGAGGCAAAAGCGGGTAAC TCGGCCAGACGCTAGTGATCTCGGATCCGACAGCCCTGGCCTTAGAGG AACCTGGGATAAGGAGCACGACGGGAAGGATGTTCCGCAAGGATTTC CCTTCCATTAGTCAGGGCTGGCAGTTGGTAATATAGCCTTTCTACACAC CACGCTTGCACCCACTAACCAAGTGGATATGCGGGTGGACTCAATG TAGAAAGGTGTTCCCACTGCTGACTCGCAACTTTATATGCTCTTGCAA CATAATGGCCCTCACTATAAACCCTAGCAACTGGTGCTCCGGCGAA GCCGGTTCTTGCCACTATTGCGCCCCAGGCTCGCC (SEQ ID NO: 1052)		ACTGTA CTACTCC GTTACTC CCAAAC ACATGG ATCTCT TCTCTCA CCAAA AGCTTTA TAACCA AGCTAA CGGTGG AAAGGA CATCAT GTCACG AGGAGT AGCTAC AGTAAC CTCTCTC TTGAGA CTGCAA AGTCGA GGATGG ATTGGG AAGGCC GCGAGG CAAAAG GCGGGT AACTCG GCCAGA CGTAG TGATCTT CGGATC CGACAG CCCTGG CCTTAG AGGAAC CCTGGG

249

NeS L	NeSL- 4_SM	.	Schmidt ea mediter ranea	TTAAATCATTTTAAATGTGTTTGAATATCTTAAATATCAAAATCATATTA ATATCAATGCTAAAAAATCGTGCKCATCAGGCGCACGAAAAATATG GACAACTCGTCGACCTGTGCACTCAGAGAACCTCAATTTGGA AGAATGGGAAGCCTATAATGCTACAATCCGCCAACCCCTATTGAAATG ACAGATAGTCAAATATCAAAAAATATACAACTGCTGTCAAGCGTGACT CACTTCTTCCAATCGAAAAATAGGAATGTAAGAAACATGAAAGTCA AGCTGAAAAACCAATAATATGCTCTAAAAATAAAACAATTTGAAAAATG CAAAAAATACCTATAAAATCACAGCCGAATAAATCCCATCCGTTCTAAG CAGAAACCGCTACGAACTACTGCAAGAAATCGGATCAAGTATATTAATT CCCCCMGGGGGAAATTAATATACTTGTTAKAAAAATTAATTTTAAATAA AAATAAATAAATCGAATAAATATAAAAAATAAAATAAATCAAAATAAACTT TTATTAACCAATAAAATCGCAGTAAGTAAATTTCCACTGTTATTAATTTA AAACAAAAATCCTTTAAAAATGCCTCTCTTTTCAGTAATAACACCTTTTC TTGCTTTTATTAATAATTTCTTGTAAGTACTGTACAAATCGAGCACAGTTAATG CAAAATAGGACATAGAAAATTCCTTTTAAAGTAAATTTAAATCCATGAGAAA TAAAAATAAAATCCTTTTGATTCAAAGTTCTATGTGCTTCTAATAGAAAT GGTGAAGCATTAAATGGGCTTGATTTTATAAAATTAATATATTTAAATC TATTAATAATTAATGTTTTTATAATTTAATTTTATAGTGGGGGAA TTAATATACCTGATCCCAAGAAATCAACTGATGATGAAGAAATGTTATTT CAAAATACATACAAGAAGCTGGAATAAAACAAATCAATCGCTACAATGAA TGTGGATCTCGATGCAACAAATTAAGATTTGGAATGAACACAAAAAGAG ACGACCTATCCAAATTCACAACTCGAGTTGAGACGACTCCCTGTACCTC AACGACTATTATGCATGCATCTTGCAACACACCGACACTATATCTTACT CTCCATTACCATCGGCTGTGCACTTCCGAAAGCCCTGCCTCGTCAATC ACAATAACCAACACAGACGATAATGCGATATTATAGAGACCCCTTACC CATTACCTCAAAACAAATGGTGACTTGAGTGAATATTAAAGGATATAGA AGCTAATAAGGACACCAACCATGTCGAATAAAGTATTGGACTGTGACTCT GACAGCGGCGATGATCGGGACATGATAATAGAAAAATGACCGAGAATCT GACATGGACCTGTTTCGCAATCTTTATTGAACACTAATCAATCTGATGA GAGGAGGGAGAAAAAACTTAACAGAAAAATGCTCCAACAGAGATTACTAC				TTAAATCA TTTTTAAAT GTGTTGA ATATCTTA AATTATCA AATCATAT TAATATCA ATGCTAAA AAAAAATC GTGCKCAT CAGGCGCA CGAAAAATA ATGGACAC AACTCGTC GACCTGCT GTCGACTC ACAGAGA ACCTCAAT TTGGAAGA ATGGGA GCCTATAA TGCTACAA TTCCGCCA ACCCCTAT TTGAATGA CAGATAGT CAAAATC AAAAAATA TACAACT GCTGTCAA		GTCTTG CCACTAT TGCGCC CCAGGC TCGCC (SEQ ID NO: 1298)	MNVLDLDAIKSIGM NTKETTPNSQLRVE TTPCTSTTMMHASCN TTSTISYPLPSAVSLP ESPASSITITTTDDNC DIETPYPLPQTNGDL SELKIDIEANKDITMS NKVLDSDSDGDDR DMIENDRESMDLFI SQSLNTNQSDERRE KNLTENAPTEITTEKS YFDIISKASDNTTSKLL LNVKNELTAGLPMP PVTNTAKFIRNVPE DIADPTLYRLDSRGKL GCRTQYKKPGCGDIA VVDYEAIVEHAAFIHT IPFNEQNNVDCQPC HPKKGDVHTVILIKY ADIFNHIEAHSHVVQ TAITDNMKTYLRLTKE NXFYCSYRNKKKNK CKKAFNLESNMMDI TEHMKHTGYSFDX NLNLCYCGIWKPFTE LIAHIKTHELQEVNSI PNKENIHNTTIVSPL NFAGILASGETQNI DEEIKPRDLPENLAF NRNIENELSWSQHLV

TGAGAAAGAGCTACTTTGATATCATCAGTAAAGCATCTGATAATACAACTT	GCCTGACT	KAYIFSVAVKTSTIFIIN
CTAAGAAAGCTGCTTAATGTAAAAACGAATTGACTGCTGGACTACCTCT	CACCTCCT	PYTCNALIQCNKYKTF
ATGCTCCAGTGACCAATAGTCAAAATTCATTGAAATGTTGACCTGA	TCCAATCG	ETFPKDFAKWNEIV
GGATATTGCAGATCTACCTATATCGACTTGACAGCAGGGGAAAGCTT	AAAAATAG	LPHNNTSSWSFFFL
GGATGCAAAACWCAATACAAAAACCCGGATGCGGGGACATAGCAGT	GAAKATGT	NKKRVAMIIDPTAD
ATATGACTATGAGCGGATAGTTGAACATGCCGATTATCCACACAATC	AAGAAACA	DSHTLHFELATDILRTI
CCATTTAATGAACAAAAATATGTGGATTGTCAACCATGCCACCCTAAAA	TGAAAGTC	LNVQIFEDLNFPLTE
AGGAAAAGATGTCCATACAAATAGTTCTGATAAAATATGCAGATATCTTT	AAGCTGAA	VEYPVCHEANLSAFX
AACCATATTGAAGCCCATAGCCAGCTGTGTGCAAAACCGGATTACAGATA	AAACCAAT	VCHFLKCLMSDLPIDI
ACATGAAAACCTATCTACGTTTAAACAAAGGAAAAATTTKTCTACTGCTCA	AATATGTC	PDIDHMKETMRPIIR
TATCGTAACAACAAAAAAGAAATAAATGCAAAAGGCTTTTAAACCTTG	CTAAAAATA	KYNCAKFPESDVRNY
AATCAAAACATGATGGACATAACAGAGCACATGAAAACTCATACCCGGATA	AAAACAAT	RVLIEDLIYQLNLDIT
CAGTTTCGACMAAAACCTTAAACATTTCTATGCTATTGTTGTTATCTGGAAG	TGAAAAATA	CEEILCEIERINGRLNP
CCGTTACAGAGCTCATTTGCCACATCAAGACTGAGCAATTTGCAAGAAT	TGCAAAAA	KRYFKESKPKTDIHL
ATATTAACCTCAATACCAAAACAAAGAAAAATATCCATAATACTACTACCATA	ATACCTAT	QKKSAELLCKVKRLKF
GTTTCCCTCTAAACCTTTGCTGGGATCTTGCACTGCGGAAACTCAAAA	AAAAATCAC	QISQKTEIGKIWEND
TATCCCGGATGAAGAAATAAATAAACCCAGAGATCTGCCAGAAAAATCTT	AGCCGAAT	DVDHRPPMARFLKT
GCCTTCAACCCGAAACATCGAAAAATGAATTAAGTTGATGGTCGAGCACT	AAATTCCT	FASQDCPVSNSTSSINL
TGGTCAAGCATATATCTTCTCTATGCTGTTAAACATCAACCATCTTCA	ATCCGTTT	PYYMDTDTDXCTDC
TCAATCCTTATCTTCAATGCTTTGATCCAGTGCAACTACAAAACTTTCT	TAAAGCAGA	ENLSHIMKNLDSSAP
TTGAAACCTTCCCTTCAAAGACTTTGCCAAGTGGAAAGAGATAGTCCTG	AACCGCTA	GMDLITGGDWKKISP
CCAATTCAACAACACACTTCTTGGTCTTCTTCTTGAACAAGAAA	CGAACTAC	KHELITAICNLCILNKV
AAACGAGTTGCGATGATTATCGATCCAACCTGCAGATGACAGTCATACCC	TGCAAGAA	CPEKWKLFRVTLILKP
TGCACTTTGAATGGCTACAGATATCTAAGGACTATATACTTAACGTCCAG	TCGGATCA	GKMSSEFRANSWRP
AATATATTTGAGGACTTAAATTTCCCTCTTACTGAGGTGCAATACCCCGT	AGTATATT	LAIMDTAYRIFTLLN
GTGTCATGAGGCAACCTTTCCGCAATTTMTGTATGCCACTTTCTTAAAT	AATTTCCC	NRLLQWIRNGNLISP
GTTTAAATGTCGGACTTGCCAAATGATATCCGGATATCGATCAGATGAAA	CCCMGGG	NQKAIGIPDGCAEHN
GAGACWATGAGACCAATTATTAGAAAAATAAATACTGCGCAAAGTTTCCG	GGAAATTA	ATLHFAIDRAKCKTE
GAGAGTGATGTTAGGAATACCGGCTACTTATCGAGGACCTGATATACC	ATATACCT	LHIVWLDIADXFGLSP
AATTGAACCTTGACACAATTACTTGTGAGGAAATACTGTGCGAAATCGA	GTTAKAAA	HDLIWTYLANMGLK
AAGAATAAATGGAAGTTAAATCCCAAAAGATATTTTAAAGAGAGATAA	ATTAATTTT	NETLTLIKELYKDVKTI
ACCAAGACGGATATAATACATCTGCAAAAGAAAAAGTCGCGGAACT	TTAATAAA	FDCQGTLPSEPVPIKGG
CCCTGTGTTAAAAAGATTGAAATTCCAAATCAGTCAGAAAAACAGAAATC	AATAAAATA	VKQGCPLSMVTLFCLSI
GGAAGATATGGGAAAAACGACGATGTGGATCAGACACCGCCTATGGCC	AATCGAAT	DYILKSILTNYPFLHHD
AGATTCTTGAAGACTTTGCGGAGTCAAGACTGCCCGTTTCGGAATACGTC	AAATATAA	LNISILAYADDLVLLSD
ATCCATAAACCTACCTTACTACATGGATCTGATACAGATAMGTGTACT	AATAAAAA	SYLEIKKSLESTVELAA
GATTGTGAAAAATTTGTGCGACATCATGAAGAACTTGGATAGCTCGGCAC	TAAATCAA	FANLKFPSKSGYLSI
CTGGAATGGACCTCATTACAGGTGGAGACTGGAAAAAGATCTCCCCCGA	ATTAACCT	NNVNSDILKLHLYNE

AGCATGAACTGATAACAGCAATCTGCAATTGTATATACTACGAAATAAGGT CTGCCAGAGAAATGGAAGCTGTTAGAACAGCTTTTATCTCTAAACCA GGAAAAATGTCGAGAGTTTCAGAGCTAACTCATGGAGACCTCTTGCAA TCATGGACACAGCCTATAGAACTTTACGACTCTGCTGAAATAACCGCCTG CTGCAATGGATCAGGAATGGCACTCATAAGCCCGAACCAAAAGCG ATTGGTATACCGGATGGATGCTGAGCATAATGCTACTCTACACTTCG CAATTGACCGAGCTAAACGATGTAACACTGAACACACATTGTTGGCT CGATATCGCCGATKCAATTGGTTCGCTGCCTCATGACTGATCTGGTATA CACTGGCTAATATGGGTCTGAAGAACTGAACACTAACTGATTAAGGA ACTATATAAGGATGTAAGACTATCTGACTGTCAGGGAACTTTGTCC GAACCTGTCCCAATTACTAAAGGAGTTAAACAAGGTGCCCCATTATCAAT GACACTCTCTGCTGCTATTGACTACATCTTAAAGTCAATACTACTGACTA ATTATCCCTTCCTTCTCATGATCTGAACATCAGTATTTTGGCATATGCTG ATGACTTGGTCTCTTCTGACTCTTATCTAGAAATCAAAAAATCTTTAG AGAGTACTGTGGAAATGGCAGCMITTTGCCAACCTTAAGTTTAAACCTTC GAAGTCTGGATACCTTGCCATCAACAATGTTAACTCCGATATCTTAAAT TACATCTCTATAATGAGGAGATACCAACGATATCCGAGAAATAACAAATA CAGATATCTGGAGTTGACTCTCTTACAAAAGAAATCAGGATGTTGAT GGACGACTTGGGTCTGCACTTGCACACCAAGATCTCTATTAATCATA CTTGCAATCCGGCGCAAAAGCTGAATGCTTACAAAACCTTCATCCACTCCA AGCTTATCTTCTCTCGTAAATGCGTGATAGGTGATAGAAATCTCGAC TGATGCGGAATAGAGTTACACAAGGTCGGGAAAAACAGCTGGGCTTT GATCAGGAAATCAAGGCACTWCTGAAACCATGATTGGAGACAAATTT CAGGCAKTAATAACTACTTCTTACACTCACTGCAAGCTGGGGGAC TTGGTAAACCTCAGCTATTGATGAATATTTGATCCAAAGCATTACCGGA ATAACAAGATTATTTCACTCATCCAACTCAGCTTCAGAAAAATGCTAAT CACAGAACTCGCTATTCTAGAGGGGAAAAAATTTGAAGCGGGCT AAAATGGCTTAACGTGAAGTTAACAGGCAATCCCCAACACCTCTTTCT TTGAAAAATCCAAAAATCGCACTTGTCTTAAAGAAAGTTCTGTATA TGCGTTAACCTTAAATTTGATAGAGGACAAATTTCTCACTTGAGATGACCTA CAAAAAGCGCACTTCTATGTAAACCATCAAAAACCTCAGCACACTTTCCA AAGAACTCCAGCACTTCTGTTGGCTTTACTATGCGATGCGGGATAGCATAA AAATGAGAGTACAAGGACACATTTGCGACTGCGATGCGGGATAGCATAA CAGCTAAAATACCTAATAGCTAGTGACATCCTTAACGACGACAGTACTAC TTCTTGGTACGTGCGAGAAATAATCTTCTGAATCTTAACTACAATGCGTA TCGACTCAAGTATAATATTGGCACAAGTGACAGCTTTGCCACCTTGAT GAAGAACTCAGGGCCCATSTGTTCAATCACTGCGGTGCCAAACCAACG CTAGAAAGAGTGAAACACGAAAAATGTGCTAGTAAGCATAGTTGCCCTTCT	TTTATTAA CAATAAAA TCGCAGTA AGTAAATT TCCACTGT TATTAAT TTAAACA AAATTCCT TTAAAAAT GCCTCTCT TTTTAGT AATAACAC CTTTCTTG CTTTTATT CTATTTCT GTGTACTG TACAAATC GAGCACA GTTATTGC AAATAGGA CATAGAAA TTCCTTTT AAGTAAAT TTAAATCC ATGAGAAA TAAATAA AATCCTTT TGATTCAA AGTTCTA TGTTGCTT TCTAATAG AATGGTGT AAGCATTA ATGGGTCT TGATTTTT ATAAATTA AATATATT TAATCTAT	EIPTISENNKYRYLGV DFSYKRNQDVGRL GSALALTRSLFKSYLH PAQLNAYKTFIHSKL IFSLRNCVIGHRILDC DRNRVTQGREKQLG FDQEKALLKTMIGD KFQAXNNYPYTHCK LGGLGITSIDEYLIQS ITGITRLFHSSNLSFRK MLITELAHSRGGKNF EAGLKWLNCEVNKA FPNTSFFVKFQKSALA LKRKFCICVNLKFVED NFSLEMTYKKRTSVW NHQNLSTLSKELHDF VGLYYAEQXCQMRV QGHIAITAIGDSITAKY LIASDILNDAQWYFLV RARNLLNLNYNAYR LKYNIGTKRCLHLDE ETQAHXFNHCRAP NARRVKHENVLSIV AFLEKIGFEIDVEKSPK YISPTKLKPDMMVIRSK RNKDIHVLDLKVYPD SGEGFEKAREDNVVK YKDLAIEIGKAFNQKA TISAVVIGCLGTWVK KNNAAASKIGLTKTEII SLARIACPNVIAIYH IYREHVSFTKSAMAL PFSLA (SEQ ID NO: 1421)

NeS L	R5	AY216 701	Girardia tigrina	AGAGAAAAATTGGATTGAGATTGATGTGGAAAAATCACCCAAATATATC TCAATACCAACAAAGCTGAAACCTGACATGGTAATTAGGCTAAGAGGA ATAAAGATATACATGCTAGACCTAAAAGTGCCATATGACTCAGGAGA AGGCTTGAAAAAGCGGGAAGACAACATATGTTAAATACAAAGATCT AGCCATTGAAATTGGAAAGGCAATTAATCAAAAAGCGACTATATCTGCT GTGGTATTGGATGCCTGGGCACATGGACAAGAAAGAACAAATGCCGCT CTTCCAAAATCGGGTTGACTAAGACCGAGATCATATCTCTGCCAGGAT AGCATGCCAAAATGCGGTAATCGCATGCTATCACATATACCGTGAGCAC GTCTCATTTACAAAAGAGTGCCATGGCCCTCCCTTCAGCCTTGATGAGT GTGCTACGAGGACGCTGGTAATTGCATCGGCGTTGCAGATTTGTGTA CGATAGATAAAAAACCAATAGTAATAAATGCTGAGCCTAGCTCGCATATC TAAGCCGAAAGGCAGCATATATAGACAAATTTAAAAAATAA (SEQ ID NO: 1053)	TAAATTAA TATGTTTT ATTAATTA TTAATTTT ATAGTGGG GGGAAATT AATATACT TGATCCCA AGAAATCAA CTGATGAT GAAGAATA TGTTATTT CAAAATAC ATACAAGA AGCTGGAA AAAAACAAA TCAATCGC TACA (SEQ ID NO: 1176)	TGATCC GTGTGT TTGTGTC GTATGA TTGTTTC CGTGTG TGTCTAT ATTTTC TTTTTA TACTTC AATTACC TCGTTGT AATGTT ATAACTT CATATG GAATAT ATGTAA TTTAGTT	TTGRNLGQWSCYSR SIQQSNYSFKLSSTEV GELVEQSPAPLQSPQ FSNNYNLNNNNLY YSLNTFNQSNLCL VNIEFFPTQHLLGDIV NSGCINYMNNYNF DNINLYNSVLSYNN YNHSLASPYTTNITE HADINMHVQEVNM QQDNNTQHAITQQV SLQATSLQHTLDEMI VQFNTAVRLKKKHV AKIFRGNHRKDLPT LPAREQYKTKPLAIR EVLHRKTTATSSPEN AIKAFFSSYSRPAELFT GOELLESWFPVHPE
				GTAGGTAACTATGACTGCAAAAATAATTTCTACACCTATTGTTGATAAC TCATCTCGTGGCAACGAGCATGTTATTTCTAATCATTTCTGCACACA GGATCTCTAATCTGATAGTAATATTATAGATAGAGATAGGAACCTTG TTGATTTAGATGCGTCAATACTTCTCTACTATTATACGCCAGAGGAT AGTAAGATATCTGAGGATGAGGACTTCACTTAGTCAATAGGAAAAAGA GCAAAAATAAGAAAAAATCTAAGAAAAACAACCTGAAAAATAAAATGAAA TTCCTATTCAAAAGAGTAAGATAAGAAAAAGAAAGTCTAAAAATTAATAC CGAAAAACTAACTGAAAAATATTACTACTTCTGAAATACCACTTGAAATTG CTCCTTCCATACCTTTAGCTTACGCAAGTACCTCGGGTTCTCAACAACCG GCCAATCTCCAGAAGACGCTACTCTAAGTGATACGGATCTCTTCTTAC ACAGGATGATCCCGATAGTCTTATCTTCTGGAAGTACTCAACCAACCT TTGTTGACCTCAACCTTTCACAGCAATCGGAACCTTCTTCAAACTAGAC AGCCAAAAGATTTGAGCGGGTGAACACCAACCAAAATCATACTCTTACA GGGATGACCTTTTCTACTCTACAGTCTTCACTACAACTCAGATACAGGT TACGGTATAAGTTTGACAAATGGGAGCAGAGGTTTCGAATTTCTTGCTA GGAATCTTGTCAGGAAAAACCAAGGATAAGTTCCCTCTTTATATGCTGG ACAAGTAAATTAGACACACAGTCTTCTTCAATCACTTCAACCAAGGCATCT ACGCCAATAATATACTGATAGTAAAGGTAATCTAATTGAGTTTCTGAT	GTAGGTAA CTATGACT GCAAAATA ATAATTCT ACACCTAT TGTTGATA ACTCATCT CGTGCGCA AACGGAG CATGTTAT TTCTAATC ATTTCTGC ACACAGGA TTCTTCTA ATTCTGAT AGTAATAT TATAGATA GAGATAG		

GATAAGCCTTTTCAAAAGTATACCGACTGACCCCAAAACTGAACCTAGAGC	GAACCTTG	TAGTTTA	DDFEFRIPGRDQJAKY
AAATTAGGAGAGAGACAACATCTAGTTGATAGAGCTCTTAGACATAA	TTGATTTA	GTTAGTT	IKFASKSAAGLDWITY
TCAGTTACGGGAACTTATATTTTAAATAAACTTAATAATAATATG	GATCGCTC	TAGTTTA	EDIKLGDPSEILQPIF
GGGTGGCGAACATTTGAAAAGGAAAAAGATCAAAAGTCAATACGGATG	AATAACTT	GTTTAGT	EYIVQNNICPSEKAS
ATGTCTCAGCAATGATGGAGACAGAAAAATAGACGACAGCAAGAAAA	CTCCTACT	TTAGTTT	RTMIPKPGKSDYSDP
CCTCGGACAAATGGAGTCTATCCGCTCAATCCAACAATCAAAATTACT	ATTATACA	AGTTAG	SSWRPITTSAYRLL
CCTTTAACTGAGTTCGACTGAAGTGGGAGAGCTGTTGAGCAATCTCC	GCCAGAG	TTTAGTT	MKYLWELYNWILL
AGTCCCTTCAGTCGCCTCAGTTCTTAACAAATTATAACAATCTAAATAT	GATAGTAA	AGTTTA	NQMLRSQKSLGKFE
CAACAACAACCTTATACTATAGTCTCAATACCTTTTAACTAGCTAATAACCT	GATATCTG	GTTAGT	GCHDHAMLMNLI
TTGCTGCTTGTAAATTTGAGTTTTCCTCAACTCAACCTTCTTGTTGA	AGGATGA	(SEQ ID	QDVRRQTNPSNPIN
TATAGTTAACTCGGGTGCATAAACTATATGAATAATTATAATAACTTTG	GGACTTCA	NO:	KNKRLYVFLDFTNAF
ATAATATTAATTTATATTAATAGTAATGTATTATCTTACAATAATTATA	TCTTAGTC	1300)	GSVPLDTLMVVPQRF
ATCACAGTTTTCTCGCTTCCCATATACTACTAACTACACAGAACATGCA	AATAGGAA		GLGTSALTLIKNLILD
GACATAAACATGCACGTGCAAGAAAGTTAACATGCAGCAAGATAACAAT	AAAGAGC		NYTNVTCGESKIENV
ACACAACATGCTATAACACAACAAGTCTCTACAAGCAACATCTCTGCA	AAAAATAA		KLKNGVKQGCPLSM
ACACACGTTGGACGAAATGATAGTCCAGTTTAAACACTGCTGTCAAGTTA	GAAAAAAT		LLFNIFINIIIRAIEAMP
AAGAAAAAGCACAAAGTTGCAAAATCTTTAGGGGACATAATCATCGTA	CTAAGAAA		DVHGYPLGDMDIRIL
AAGACCTTCCAACATTGCTGCTAGGGAACAGTATAAACTAAACCGAA	ACAACTGA		AYADDIALISDSHKDL
ACTTGCAATTAGAGAGGTAATTCATCGAAAAACAACAGCTACGTCTTCCC	AAATAAAA		QEMVYKAEYIGRILGL
CTTCTGAAAATGCAATTAAGGCTTTTCTCTCTCTACAGCCGTCAGCT	ATGAAAT		LFNPSKCALMDIPHD
GAACTTTTCACTGGTCAGGAACCTCTTGAATCATCTTGGTCCAGTACA	CCTATTCA		KKRTPPILVNGEMIKC
CCCGGAAGATGACTTTGAGTTTGAATTCGGGTAGAGACCAAAATAGC	AAAGAGTA		VGKADPYKVLGTFRS
GAAATACATCAAGTTTGTAGTAAATCAGCTGCTGTTGACTGGATC	AAGATAAG		WFRKLDIKELLQMM
ACGTACGAGGATATTAAGTTAGCGATCCGTCGGGGGAAATTCCTCAAC	AAAAAGA		MDETKLITESNLPH
CCATTTTGAATATATAGTACAAAATAACATATGCCCATCCGAGGGGAA	AGTCTAAA		QKIHAYETFIHSQLPF
GGTAGTAGGACCATATGATTTCCAAACCGGGGAAAAAGTGACTATTCA	ATTAATAC		HLRHSRIPFSDFITNR
GATCCTCTTCTGGCGGCCATTACAAATTACCAGCGCAGTATACAGACT	CGAAAAAC		KTNKTTNNSNDSEKS
TCTCATGAAATATCTTACATGGGAGCTGTATAAATCGGATTTCTTTAATC	TAACTGAA		IQKAYDPESGQLFLN
AGATGCTGTCCAGGAGCCAAAGAGTTTAGGGAAGTTTGAGGGATGTC	AATATTAC		TFALPSGCAKDFYIT
ATGATCACAAACGCAATGTTGAACATGCTCATCCAAGACGTTAGGAGACA	TACTTCTG		KDAGGPQLTSGLDEY
GACCAACCCGCTCTAATCCAATCAATAAGAAATAAGAGGCTATACATAGTC	AAATACCA		LIQSIMYIFRLLGSEDP
TTCTAGACTTTACGAATGCTTTCGGGTGCGTTCCGTTAGATACTCTCAT	CTTGAAAT		TLNSAIKHDILSHLNL
GTATGTCCCTCAACGCTTTGGCTTAGGCACCTCTGCTTTAACGCTGATTA	TGCTCCTT		KGFVNINFSQAISIFN
AAAACTTATATCTAGATAACTACAAAAATGTAACATGTGGGGAAGCAA	CCATACCT		SNFTDRDHFHLSR
AATAGAAAAACGTAAAAATTAATAAAGGGGTTAAGCAAGGCTGCCCTCTA	TTACCTTC		TEWARLQLARKKLKS
TCTATGCTGCTTTTCAACATTTTATCAATATTATAATTAGGGCAATAGAA	AGCAAGTA		TLAIQTNVCLINGHLV
GCTATGCCAGATGTCCATGGATACCCACTTGGAGATATGGACATCCGGA	CCTCGGGT		LTLSLENNVLLDSKEK
TACTGGCATATGCTGATGATATTGCTCTAATATCTGACTCCCAAAAGAC	TCTCAACA		GDVKKIHASLMGFLR

CTGCAGGAAATGGTCTACAAGGCGGAATATATCGGTGGATTCTTGGAC TACTCTCAACCCGTCAAAATGTGCACTTATGGACATTCGCGCAGACAAG AAGAGGACGCCCTATCTCGTCAACGGTGAGATGATCAAGTGTGTTG GAAAGGCCGACCCATACAAATATCTTGAACCTTTAGATCCTGGTCCG GAAGCTGATATAAGGAGCTCTCCAGATGATGATGATGAGACTAA ACTCATCCGAGTCAAAATCTACATCTCACCAAAAAATCCACGCGTATG AGACCTTCATTACAGCCAGCTCCCATTTACCTTAGACACAGCCGAATT CCGTTCTCAGACTTATAACAAACAGAAAAACAAACAAACAAACA ATTCAAACGACTCAGAAAAATCTATACAGAAAGCTACGATCCGGAATC AGGACAAATTATCTCAACACCTTCCCTTCCAAAGTGATGTGCTAAGG ATTTCTTTTACATTACAAAAGATGAGGTGGACCTCAACTCACAAGCGG ACTGGATGAGTACTTAATCCAATCAATTATGTACATCTTCCGACTATTGG GCAGTGAGGACCCACCTTAAACTCTGCAATAAATCAATGATCTCATTTCC CACTTAAATTTAAAGGGTTTGTAAATATTAATTTTCTCAAGCCATTTC ATCTTTAATTCAAATTTTACCGACCGAACCGATCACTTTTACATCTTAGC CGCACTGAATGGCAAGACTTCAATAGCTCGGAAAAAATGAAGTCAA CCTTAGCCATCCAAACTAATGTGTCTGATAAATGGGCATCTTGTCTTA ACTCTTCGTAGAAAAACAACGTTCTGTTAATTGATAGTAAAGAAAAAGG GGGATGTCAAGAAGATCCATGCATCCCTCATGGGGTTCTTAGTTAGC TCACCTTATCAGACTGCAAAACATGGATGGTCAAAACTGCTCTTCAGT GCGACCACTCATCAGAAATACTAAATAAGCGTATCTTGAATGGTCACG TCCCTTATAAGATTGGTACTTTATCATAGGCGCGCTGGGTGTTG CCTACTAACTCTTAGTGTAGTAACCTTTGTAGGAAGTGGGGGGA AGAAAGAGACCATGTCGCATGCTTGGTCAACTGTCCAATGATGCAGAC CCTCATTAAATGAGAGACATGATGCTCTTGAATCTCCCTGTACAAATTC TTTCTTAAATTTCAAGGTACGGTTATAAGGCAAAAGACCTATGTCAAC GAGTTAAGACCCGATATCACAATGGAATCGGATACCAATATTTATCTTG TTGAGGTAAATGCCCTTTGACACGAAGATGAGTTTGAATTGAGAAC ACAACAAACTACTGATAAATAACAATTTATTTGAATATTAGAAGATG TAGACCTGGGAAGGAGGTGCGCTTGTACGTTTATTGTAGGCACCTT AGGCTCATGGGCCCGCAGAACTCGGACTTTTGTAGAGATCTGGGATTTC TCCAAAGACGAAATCGACAGGTGAAGACGCGGCTGATGCTTCAGAAAT ATCAATTCCTCTGCGAGCAGTGGAAAAAGATTGTGCAATATGCACCCA CAATTACACCTGGGCCGATTCCAGACGCGGAGAGCGAGGACGATCAGG GGACGAGCGACAATGGGCCAACAGCTGTACAGTGAAGGACCGGTGA TTGGCGATGAGGAGGAGGAACTTCAATCTACGATTCGCGCCTTGACG AGTCCAGCGATGATGAACCCGACCCAGATGATGCTGAAATTACTTTTAC AATTGACATAGAACAAATATTGAATTTCTGTGATAACAGACTGATCCGCTG	ACCGGCCA ATCCTCCA GAAGACG CTACTCTA AGTGATAC GGATCTCT TCCTTACA CAGGATGA TCCCGATA GTCTTATT CTTCTGG AAGTACTC AACCAACC TTTGTTGA CCTCAACC CTTCACAG CAATCGGA ACTTCCT CAAAATCT GACAGCCA AAGATTG AGGCGGG TGAACAC CCAAATC ATAACTC TTACAGGG ATGACCTT TTCTACTCT ACAGTCTT TCACTACA ACTCAGAT ACAGGTTA CGGTATAA GTGTTGAC AATGGGG AGCAGAG GTTTCGAA TTCTTGCT	LAHLRLQKHGWSKL LFSATTHHEILNKRIL NGHVPYKIWFHRA RLGLPTKLFVSNLCL RKCGGKETMSHAL VNCPPMQTLINERH DALEISLVQILSSKFQ GTVIRQKTYVNELRP DITMESDTQYLYLEV KCPFDTKMSFELRTQ QTTDKYNIIEILEDVH PGKEVRLVTFIVGTLG SWGPNQNSDFLRDLG FSKDEIDQVKTRLML QNINSCEQWKRFV QYAPTITPGPIPADES EDDQGTSDNGPTAA TVQGPVIGDEEEELQI YDSGLDESSDDEPDP DDAELLFTIDIEQYLN SVITD (SEQ ID NO: 1422)

	GTTTGTGTCGTATGATTGTTCCGGTGTGTCGTCTATATTTTCTTTTTTATAC TTTTCAATTACCTCGTTGTAATGTTATAACTTCATATGGAATATATGTAATT TAGTTTAGTTTAGTTAGTTTAGTTTAGTTTAGTTTAGTTTAGTTAGTTTAG TTAGTTTAGTTAGT (SEQ ID NO: 1054)	
		AGGAATCT TGTCAGGA AAACCAAG GATAAGTT CCCCTCTT ATATGCTG GACAAGTA ATTAGACA CACAGTCT TCTTCAAT CACTTCAA CCAGGCAT ACTACGCC AATAATAT AACTGATA GTAAAGGT AATCTAAT TGAGTTT CTGATGAT AAGCCTT TCAAAGTA TACCGACT GACCCAAA AACTGAAC TAGAGCAA ATTAGGAG AGAGAGA CAACATCT AGTTGATA GAGCTCTT AGACATAA TCAGTTAC GGGAAACT TATATTTTA AATAAACT TAATAATA ATAATGGG GGGGGTG

NeS L	Utopia - 1B_CP B	.	Chryse mys picta bellii	GTTTAAATCCTTCTGATGGACATCTGCAACACCCGTCCTGAAGATGGAGTCTCTGCGWKCATTITTTGAAAAAATTGATGCTGCTTGMAGATATACTCGCTGCTGCTGTTTTGGAWGAMAAATTTCTCTCTCTCTCACTCWAGMTGCASTCWGTCAMCTSCTGCTGCTCTCGTCTCTACTCCCCAGAAA ACTCAGWGGAAAGCCTATCCCGAAKACCCTCTTGGTGCTCACGGGAAGAMCMGGACCAACASCAAGGATGAAAAAMATCAGSASCTGGCKGAAGAAAGCCCCCTGTGGATACCTCTKAGGGAGAMCTAGCACCAAGAGGACAGCTCTTGGGACCTCACATCCAGGAGCAGKAATATCWCAMCAGCTCTTCA GGAGGAGGACCCCGGAGAACCCCTCCCWCTTCCCGGGACCAAGGATGCTGAGCGCGCTGCTGAGAAAGGCTGTCTACCAAGAGGAGGCCCCCGACGATCCAGGACCAAGGATGCTGATCGCTGCTGCCCTGCTGGGAGGGATGCCACCGGAGGAGGCCCGGACGACCCAGGACCAAGGATGCTGANCCTGCCCGCTGGGAAAGGATCGCGCGGAGGAGCCCCCGGACGACGCCGGGACCAAGGATGACCGCGCGCCCGCTGCTCCAGAGAGGGATGCTCCMGAAGGAACCACTCTCAACCCCGGACCCCGAAACCACTTACCAACCCGCTGTCCGGAGGAGGGCGCTCCAAGGGGAACGCACTCCGSAGCCCWGATCTCGATCCGCGACGCTGTCTTCCGGGCAAGAGACATCGTGGCCA GTGAGTCCAGCACCCCGGAGGCGACTTCACTCTCTCAAGCTTCTCTGCCAGACCSAGAGGAATCACCTGCCGAGTCKGAGGGCAACAACAGAGGTCGCCCCACAGAGGGTGAGGCAGGGGAAGACGACTGCATCTACCTCCAGTACCCGMTCCCTACAGGCTCTCTCTGCCCCCTTCTGCTCTCCCKTCCATGGAGTCCAGACCTCGGGGCCCTCAGCAAAACACGTTTCGKAAGGCCACAAACAAACGATTGCTTCGGGTGAGCCGCTCGCATGACCCMTTCGAGACTCAAAAAGAAATGCAAGTMCCATCAMGCCACATGCAAGGACCCCTCAACACCGCGAAGGTSAAACCCCACTGACACCTGCGGGTTCACACCCCGACCCCAACCGACGGTCCAGCTTCAGCACCCCGACCCAGCATCCCCAGAGCC	GCGAACAT TTGAAAAG GAAAAAG ATCAAAGT CAATACGG ATGATGTC TCCAGCAA TGATGGAG ACAGAAAA CATAG (SEQ ID NO: 1177) GTTTAATT CCTTCTGA TGGACATC TGCAACAC CCGTCCTG AAG (SEQ ID NO: 1178)	TAACCG AGACCG CCGACC AGGAA ATAACC CACTTC TTCCTG ACGAAC CAAGGG ACGCAC CCCAACC ATGTACT TATTGCG TACACC GATACT GACTTG GACTCCT TATACAT TCCATG GGTGGC GTACCC GAGCCC ACTTATC CACTGA CACTTA AAAACT	MESPAXIFEKIDAAIX IYSAAAXLXNSLSLSP XXAXXXXAAPASST PQKTQXKPIXTTLG ASRKXRTTXKDEXIXX WXKKAPVDTXGRX STRRTALRDLTSRXN IXXALQEEDPRRTPPX SRDQDAERRPAAPEK AATRGAPPTIQDQD ADRCPAGRDATGGA PRRPRTRMLXAAPLG RMPPEEPPTTRDQ DADRRPAAPERDAP EGTTSSTDPETTYHP PVRRAAAPRGTHSXA XDLDAARCPGQORDI VASESTPPGATSP QASLPDXEESPAESA GTTEVRPTEGEAGED DCIYLQYPXTGLLLC PFCLPXHGVTLGAL SKHVRKAHNKRIAFR CSRCDAFETQKKCK XHXATCKGPLTTAKV NPTDTRLVPPTPTD

ACAGCANGTAAGGGGGACCAACCGCCCAACCGAGGGGAAGCGCAACCC	CTTGAC	GPASAPQPASPEPQX
CCGCTCGAGGACTGACGATGCCACCAAAAGGACCAGCCCGCCTCCAG	CCCAATC	VRGDQPTEGSATPA
AATCCCCACGCTGGACCTGCCGTGAGGGGATCACCGCCACCTCSAG	TGGTC	SRTDDATKRTSPASRI
GTACGGACCTCACCAGATGCTCAGCGACCTCATATAAAACCATCCGGC	TATGCC	PTLDPVAVRGITATSQ
ACAAACGGGATACGAGCGGKGCAGCGCTCCCCACAGGTMACCTCAT	GGTTAT	VSDLTRCLSLIKIRH
GCCGCTGCCGTAGGACAACTAGCACTGCCCCCAGGCGCACGGC	GCGATA	NTDTRXSAPPQVTS
GAGACCAACGAGGAGCCTCCGAGCCCCCAGATCCACGGC	TGTATGT	CRPAVGATSTAPQA
CGACCCCGCCCCGGGAGACCAACACCTCTCCAAGGTTACCCAAG	ATCTCT	ARRDPANGGASRSP
AGATCCGACCGCCAAACCCCATGCCACCGGAGACCCCCAGCCG	CATCCT	QIPRPDPAPGRPNTS
GATACACCGCAGAAACCAAGAACCATCCCCAGCGCTTCCAAACAG	GCAACC	SKVTQRDSDRQKPH
ACCGCCCCCAAAAAGCCCAACACCGGTGTTCCAGAACCCCACTCCT	GATACC	APRTPQPDTRRRT
CCGGAAAGATCCAGCGCTGCTCGGAGACACCGAGGCTGCCCCCCC	TGTAATC	RTIPSASKHDRAPTKP
CACGCAACGAGACCCCGCTCAAGACCCCACTGAACACCGCTCCCA	CCTCAT	NTGVSRTPLPPGRSS
GTCCGAGGACAACAAAGGCCACAGACTGTCTCCGAGCACCTGAACCC	ACCCAA	AASETPRAAPPPPHQ
GCGGAGACAACGACGAGGAGAAACGACGGCCGCGAGCAGAGGGTGG	GCCTGA	DPRLKTHLNTAPQE
CCACGCCGTGGCAATCCGCTGGATGGAGGAGTGGCAAAGGTGAGG	CCCCAG	GQQGHRLSPLHNP
ACTTCGAGACCTTCGACACCTGATGGACAGACTGACCGCAGAACTGTC	ATGTAC	RRQRSRRNDGREQR
TGCGAAATACAGCCAGAAAGGAGGAAACCCAGGAGGCTCACGGG	AGTACC	VATPWQSAWMEEL
CCACGCGCAGATCCCCGCGCCGACCCGTAAACACACTGCCAGAGAAG	TTCCCTC	AKAEDFETDLMIDR
GCAGGAGAGGGACGTGCGCCCGCTACGATCCGCGAGCTGCTATCC	TTAACTC	LTAELSAEIARRREP
GTATTCAGAAACTATACAGGACGAACCGGACGAAAGCCATGAGGGAGA	GTGTAT	QEASRATRRFPAPTR
TCCTCGACGGACCTCCTCTACTGTGCCATCCAGCCCGAGAGACTCTAC	ATTTAAT	NNTAREGRRGDVGR
TCCTACTTCAAGGATGTTCGACCAAGAGGCCCCAGACCAACTTGCAAC	TTTAAAC	RYDPAASRIQKLYRT
GCCAGAGTGCCTTCTCCGCTACCCGGATCAACCTCAGGAGGACCT	ATTAAC	NRTKAMREILDGTSS
GGAGCGAGATTTTCCCGCAGGAGGTGCAGGCGAGGCTGATGAGGAC	TTAATAA	YCAIQPERLYSYFKDV
CAAAAACACTGCCCCGTGAAAAGATGGCATCCGCTACCACCTGCTGAAG	AATTTT	FDHEAQTNLQRPECL
AAGCAGACCCGGCTGCTGGTGTGCTGCCATCTTCAACCAATGCA	AAA	LPLPRINLTEDLERDF
AGCAGTTTCATCGGTTCCCGCTCTGAAAAGTCCATGACCGTGCTC	(SEQ ID	SPQEVQARLMRTKN
ATCCAAAAAAGCGAGCGAGACGACCCCGCAACTGGAGGCCCATC	NO:	TAPGKDGIYHLLKK
TCCCTCTGCTCCACCATCTAAGCTGTATGCCAGCTGCTCGCGGCAAG	1301)	RDPGCLVLAIFTCK
GATCACGGACTGGTCAAGTGTGGGGGGCGCGTCAGCTCAGTACAGAA		QFHRVPRSWKKSMIT
GGGTTTCATGCTCGGAGGATGTACGAGCACAACCTTCTCTCTTCAG		VLIHKGERDDPGN
ACGGCCATCCAGGAAGCCAGGAGGTCCAAAGGAGGAGTGCGCCGTAGCA		WRPISLCSTIYKLYAS
TGGCTTGACCTGACCAAGCCTTCGGGTCCATACCCCAACCATCATCTT		CLAARITDWSVCGG
TGCCACCTGGGAGAGTTCGGGATGCCAGAAACCTTTCATCCAGATCCTC		AVSSVQKGFMSCEG
CGGGACCTCTACAAGGACTGCACCAACCACTCCGCGCCACGGACGGA		CYEHNFLQTAIEA
GAGACGGACGCCATCCCCATCCGCGCGGCGGTGAAACAAGGATGCCCC		RRSKRQCAVAWLDL
CTCAGCCCCATCATCTTCAACCTGGCCATGGAACCGCTCATCCGAGCCAT		TNAFGSIPHHIFATL

CTCAGCGGGCCGACCGGCTTCGACCTGCACGGCAAGAAAATCAGCAT CTGGCTACGGGACGATCTGGCCCTGGTCGCCGACAGCTCGAAAGC CTCAGCAAATGCTGACGTACACGACCAAGCCGCCGAGTGGATGGC CTCCGCTCAACCCCAAAAGTGTCCCTCCCTCCACGTGACGGTGGC CCAGGGCGTGTCCGGCCATCAGGATCTGTATCCAGGGGAGCCCAT GGCTCCCTCGAAGAGGAGAGGTATACCAACACCTCGGCACACCCAC AGGATCCGGTCCGACAGACCCCGAAGACACCATCGCGGAGATCCT GCGAGCGGGCCAAATCGACTCTCTCTGCTCCCTGCCCCCTGGCAAAAG ATCAACGCCCTCAATCTCTGATCCCCGCTCTCTTTGCTCTCAGG GGATCGCGTAGCCAAGGTGCCCTGAAACAGGCCGACAGCACCATC AGGAGCTGTGAAGAGTGGCTCTACTTCCCAAGAGGGCCAGCACG GACATCATCTACATTTCCACAGGAGGGGGCGCCAACTACCTCGGA TGGTGACCTGTGCGACGTGGCGGTGATGACCCACGCTTCCGCTCCT GACGTGCCCGGACCCGACGTGAGGAGCATCGCGCAGGAAGCCGTAC GGGACGTGTGAGGAAACGATCGCCAGGGCCCTCCGAGCAGGACA TCGCCACTTACCTCAGCGGCTCCTGGAGGCTGAGTTCGGGAGAGAGG GGGAGACCTGTCTCTCTGTGCCGCGCCGCAACGCTCGAGACG CCTGGTAAGAGGATCGGCTGTCTGTGAAGTGTGCGAGGAGCGCC GGGAGCTGGGAATCTGTGCCACGCTAAAGACCCCGACACACCA TCGTACCCCGACCCGACGAGCTATGCTGGAAGGACCTGAAAGACG CCATCCGCTGCCATATGCCGAGAACCTCAAGCGGAAGCCGACGAGG GCAAGGTGTCAGGTGTCCAGCAAGTGGACGCCAGCAACCACTTC TCCCGGGGGAGCTTACCAGGTTCGCCGACTGGCGGTCTGTCACAG GGCCGACTCACTGCTTCCCTCAACGGAGCCATCCGCCACGGCAAC CGGACAAGCGCTGAGGAAGTGGGCTACGCAAAACGAGACCTGCC CAGTCTGTGTGATGCAACCGGTGTAAGGCCATCCCGCGTCCCTGG CACAAGCCATCCAGAACCGGTGTAAGGCCATCCCGCGTCCCTGG GGAAGTCACTCGTCCGCTTCCGCTGTAAGGCCATCCCGCGTCCCTGG GACCGACATCGTGTGACGACGCAAGAAAGAAAGAGTCTCATGG TAGAGTCAGGTGCTTTTGAAGACAGGTACCGGCTTCCAGAGGC CCGAGCAGGAAGGCTTGAAGTACACCGCTGGCCGAGACCTGAG AGCCAGGGCTACGAGGTCCAGATACACGCCCTGATCTGGGAGCCT GGGCTGTGGGACCCCAACGAGCGGTCTGAGAGCGTGGGAGT CGGTGACGCTACGCCCGGTGATGAGACAGCTCATGGTCTCCGACAC ATCAGGTGGTCCAGAGACATTTATACGGAACACATCAAGGACACGTC AATACCACTGAGTAACCGAGACCGCGACCCAGGGAATAACCCACTT CCTTCCCTGACGAACCAAGGAGCGACCCCAACCATGTACTTATCGCTA CACCGATACTGACTTGACTCTTATACATTCATGGGTGGCGTACCCGA					GEFGMPETFIQLRDL YKDCITTIATDGET DAIPRRGVKQGCPLS PIIFNLAMEPLRAISS GPTGFDLHGKKISILA YADDLALVADSESL QQMLDVTSSAAEW MGLRFNPKKASLHV DGGARALVRPSRFLI QGEPMASLEEGEVY QHLGTPTGVRVRQT PEDTIAEILRDAQID SSLAPWQKINALNT FLIPRISFVLRGSAVAK VPLNKADSTIRQLVK KWLYLPQRASTDIYI SHRQGGANVPRMG DLCDVAVMTHAFRL TCPDPTVRSIAQEA RDVVRKRIARAPSEQ DIATYLSGSLEAEFR EGGDLSSLWSRARN ASRLGKRIGCCWK WCEERELGILVPRIK TPDHTIVPTARAML ERTLKDAIRCHYAENL KRKPDQKQKFEVSSK WDASNHFLPGGSFT RFADWRFFVHRARLN CVPLNGAIRHGNRDK RCRKCGYANETLPHV LCGCKQHSAGWRHR HNAIQNRLVKAIPPSL GKITLDSAIPGTD SRL RPDIVVTDAEKKVL MVDVTVPFENRSPA FHEARARKALYTPL AETLRAQGYEVQIHA

NeS L	Utopia - 1_ACa	.	Acantha moeba castella nii	GCCCACTTATCCACTGACACTTTAAAACTCTTGACACCCCAATCTGGGTC TATGCCGGTTATGCGATATGATGATCTCTTCATCTTGCAACCGGATAC CTGTAATCCCTCATAACCCAGCTGACCCAGATGATGACAGTACCTTCCC TCTTAACCTCGTGATATTTAAATTTTAAACATTAACTTTAAATAATTTTAA AA (SEQ ID NO: 1055)	CCCCTCAA GGGTGCTC CACGAGAT CCCTGTCG CTAGCCGA CCGGTTTT ACCACCCC ACCCCGCC CGGACAAC CAGGACC CTGCTCCG CAGCAGG ACCCACG CAGC (SEQ ID NO: 1179)	TAACAA CCATGT ATGGTG AACCAC ACCTCTC TCGATCT TGATTC TGTGATT GGACAT CAGAGT TCCTGC GAAGGG ATACACT CTGCCA ATCTCGT GGGTG TAATAA ATCCAC ACCTTCA ACA (SEQ ID NO: 1302)	LIVGALGSWDPHNEP VLRACGVGRYARL MRQLMVSDTIRWSR DIYEHITGHRQYHTE (SEQ ID NO: 1423) MAAKSVACPHDGA NKYASEASLRHIKNK HATDEEGDETSCHP HCHRPFTARGLSVH IGKSHRQAPPEPTRP PPAPAPADPGLDPPD GPTVTPPSRDEDEDE EPDDDPVEIADLSCP HCAQALPSAHGLAN HLRACKDHRVPAPG APRSGPPSSRYWTAV EHRYVEAMARFAD HPDLLARAAAHIGTR TYKQVDSHRTKVAIA EREGRPVRTLDPTM DWRMRPYCASTAR WLAEQGRSPVAPRS PCPEPHAPPPAAALL YIPATPPAPTTPRAPVA PPKLAPPAESTVPATP DGNPEAPAPPFSAAG PPTPKALPPPPPSRR NLRHLVPKDAWQWQ VADAVAPAAASRLRT PLAHLSTEQWATFEA ALAGLEATLHHAARS AEAVPTRCASRARED AERQLREARKTREIFG KAAALYAAAGKDPTAT IERIPPEVRLHLPTGS AEWPARAAAAARRVI RRAVARADRLRKRM GILSDRDLQLRFNA

TCGCCCGTCCACCAAGGCGCGCGGTGCCAGCTAGACCCAGCCGCGCGT CGAGGAGGCTTACATCCAGACCTCGCAAGCCGCGCGCGCGGATCGACCCC AGCCCCCGTGAAGAACTCCGTCCAGTGGCCCCCGCGCGCGCGCGCGG CCGATGACGGAGGAGCCCTTACGCGTCCGCGAGGTCCGGGCGCGCGG TCCGCCGACTGCCAACGGTCCGCCCGGAGGATCGATGGCATACCGTA CGAGGCTACAAGCGTACGAACTGGACGCCACGCTGCCCATGCTTTC GAGTCTGCGGCTGAATGCGGCGCTGCCAGCTCGATGGGATGTGGCG CGACGGTCTCTCTACAGAAAGGCGACCTTAAGATCTTACGCGCGGCACT GGCACCGATAAGCCTCCAGGTACCATCTATAAGATCTTACGCGCGGCGG CCTGTCGAAGCGGCTATCTCTGGGCTGGCAAGCACAACTTTCTTCC GCATCGCAGAAGGATTCCTACCGGCGGAAGGCTGCCACGAGCAGCGG TTTGCTTGGAAAGCGTCTTGACGACGCCGCTCGGCACAAAGCAGAAGC TGTAACCTTGCTGGTACGATCTGCGCAAGCCTTCGGATCGGTGTGCGCA CGACCTCATGCGCTGTGCGCTGCCATGTTGGGCTGCCCGCGCTACCTCC GGGATGCCATCGGCGCAATCTATCGGCACTAGCGCTCTTCTGTCGCAAGT TGGGATCAGGAGACCAACCGGCGTCTTCTATGCGCTGCGGCGTCAA GCAGGGCTGCCCTCAGCCCCCTCTCTTCAACCTGTGCGTGCAGCGG GCCCTTGGCTGCTACGCCGCAACCGGCTACCGGACAGTCTACGGCACGT CGATCACCGTCGAGGGCGAGGCTACGCCGACGACCTGCTCACTGCCGCG GCCCTCCGCTACCATGCGGCGCGGAGGTGCCACGATCGAGGAATG GGCCAACTGGCGGAGTCTCTTCTGCTGCTCCAGCCCTCTCCCTGGAT GCGCGCGCGCAAGTGTGCGGCCCTCGCGATCAACTTCGAAGGTGGT CTAATGCACTCTATGACCTGCCCTCAAGGTCCAAGGCGGCGAGCCATCCC GGCCATGTCAAGAAACAACGTGTACCGTACCTCGGAGTACATGTCGGT CTCACAGATGCGCTCGGCAAGCAAGAGCTCTCGAGAAGGCGCTCA CGGATGCAGCACGATCTGTGCTCTGGCTCGAACCCCTGGCAGAAGG TGTCGCAATCAAGACCTTCACTCTCCGCTCGAACCCCTCTTCTTCCACA ACGGGAAGATCCAGAGGGCGGATGCCAGCAATTCGACCGGAGCTTC GAGAAACCTGCGGCGCGCTCGACTCCCGTCTGCACACGAAACGC CTTCTTCCATCCCGCTGGCTCAGGCGGCTTGGCATCTCTGCCATCG CGAAGAACAAAGTCTACTGGCAGCCCACTGTTCAAGTCTCTGAC TTCGCCAGATCTGTCATCGCGGCTCGCCGACACCAACTTGCAGG GTCAACCCAGCGGACACACCAAGCCAGTCCAGGAGCGGGAAGCGTCA CCCTTCTCGATGCTCATGCGGGGCGAGGAGTCCGATCAACTACCC CCTCGGCTGACGTCAATCTGTTGCGAGTGCAGGCGGCGCTACTC GAGGATGGGATGTCAGTCCGCGATGCACTCCACCGGCGCTGACAGT TGGTCCGGGCGTCAATTCGAGGCGGATTCACACGTCGCAACGTCATC CCAGCTCTCCGGGCTAGCGCTTTCCCGCATGCTGTGGAATGGAGTG			NQKAVRQJLAPSTK APRQLDPAAVEEAY IQTIAKPPIDPSPW KNSVQWPRPPTAAD DGGSPFSAEVAQL RRLPNSAPGIDGIPY EAYRTKLDATLAHV FEVRLNARLPARW DVARTVLLYKKGDPN DTGNWRPISLQVTY KIFTAALSKRLISWAG KHNTFSASQKGFLPA EGCHEHAFVLRSLD DARRHKQNVYLAWY DLRNAFGSVSHDLIA WCAAMLGLPRYLRD AIGAIYRHSALFVQV GDQETTGVPIMRCG VKQGCPLSPLFNLC VEPALRCLRRITGYKF YGTISITVEGQAYADD LLTAAPSAHYAARQV ATIEEWANWAGVSF VVQALSLDAPAGKCA ALAINFEGGLMHSID PALKVQGAAPAMSR NNVRYLGVHVGLT DALQANELLEKASR DARTICASGLEPWQK VVAIKTFILSRLPFFH NGKIQRGRCQFDR ELRENLRALRLPVCT TNAFFHSRVASGGGLG ILPIAEEQQVYLAHV FKLLTSPDLISIRAIARH QLAEVTHARHTTPV QDGEASPFEGWLMR GOEVASTTPSGDVSS

NeS L	Utopia -1_AEc	.	Acromy rmex echinati or	CCCTCCGACCCAGGGTCTGAGCAGCAGCCTACCAACATGCCGTCCACCC TGCAACGCACCACTGGGTCCAAACAGCGCTGGCTGACGACCAAGGA GTACCGATTCCGATCAAGTGTGATTGGTCTCTGCCAGCGGAGCA GCTCCACACCAACCGCAATGGCCAAACAGCGTGCAGGGCGTCTCTAC GCCCGAGAGCGGCAACCATGTTCTCGGACACTGCCGCTCTGCCAGCAAG GCCGAAGTATCGCGGCCCAACAGGATATGCCAGCTCTGCCCGCAG GCGGTGAAGCTCATGGACGTCTCTTGAAGACGTCCCGATCCCGG GGGTGACTCCCCCTACGACCCGACATCTACTGCTCTCGCGCGGCCA GTGTGCCATCATCGAGTCCGGTCTCTACGAGGACGCTTCAACGCT TCGATGAGGGCGGCGAAGCAGAAACGACGACCAAGTACGCTGGCCT GGCTGTACCGTCTGAGGAGAGCTGGCGTCTCCAGCCTCGGTAAACCGCTA TTCTGTGGTGGCTTCTGCGCTGCTCCAGCCTCGGTAAACCGCTA CGGCCACCTCCCTTGATCTCCCCCAAACTTGAATGTCTTCTTAAA CGTTGTGTGCTGCTCAATCAAAGGCGATTACACAGCGTGGAGAAGAT TCCGGCGCTCTACTCCATAACCAACCATGTATGGTGAACCACTCTCTC GATCTTGATTTCTGTGATTGGACATCAGAGTCTCTCGAAGGGATACAC TCTGCCAATCTCGTGGTGTGTAATAATCCACACCTTCAACA (SEQ ID NO: 1056)			IWFAAAGAYSRMG WSVRDALHPTLTVG PGVQFEGRFQRANVI PALRASAFSRHAVEW SALRTQGRAAAAYQH AVHPATHHWVHNS AGLTKEYRFAIKRL GLPTRAAPHHRNGP TACRACSYARETANH VLGHCPATKAEVIAR HNRICRALAQAEAS WTSVLEDPPIGVDS PLRPDIYCSRPQGCAI IEVAVSYEDAFNASM EGRAKQKTDKYAGLA ATVEEQRLRLQTRHAA FVVGFGVVLPAVS ATATSLDLPKTNWV LLKRCVAASIKGSYTA WRRFRSTP (SEQ ID NO: 1424)
				GGTGCAACCGGATGCATCATACGTGTACCGGAGCATACGGGCTGTCA CGCGGGTGCATGCGGATCTAGCTCGGAGATTTTATTATTATTTATT AATTTATTTATTTATTCATCGAGTGTGAGTGTGCGGTTTTCGCGAGAAG CGATTTTCGTTAAGTGATACGCGCGCGGTTCTAGGTTAGGTGTGCGAGT GTACGTGGCTGTGTCGCGAAGATTCTCGAAGATTCTACAAGTTCAAAT TTCCGTTAAATTCGTTAAAGTTCTAAACCATCTGTGATCGGGAGTGCG TTCCAAAATATCATGATTTCCGGCTGTTCCGAAAACCACTCCCGGTCGAC CCGCGTTCGAAAACCAACAGACACGCGTACAGCAATACGATTGGTCGG TACACTGCCGCGAGTGCGAATAACTATTTAACTGTTATTACTGGAAA TTACACTGTATTCGCGCAGTGATCTGTATTCGCGAGTGCACTTGGCTTT TAAGTAAATTTGTTAAATTTTCTTAACTATAATTTGGATACTTTTCCAAAT GCGGCTGCTGTTATTTATGAGGGGCCAGTATCTAAGGCCCTTTTAGTC AGACTAAACCGTAGTGCACTCGCCATGGCGGATCCACAAGTGCACTATA TAGACTACCCGCTGCCCTAGAGTCAAAATGCGTAAAAATGTTTCGGTGC TGAGGGGGCAGGCAAAAGTAAAGGGCGAATACAGCGACCCCGCGCATTT AGCAAAACATCTGAAAAAGTGCCACCCGGGAGACACATTAAATTATAAA TGCTCAATTTGTGATCTAAGGGGGACCGGTAAATACCCCTTAGAGATG	GGTGACACA ACGGATGC ATCATACG TGTACCGG AGCATACG GGCTGTCA CGGCGGCT GCATGCGC GATCTAGC TCGGAGAT TTTATTTAT TTATTTAT AATTTATTT ATTTATTC ATCGAGTG TGAGTGTT CGCGTTTT	TAAATTA TTTTGTC TTTGCT TGGCCC CCCTTT TTAAACC AAGCAG GAGAGA GTGGCC CAATGC CCAACT ATTATAT ATTAAC ATTACT GTGATA TTTATTA TTTGACT	VCSVRGCRREDSRRF YKFKPLNFVKPKTI VIGSAFQKSSVSARS QNHSTRVPKTRQP RTSNTIGRYTAASAN NYLTVITGNYTVFAQ WICYRECTWLLSKFV NFFLTIGYFFQLRLV IYEGPVLDTFSCGS SLFMRGQXSKALLVR LNRSALAMADPQVH YIDYPLPRVKCVKCF GAEGAGKVKEYS PPHLAKHLKKCHPGD TLNYKCSICDLRGTK YPLRDVKAHYAECHV SPAVIDAAGPSTRGSL

TTAAGGCACATTACGCCGAGTGCCATGTGTCTCCCGCAGTGGATGCGGC	GCCGAGA	GTGGG	GECSGAGQPTASRA
GGGTCCAAGCACTCGCGGAGCTCGCGAGTGCAGTGGTGC	AGCGATT	CGGGC	AKATRLAETVGGTD
ACCGACAGCCAGCCGCGGCTAAAGCGACACCGGATTGGCGGAGAC	TCGTTAAG	CCTCTCT	KRRAATSGSRQLTLP
GGTTGGGGTACGATAAGCGCGTCCCGCGACATCGGGATCGCGGC	TGATACGC	GCTGGT	FAATPSPSTAAGEAR
AACTCAGCTGCCGTTCCAGCCACCCCATGCCATCCACAGACGCCG	GCCGCGTT	TTTATT	APRSXTTPTSRPSY
TGAGGCAAGGGCCCAAGAGCGMGTCAACGACACCCAGCAGCAGGT	CATAGGTT	ATATATA	AAVTAGPPSMRSTTT
CCCCCTCATATGCGGCAGTCACTCGGGGCCCGCATCGATGAGGACAC	AG (SEQ. ID	TTTTTA	STTARSKTVAKGAAP
GACAACTCCACACAGCCCGCAGTAAGACTGTCGGAAGGCGCCGC	NO: 1180)	CTCGC	NTTTTTARRSGEAA
GCCCAACACMACGACGACCAACGGCCAGGAGATCCGCGAGGCCG		TACTTTT	ATRPPTTATVSKPR
CCGCAACGAGGAAGCCGCTACGACCGCCACCGTGAGTAAACCGCGTG		TGTAATA	VLSVETVRLPVDDIQ
TGTTGTCGGTAGAACTGTTAGTGTCTGTCGACGACATCCAGCGAGC		CTCTATT	RAGVQNAAKPARAP
AGGCGTGCAAAACGGCCAAACCGCGCGCTCCCTCTCGCCCCC		TTTCTTT	SRPPQRTSPEAGGPR
GCAGAGAAATCACCAGGAGGAGCATACAAGAGTTGCCGCAACAGC		TTATTTT	TTGAKEKCGEGAYKK
AGGAGAAATGCGGAGAGGAGGAGTCAAGAAAGTTGCCGCAACAGC		AGCTAT	LPANSGNPISTRRR
GGCAATCCAAATTTCAACAGGAGGAGGCGGCAACTAGCGTGCCGTT		GCTATTT	ATSVPEKSEGTARR
GAGAAGAGCGAAGGACGGCAAGACGGGAGCGCTCTCCACACCT		TTATCTC	ERVSPHPPPKIDILS
CCTCCAAAGGAATTGATATCATCTATCGTGCACATCGGAGGAGG		TTTCTTT	STSEEGTPYQPGV
GCACGCCATACCAGCCGCGGCGTGGGAGACTAAGACTAAGGAGG		GTCCTTA	GRLRLRKKVTGPPP
AAAAAGGTGACCGGACCAACCCCAAGATGACACCCAGAGAGGGGT		TTTTCTT	KMTPREGVTVRARR
GGTCAAGAGAGCGGCTCCACAGCGCTCCCGTCGAGAAGAGTGC		TCTTTTT	STSAPEKSALDARLT
CTTGATGCAGCGCTGACGCTCTGGACCGGACATCGTCCAGAGCGAC		TTCTTTC	ALDRTSSRATGNPTS
AGGCAACCCGACGTCGCAATCGCAGGGGGCTTTACACAGTAGAGG		CTTTTCT	QIAGGLYTSRGQPER
CCAACCGGAGAGGACCCCTCGAGGCTCCCGCTCCCGCTCTCCGACC		TTTCTTT	TPPARLPSLPTTRGS
ACCAGAGGAGTCCATCGGGGAGCTAGGCGAGATACGGACACCCATC		CTTTTAT	PSGSLGEIRTPISPATS
TCGCTGCGACGTGCTACCGGCAACGCTCACCACCTGCACGGTGACGA		TCTTCTT	LPATLTCTVTTTTCG
CGACCACTGTGGAAGCCCATAAATCCACGGGCTTACAGGTGGCGT		TTATTTA	SPITSTGFTGGVGRLI
GGGAGGCTGATAACACCCWCCGAGCCTCCCCAACGAACATCTCCC		TCTTTT	TPSLPQTNILPTIGEE
GACCATCGGGAGGAAGAACGTCACCGTGCGTGGCGGTGTCACCCAC		TTCTTTT	GTSPCVAVVTHPRP
CCATCTTAGGGGAGGAGACGTCCTCCGCTCATCTACCAAGGCCAA		CTGTTGT	TGEDAPCEAPQPVSD
CGACACCGGAGCTGAGCGGGACAAGAGGAGCGGCGGTAGAAAGC		GGGGCC	HQRQSIGEPRRDIT
GCGGCGAGCCACTACACACCCGTCGTCGAGGGGGACAACCAAGTG		CTGACC	HLACDVATGNAHHL
GATGGCCAGTGACGGTGAGCGTGAGGAGAAAGCGAGGAGGCAACA		GTCCGA	HGDDDLHLWKPHNIH
ACTGAACGATACATCCCCCTCCAACTCCGAGTCCCCGCCAACCGCTGGA		GTGTGA	GLHRWRGEADNTXE
CCATCGCGTTCGCGCGCATAGCCCCACTATCTGCGCTGATAGCGCGT		ATGCCG	PPNEHPPDHRGGR
CGACGAGCGCCATGAGACCACTTAAATCTCAATTGCAGCAACGGCAA		CGAAAA	NVTVRGGRRHPSLG
TATTTGCATGGACCGAACTCCGCCCGCTTAAATTTTCCCGGTARGGGCG		ACAATA	GRSTSPILPRPTTPEP
GAGCGCCGTCGCGAGACATCGCCACAGGATCGCGTGAGGGGAGACATC		TTATGTT	ERGQEERRLEGAAQP
GGTTATGTTGCTGGAAAGGTAAGTGCAGCAACACCCGAGTGCTCCCGTA		TTATACG	PTTPVVEGDNQWD

AATGTCGGTGGTGTGATGTCGAGGGAGAGCAACCGCGTCATCCATCG TGCCACCGGAGCAACCGTGGGAGGGCGGTGCGGACGATCATAGTC GGCGCGTCCGACGCTCTGTGCGTCAGCCGTCGCGGGATCACCCGG CGCTGCGACTGTCGAGGAGCGTAGCGGTGAGCGGGTGGCGGCC GCGACGCGTCTGATCGGCTAAGGACGTCGCTACGATTGCGGATC TGGAGGCTTCGGGCTTCGTCGCGGCTTCCTCGGGAGGATGTCAT CGGCACTGGTGTGACGCCGCGCTCGGATCGTTCGGTACGCTCAG GGAGCGGTGCGCTGCGGGGTGAAGGAGGTGAGCGTCCGGAG AGAGGGCGCGGTAGCGCGGGTACGCGCGGTGACCCCGGAGC GTCGGAGAGCAGCGGGGACTGGTGCAGGAGGCGCAACGCTTGC AGCGCTGTACAGGCGAACCAGCGCAAGGAGGAGTGCAGAGGTGCTC CAGGGACCTGCCGATCAGTGCCAGGTGCCTAAACGTCAGGTCCAGGAG TACTTCGAGCGGTGTACAGCGCGGGGAAACCTGGCTGGCGCGGC GTGGAAGCCGAACGCCCTGACCCCTCGAGTCCGCGTGAGGTATCTGCG GTCCTGGTCCGCTCGGAGCGAGAGGTGACCGTCCGCTCCGGCGT ATGAATAACTCTGCGCGGTCCGACGGTGTATCTCTATCGTGACCTC GTGGGGGACCGGGAGCGGCTCTCACGGCGCTCTACAACATCT GCCTGCGCTCGAGGAGTCCCGCTCTCTGGAAGACCTCAACACTGT GTTGATACAAAGAAAGGAGACCGGGCATGTTGGAGAACTGGCGCC TCTGCTCTGGGGACACGCTCCCAACTCTTCGCGGCTCTTGCGCC ACCGATTACCGACTGGCGGTACCCCGGGAAGCTCTGCTCCGCGC AGAAGGCTTCTGCGGACGAGGGGTGCTACGAGCACAACCTTCGTC TGCAAGAGTCTGACGACCGCAAGCGCTTAAGCGCAGCGCGGTG TCGCTGGCTGGACCTGTCAACGCTTTGGATCGATCCCGACCGAC GATCCCGCGCGCTTAAGATCCGCGTGCCACGCGGTCTCATAGCG ATATGGACTCCATGTACGATGGTGCACGAGGGGTGCGAACCGCC GAGGTACACAGACCCATCCCATCCGTCGGGCGGTCCGTCAGGTT GTCCGCTAAGCCCTATTATCTCAACCTGGCCATCGACTCGGTCCGCT GTGGCGCGAGWCGAATGACGGGTATTCCTCCACGGAATACTGG TCGGCATTTGGCTTACGCGACGACATCGCACTACTGCCCCAGACGCCG AGGGATGGAGGATGTAGCTCTGTGGAGGCGGAGGAGCGCTCG GTGGGCTGCGGTTCAACCCCTGCAAGTGTGCCACCTGCACGTCCGTTG CGGGAATGGCGGAGGTCTACCGACGTCATTCCAAAATCCAGGGGG AGACGATCAACCCCTGCTCAGGGTGAGTCGTACACCCACCTTGGCGT TCCAACGGGTTCTCGTGACGAGCGCCCTACGCGCGCTCGGGGA CATGCTCTCGGACCTGCGGCTGTCGACCGCTCACTCTTCCCGCTGGC AGAAGATAGAAATGCTGGGACCTTCACTCTATCAGGCTGACTTTCT GCTCCGGGGGCCAGAGTGTCAAGGGTCCCTCAACGGCGCGTGGACCT					AGTGTG CATGTG CGTGAT ATATTA TCTATT TATTTA TTTATT ATTATA TTATTG CCGCGC GCGCTC CTCCGG GACTTTT ATTCGTT GACAAT ACTGTG ATATTT TCTGKC AGGCTG GGGGG CTTGCC CCGAGC CCCTAG TTTTAT TGCCTAT GCGGG GGGCT TTTGTC CCGCA AATGTA TATATAT ATATAT TAGCGC GCGGCT TAGCCG CTTTGT TTGTATT ACCCCA	GQWTVSVRRRRRQ QLNDTSPNSESPPT AGPSRPRIAPLSALI AASRSHETSLNLC TNGNICMDRTPPRNI LPVXAERRRETSPQD RVEDIGYGAGKVSA EHPAPVNVVRGVM RGRATASSIVPPRAN RGEGRQHHSRRRP DAPVGPQSRDHPAP ATVARQRRRERVAA RDALLDRAKDVATIA DLEAFAASVAFFGE DASATGAAARARDR SVRSREAGARRGVKG GERPEREGAGRP PADPGASGEARGDW VREAKRLQALYRANR RKAVREVLQGPADQ CQVPRQVQVEYERL YSGGEDIAGAGVEAE RPDPSPREVSAVLG PLAEREVDRLRRM NNSAPGPDGVSYRD LRGADRGARLLTALY NICLRLEAVPASWKT SNTVLHKKGDRGML ENWRPLALGDTVPKL FAALLADRLTDWAVT RGKLCSAQKGLRDE GCYEHNFLQEVLT AKRSKRQAVVAVWLD LSNAFSGIPHATIRRA LIRSAVPRGLAIWDS MYDGGCTTRVRTAEG HTAPIRSGVRRQGC PLSPIIFNLIDSVVRV

265

NeS L	Utopia - 1_AMi			ATGCGGGGGGGCTTTGTCCCGCAAAATGTATATATATATATATTA GCGCGGGCTTAGCCGCTTTTGTATTATACCCAGAGGGGAATTGTC CCTCTGGGAAAAAATGATTGAAAAATAAAGTGAGCTAA (SEQ ID NO: 1057)	TGCTGGAAACGAGGAGAACCGCTTCTTTTCCCTGCGCCTGGCCTGGT ATTGAGTACCTCAGGATTAGCGCAACTAGTCCGGCAGACTGTCGGA ATACAGCAATAGAAAGWAGTACTAGTACGAGCTTCTTCTCTCTCG GTGACGATGGTCTCTGTCAGTCMTGACGGGTAGGGAAGCGGTG CTGCCAGTACGTCGAAAGAGTGCCTGTTGCGAAGCGACGCGCCAC TCAGGTGAGTAGCCAAAGGCTTTACAGTTACCGGACCCGAWAACGCG AAAACCCCAACTCGGGCTAGTAGCCGAAGACCTGGTCCCGCCCMGGT CAGAGTAGGCGAACGCCWKGCTCAGAGGACGGAACGCGGAAACAC CCCAGGTCCCAAGGACGCCCTGATCCACTGACAAGAACGCTCGAGGCA CGCCAGGAGACCCCGAGTGGTGACCGCGACTGCGAGTCCGGAG GACCTCCAGGAGGTGGACAGCGAACCCGACAGGGCGGTGGACCA TGACGACCCCCACGATGTACGAGACCCGACAGGGCGGTGGACCA CTGACCATGACCGACCCAGAGGAGGAGAGACTCTCAGAGCCCGGA ACCCGGCTGACGAGAGCCCTCCCGGAGGAGACCCCGAGCCTGA GGATGCCCCCGGATGACGGCGGAGCGCCGAGCAGCAGCGGACCCC TCCGGACCCACCGCCCTCGGTGACGATGGCGGGCCCCGAAACGAG ACGACCCCGGACCCCGGCTCCGAGGAGCGCCCCCGAGGGTCT CCCCAGCTGGTGAGGAGCCCCGACCCCGACCCGACCGACCCCGCC ACGGACACCCAGGCGAAGCGGTAGACATGACAGCACTACGTTCTCT CCCTTCCCTCCCGGGAAGCTGTTCTGCCGACCTGCCACCCGCCAAG ACAGTACAGGTCGACGGCGACATGAACAGCACCTACGGCGCTTCCAC CAGTGGCCTAGCCTTCTACTGGCCTCTCGGCGACCGAGTACGAGG CCCTGAAGCTCCTGAAGAACCAAGAGGATGCGAGGGCCACGAGG CCGAGAGGAGACCGGACGCTGTGAGGTGCTGCTCCCGGCCCGCC GGACCAAGCGCGGTGCGAAGGCCCGCCAGACTGGCCACCCCGCCGA CAACCCACCGGACGAGCTCCAGGGACCAACCGAGAGAGACCTG CCCAGTGATGCCACCAAGGAGGCTCCGCGGAGGAGCCCGGAGGAGG CAGACGCCCGGACGAGAGGAGGCTCCGCGGAGGAGGAGGAGGAGG GATGCCCGGAGCGCGGGGATCCCGGAGGAGGAGGAGGAGGAGG GCCCGGGAGCGCCCGKCCCGGAGGAGGAGGAGGAGGAGGAGG GGAGCCTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG GGAGGAGCTCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG GCCCGGCTCAGAGGAGGAGCTCTCGGCGCGCGCGCGCGCGCGGCGG	TGCTGGAA AGACGGA GAACCGCT TCCTTTTC CCTGCGCC TGGCTGG TATTGCAG TACCTCCA GGATTAGC GCCAACTA GTCCGGCA GACTGTG GAATACAG CAATAGAA AGWGAGC TGACTAGC AGCTTGT TTCCTCT CCGTGCA GCATGGT TCTCGTA GTCMTGA CGGCTAG GGAAGC GGTGTGC CAGTACGT CCGAAAGA GTGCGGT TGCGCAAG CGACCGG CCACTCAG GTGAGTAG CCAAGGT	TGAACC CCCCCTC TGACCC AGATGG ACCTTCA CTTCGA GAGGAT TCTTCAG CAATGG ACGACC CCGCTCC ACCCGA AGAGGA CCCCCG CGATGA GACTCT ATATGG ACTGAG ACACTT TCTTCG AACCAC TTCCTCC ACCAT GCGGAC CAITGTA ACGGGT TTGTGT GTATCTA TCTTCT TCTCTCT CAGCGT CGCGAA CCCCCTC	GTRQYLAPSRPSGDL ATPPRAVRRRWLAEE RSAQDAARRGSDSV SVA (SEQ ID NO: 1425) CHHAGLRPGTPNRT RRPDQTAPLPDRG HPMPNRRGSRSP EESRREPPXPACQ GLRVWSPQQRMP PWQTLWLEELSRATT FKAFEASVARLTEL AAARPGQPRGNNR PATRRDHRLLQPRR PRRQRYDPAASRIQ KLYRANRPKAVREIL GPSAFQVPRETLFN YFSRVFNPPAEAAAP RPATVEALTVPPEAE GFEDAFTPQVEEARL KRTRDTAPGRDGIRY SLLKRDGCLVLSVL FNRCFRRTPTTWK RAMTVLIHKKGDPTD PGNWRPIALCSTVAK LYASCLAARITDWA TGGAVSRQKGFMS TEGCVENFTLQMAL DNARTRKQCAVA WLDISNAFGSVPHRH IFGLRELGLPDGVID LVRELYHGCTTTVRA TDGETAEPIRSGVRQ GCPLSPIIFNLAMEPL LRAVAGGPGGLDLY GQKSLVAYADDLVL LAPDATQLQQLMDV TSEARWMGLRFNV

267

AGGAGCGCGGGAGCTGGAGTCTCCCTGCAACCAAGCCCGCAGCGCGG ACCGGTACCGTGACGCCCCGACGAGACCTTCTTGAGAGGTTCTT GAAGGACCGCTCCGAAACAAGTACGCCGCGACCTGAGGGCCAAACC CGACAGGGCAAGTCTTCGACGTACCTCGAAGTGGGACTCCAGCAA CCACTTCATGCCAGCGGAGCTTACGCGCTTCGCGGACTGGCGCTTC CTCACCGCGCCGCTCAACTGCCTGCCTCTGAACGGGGCCGCTCGCTT CGCCACCGGACAAAGGTGCCACGGTGCGGTACGTGGCAGAGA CCCTCCCCACGTGTGCACTGCAAGCCGACGCCAGACCTGGCA GCTCTCCACAACGCTGCCAGGACCGCTGTGAGGGCCATCCCGGCC GCAGCGGGGAGATCTCGTGAAACCGCACCGTCCCGGGCTCGAGAGC CAGATGCGCCCCGACATCGTCAACCAACGAGGAGGCCAAGAAGTTC GTGATCGTGACGTGACCATCCCTTCGAGAACCGCGCCCAAGCCTTCA CCGACGCCCGGCTCGCAAGCGGGAGAGTACGCCCGCTGGCCGACA TCCTGAGGGGCCGCGGTACGACGTGACGGTCGACGCGCTCATCGTGG GAACGCTCGGAGCCTGGGACCCCAACGACGAGAGCGTCCTGCATGCT GCCGCTCTCCCGCGCTACGCCAAGCTGATGCGCTGCCTCATGTGTGTC CGACACCATCCGTTGGTCCGTGACATCTACGTGGAGCACATCACGGGC CACGCCAGTACACCGACCCACAGACGAAACCGCGCCGACCGGAC CCAGAGGGGACCGCTGAACCCCTCTGCACCAATGGACCTTCACT TCGAGAGGATTCAGCAATGGACACCCCGCTCCACCCGAAGAGGA CCCCGCGATGAGACTCTATGGACTGAGACACTTTTCTCGAACCCAC TTCTCCACCATTCGGACCATTTGACGGGTTGTGTATCTATCTTCT TTCTCTCAGCGTCGGAACCCCTCCCTCCCTCCCTCCCTCCCTCCCT CCCACCCCGGGCTTAGTTGGCTAACATTGTATCTCTGTAACTAGTTG CGTTCCCTCTCACCCCATCCCTCTATTGTTAGTCCCTCGCTCGGGCGC TCTGTAATTCCTACCGGCTTGTATCTTTTGGATTACAAATCCTAAA CATCTACTAATAAAGTCAATC (SEQ ID NO: 1058)	ACCCCGAC AGCGGC GGTGACC ACTGACCA TCGACCGA CCCCAGA GGCAGAG AGACTCTC AGAGCCCG GAACCCCG GCTGACGA GAGCGCC TCCCGCG GAGGACCC CGGAGCCT GAGGATG CCCCCGG ATGACGGC GGAGCGC CCCGAGCG ACAGCGG ACCCCTCC GGACCCCG ACGGCCCG TCGGTGAC GATGGCG GGCCCCGA ACGACGAC GACCCCG GACCCCG CGGTCCCG AGGACGCC CCCCCCGA GGGTCTCC CCACGCTG GTGGAGG AGCCCCGG ACCCCGCC	

269

270

GSCGACCTCGAAGCTGCGTCAAGAAAGCCTGCCCCCGTGCAGGCGCA	CTGCTGCT	ATCGTAT	SRPVSPSHVAKXIS
GCGGAACGGGATGCTGCGATCGAGAAAGTACCTGCTGCCCTCGGGGAAC	GCTCCACA	CCACTCA	MLRRLSAASPPVQH
ATCACCAGGTCTCTCGCAGCAGGAGGCCGCTCTACCCCTCTCATGTG	GAAGGTTT	TTATACT	VPVPRRISAPPRIAAR
CCAAGMAGATCTCATGCTGAGACGACTCAGTGTGCTCGCTCGCCACCTGT	CTCGGACA	CCACAC	DPVAGRASAAAPQTA
CCAGCAGTCCCCTGTCAGAAAGGATCAGCGCCCCACCGCGCATAGCT	(SEQ ID	CTGAAC	LRTPAAGGASTTPQT
GCTCGAGATCTGTGCGCGAAGAGCCAGCGCCGCCCTCAGACCGCCC	NO: 1182)	ATAGCC	ALRTPTAGGASAMP
TGCGAACTCCAGCCCGGAGGCCAGCACCGCCTCAGACCGCCCC		ATTATAT	QTTLPXPRRPDWRN
TGCGAACTCCAAACCGCGGAGGAGCCAGTGCCATGCCTCAGACCACTT		GAACAA	QPRSHKAPGLHRQT
GCCAGMCCCAGAGCTCAGACTGGAGGAACAGCCCCGAGCCACAG		CATACCC	DQHGPQVHSAGHCL
CAAGCACCGGCTTATCGCCAGACGACCGACGACGCGGCCCAAGTC		CCATATC	REISRSSNRLGSSHS
CATTCTGGGGACACTGCTACGGGAGATCTCACGCTCCAGCAGCAACC		TCAATGT	AAATHRRRTGGVPAT
GCCTAGGCAGCAGCCACTCATAGAAAGGACCGCGGTGTCCAGCAACC		CTGTACT	PEPDRVSPPTSNAVIP
CCGAGCCGACCGGTCTCTCCGACCAACCGAAGGAAATCTGACCCACGAG		TTGACCC	PEIPPQHPTEGNPD
CAGAGATCCCCCGCCAGCACCATACAGCAGGCTCWGAGCCTGCACACGAC		GTTAAC	RDRRQADHTAGSEP
ATAGACGCGAGGCCAGCATACAGCAGGCTCWGAGCCTGCACACGAC		CTTTTAC	APDEVEDXEGQRPM
GAGGTCGAGGACCMGTAGGGCCAGCGGCCGATGTTGAGGGCTGCCAC		CCCCAAT	VRAATPWQTAWTEE
NCCGTGGCAGACTGCTGACCGAGGAGCTACAAGCGGCAGCTTCCTT		CGGGGA	LQAAASDFDLDLVD
TGACGACTCGACCTCTCTGACAGGCTCACCCGAGAACTGTCTGCG		TATTGCA	RLTRELSAEIAPRRSS
GAAATCGCTCCAGGAGGAGTTGAAACAGGAGAACGCCCGCTGCTGCC		GATTAT	NQENAPPAHRTAP
CACAGAACGCTGTCTCCGAACCAACACCAACCAAGGGGAGCCAGA		GTAATCC	NHNTTTRGARSRA
AGTAGAGACGCCAGCCGCTACGATCCAGCAGCGGCTTCAAGGATC		TTACGCC	SRRYDPAASRIQKLY
CAAAAGCTGTACCGGGCAACCGCTCCAAGGCCATGAGGGAGATCCTA		ACCCGA	RANRSKAMREILDGP
GACGGCCCTCGCCCTACTGCACGATCCCATCTGAGCGTCTACAGCT		TCCTAAA	SPYCTIPSERLYSYFKD
ACTTCAAGGATGTATTGACCGCATAGCCCGGAATGACGGCAGCGCCC		CCGAAT	VFDRIARNDARPEC
AGAGTGCTCGCCCTGCCCCGTGTCGACGAAGCAGGTGTCCTGGAA		TTGCGAC	LRPLRVDEAGVLET
ACTGACTWTAGCCCCAAGGAAGTATGGCCAGACTCTCAAAAACAAA		CCCTTGA	DXTPEV/MARLSKTK
AACACAGCTCTGGGAAAGACGGCATCCCCTACAGCCTCTGAAAAGC		TAATCTG	NTAPGKDGPYSLLKK
GAGATCCCGCTGCTGCTCGCCACGCTTCAACCAAGTCAAGCG		TACCTTA	RDPGCLVLATLFNQ
ATTCTGCCGACTCCAGCTCTGGAAGAGGCCATGACGGTACTGGTG		TTCCCTG	KRFRTPSSWKKAM
TACAAGAGGGCAGCGGATGACCCAGCAACTGGAGGCCATCTCC		ATAACC	TVLVYKKGERRDPSN
CTCTGCTCCAGATGTACAAAGCTCTATGCCAGCTGCTGGCGTCGAGGA		AGAAAC	WRPISLCSTMYKLYA
TCACGGAGTGGTGGTGAAGCGGGGAGCCATCAGCTCCATCCAGAAAG		TTCTATG	SCLASRITEWSVSGG
GCTTCATGCTGCGAGGGCTGTACGAACACAACTTCGCTCTCCAAACC		CTTAAAC	AISSIQKGFMSCEGY
ACCATCGAAACGGCCAGAGGGCGGAGGCGAGTGCAGCGGTAGCGTG		TCTGTAC	EHNFVLQTTIETARR
GCTCGACCTGGTAACGCTTTGGGTCCATGCCCCACCAACCATCTTTG		CGTTTTT	ARRQCAVAWLDLAN
CCACGCTCCAGGAGTTTGGGATGCCAGAGAACTTCTCTGTTGATCCG		TTTTATT	AFGSMPPHHIFATLQ
AGAGGTGTACGAGGGATGACAGCACCACCATTCGCTCGGTGCAAGGGGA		TCAACAT	EFGMPENFLRVIREV
GACCGCGAGATCCCGATCCGAGCGGAGTTAAGCAGGGGCTGTCCCT		CATCTTA	YEGCSTTIRSVEGETA

272

[illegible]

CGGACAGAGACGCTGCAGCGCTCCCCACAGGTAACCTCATGCGGCC	GACCCGGA	GAACCT	SETPRAALPTTPGPPP
TGCCGTAGGAGCAACTAGCATGCTCCCCCAGGCTGCACGGCGAGATCC	SCACCTGC	CATACCC	QDPPEHRSTVRGTTT
AGCCAAAGGAGAGCTCCCGTAGCCCCCAGATCCACAGCCAGACCC	AAGAACCA	TCATATC	PQTVPAAPPAETTQ
GCCCCGGAGACCAACACCTCTCCAAGTTACCCAAAGAGCCTCTG	AAACATTA	TCAATGT	QEERRPRARVATPW
ACGCCAAAACCCATGCCCCACCGAGGACCCACAGCCGGATGCCG	GGAGCTG	CTGTACT	QSAWMEELAKAEDF
CCGAGAAAGAACAGAACATCCCCAGGCTTCCAAACAGACCGCGCC	GCTGAAGA	TTGACCC	ENFDLMDRLTAELS
CCGACAAAGCCAGCACCGTGCTTCCAGAACCCACTCCCTCCCGGAA	AACCCCCC	ATCAAC	AEITARREPEQAAR
GATCAGTGCTGCCTCGAGACACCGAGAGTCCCTCCCCACACACC	GTGGATAC	CTTTTAC	ATRRFPAPSRNNTAR
AGGACCCCGCTCAAGACCACTGAACACCGTCCACAGTCCGAGGG	CTCWGCA	CCCCAAT	EGRRGDVGRRYDPA
ACAAAGGCGCAAAACGTCGCCGACGAGCGAGGTCGCCACGCGTG	GGGAGAC	CGGGGA	AASRIQKLYRMNRTK
ACGACGAGGAGGACGACGGCCACGAGCGAGGTCGCCACGCGTG	CTGGSTCC	TATTGCA	AMREILDGTSSYCAI
GCAATCCGCTGGATGGAGGAGCTGGCAAGGCTGAGGACTTTGAGAA	AGMAGGA	GATTAT	QPERLYSYFKDVFH
CTTCGACACCTGATGGACAGACTGACTGCAGAACTGTCTGCGGAAAT	CAKCTCTT	GTATTCC	EAQTNLRRPECLSLP
ACGGCAGAAAGGAGGAAACCCAGGAAGCCGACGGGCCACTCGCAG	CGGGACCT	TCATGCC	RIDLTEDLERDFSPQE
ATTCCCTGCGCGAGCGTAACAAACCCGACGAGAAAGGACGAGAGG	CMCATCSA	ACCTGA	VQARLSRTKNTAPGK
GGACGTCGCGCGCTACGATCCGGCGGCTGCATCCCGTATTAGAAA	GGAGCAA	TCTTAAA	DGIRYPLLKKRDPGCL
CTATACAGGATGAACCGGACGAAAGCCATGAGGAGATCCTCGACGG	GAATATCT	CCAAAC	VLAIFNKCKQFHRV
ACCTCTCTACTGTGCCATCCAGCCGAGAGGCTCTACTCTACTTCAA	CAACAGCT	TTTGAC	PRSWKSMITVLHKK
GGATGTTTGTATCAGAGGGCCAGACCAACTTGGACGCCAGAGTG	CTTCAGGA	CCTCGAT	GXRDDPGNWRPISL
CCTTCCCGTACCCCGATCGACTCACGGAGGACTTGGAGCGAGAT	GGGGGAC	AATCTGT	CSTYKLYASCLAARIT
TTTTCCCGCAGGAGTGACGGCGAGGCTGTCGAGGACCAAAACACC	CCCCGGAG	ATGTTAT	DWSVCGGAVSSVQK
GCCCTGAAAAGATGGCATCCGCTACCCCTGCTGAAGAAGCGAGAC	AACCCTGC	TCCCTGA	GFMSCEGCYEHNELL
CCCGCTGCTTGGTGTGCTGCCATCTTCAACAAATGCAAGCAGTTCCA	CCGCTTCC	TAACCA	QTAIQEARRSKRQCA
TCGCGTCCCGCTCTGGAAAAGTCCATGACCGTGCTCATCCACAAA	CAGAACCA	GAAACT	VAWLDLTNAFGSIPH
AAAGCGAMCGAGACGACCCCGCAACTGGAGGCCCATCTCCCTCTGC	GGATGCTG	TCTATGC	HHIFATLGEFGMPET
TCCACCATCTACAAGCTGTATGCCAGTGCCTCGCGGCAAGGATCACAG	ATCGCCGC	TCAAAC	FIQLRDLYKDCCTTIR
ACTGGTCAGTGTGCGGGGGCGCGTCAGTCACTGAGAGGGTTTCA	CCCACCGG	CTGTTCA	ATDGETDAIPIRRGV
TGTCCTCGAGGGATGCTACGAGACAACTTCTCTCTTCAGACGGCCAT	GAAGGAT	CTATTTT	KQGCPLSPIIFNLAME
CCAGGAGGCCAGGAGTCAAGAGGCGAGTGCAGTAGATGGCTTG	GCCACCGC	TTTTAAC	PLIRAISSGPTGFDLH
ACCTGACCAACGCTTTGGGTCCATACCCACCATCATCTTTCGCCACC	AGGAGCCC	ATCATCT	GKKLSLAYADDLVLT
CTGGGAGAGTTCGGGATGCCAGAAACCTTATCCAGATCTCTCCGGGACC	CCCCA	TAATAA	ADDPESLQGMLDAT
TCTAAGGAGTGCACCAACCATCCGCGCCACGACGGAGAGACGG	(SEQ ID	AATTTT	SRATDWMGLRFNAK
ACGCCATCCCCATCCGCGCGGCTGAAACCGATGCCCTTTAGCCC	NO: 1183)	AAATCT	KCATLHIDGSKRDSV
CATCATCTTCAACCTGGCCATGGAACCGCTATCCGAGCCATCTCCAGCG		GTT (SEQ	QTTGFIQGEPIPL
GCCGACCGGCTTCGACCTGCAGGCAAGAACTCAGCATCTGCGCCTA		ID NO:	AEGQAYQHLGTPTG
CGCGGACGATCTGGTCTGACCGCGGATGACCCAGAGAGCCTCCAAG		1306)	FRVRQTPEDTQIEILQ
TATGCTAGATGCCACCAAGCGAGCTACTGACTGGATGGGCTCCGCTTC			DAAKIDASLLAPWQK
AATGCGAAGAGTGCGCAACTCTGCACATTGACGGCAGCAAAAGGGAC			INALNTFLIPRISFTLR

TCGGTGCAGACAACGGGGTTCCAGATCCAGGTTGAGCCCGTCATCCCC TGGCAGAGGGGCGGATACAGCACCTGGGCGACCCAAACAGGTTCC GTGTCCGGCAGACACCGAGGACACCATCCAGGAGATCTTGAGGACG CCGCCAAGATTGATGCTCCCTGCTGGACCGTGGCAGAGATAAACGC CCTGAACACCTTCTGATCCACGATCTGTTACCTAAGGGGATCCG CCGTGGCAAGTGCCCTCAACAGGAGAGACAGATCATCCGGAAGC TGGTGAAGAAAGTGCTTCTTCCCCAGAGAGCCAGCAACGAGCTGG TCTACATCGCCACAGGCGGGCGCCAACTGCCCGCATGGGTGA CCTGTGCAGCTCGGGTGATACCCACGCTTCCGCTGCTGACATGT CCGACGCCACGGTGAGGAACATTGCGGCGAACGCCCTGCGTGATGCG ACAGAGAAGCGGATCGGAGAGCCCCCTCGAACCAAGACATCGCCACC TTCCTGAGCGGCTCCTGTGATGGGAAATTCGGACGGGACGGGCGCGAC ATCGCTTCACTGTGTCGCCGACTCGCAACGCCACGCGTGCCTGGGGA AGCGCATCGGCTGCCGCTGGGAGTGGTGCGAGAGCGCCAGGAGCTG GGAATCGGGTGCCGCGAGATCAGGTCCGACGACCAACACCATCGTCACC CCGACGGCCAGGGGCTTGTGGAGAGGACTCTGAAGGCCGCCATCCG TCGCTGTACGTGGAACCTGAAGCGTAAACCGGACCCAGGGTAAAGCC TTTGAGTTGACCAGCAAGTGGGACGCCAGCAACCACTTCTCGACGGG GGCGGCTTACCCGTTTCCCGACTGGCGGTTTATCCACCGTGCCCGG TCAACTGCTCCGCTCAACGGAGCCGTCGCCACGGGAACCGAGACA AGCGTTGAGGAAGTGGGCTACCCCAACGAGACCTGCCACGCTCT GTGAGCTGCAACCCCACTCCAGAGCTGGCAGCTGCCACCAACGCC ATCCAGAACCGCTGTGTAAGCCATCGGCCACCGCTGGGGGAGATC TCCGTAACTGCACATCGCCGTACCGACAGCCAGCTACGACCTGACG TGGTGTACCGACGAGGCCAGAAAAGATCATCTCGTCGACGTAC GGTCTCTTTGAGAACAGACCCCGGCTTTCGGAAGCCCGAGCTCGT AAGCTGGAAGATACGCCCTGGCTGACACCTGAGAGCGAAGGGC TATGAGGTGCAGTGGACGCCCTGATTGCGGAGCCCTGGCGCTGG GACCCCTGCAACGAGCGTGTGTCGGACCTGCGGGATCGGTGACGC TAGGCACTCTATCGCGGCTCATGTTCTCAGACGCCATCCGATGGT CCAGGGACATCTACATCGAGACATACCGGCCACCGACAGTACAGGA GGCGTGAGCAGAGTGCATCGTTCTCCCACTACGAGAAAGGACCAA GTGACCTTCTCGTTGGATCATATGAACCTGGAACCATAACTCCCTGAAC ATTAAATCTACCAAAATGAGGGTCAATCCATCTCATCATCATCACT CATTATAMTCCACACCGAACACAGCCACTCTATGAACCTCATACCTCA TATCTCAATGTCTGACTTTGACCCATCAACCTTTTACCCCAATCGGG ATATTGCAGATTATGTATCTCATGCCACCTGATCTTAAACCAAACTTG CACCTCGATAATCTGTATGTTATTCCTGATGAACCAAGAACTCTATGCT			GSAAKVPKADKII RKLKWLFLPQRAS NELVIAHRHGGANV PRMGDLCDVAVITH AFRLTCDPDAVRNIA ANALDATEKRIGRA PSNQDIATFLSGSLD GEFRDGRDIASLWS RTRNATRLGKRIGC RWEWCEERQELGIR VPQIRSDDNITVTPTA RGLLERTLKAIRSLY VETLKRKPDQKAFE LTSKWDASNHFIDG GGFTRFADWRFIHRA RLNCVPLNGAVRHG NRDKRCRKCYPNET LPHVLCSCKPHSRAW QLRHNAIQNRVLKAI APRLGEISVNCTIAGT DSQLRPDVVVTDIA QKIIILVDVTVSFENR TPAFREARARKLEY APLADTLRAKGYEVQ MDALIVGALGAWDP CNERVLRCTGIGRRY ARLMRRLMVSDAIR WSRDIYIEHITGHRQ YQEA (SEQ ID NO: 1428)

NeS L	Utopia - 1_DYa k	.	Drosop hila yakuba	CAAACTCTGTTCACTATTTTTTAAACATCACTTAATAAAATTTTAAATC TGTT (SEQ ID NO: 1060)	AAAGTGTA GTTCTTTTC TGTTTTAG TGAGTGG GAAGTCTG TTTCTTTT ATTATGTT TTTTACGA AAAAAGTCC TGGTCTTT GAAATTCA TTGTCTAA ATTTTAAA TAAAAATTA TAAAAATTT AAAAAGA AAATTAAT TAAAGAAG CGATGAAA TATCTCTG AAATTCAT TCAATCAA TTAATCAT GGCGTCTC AGCGAGT GCACGTAT TTGCCTAC CCCTTCGT GGGACCAT TCCGGTGC TCCGTATG CATGGATG CGTCCGGG ATGTCATC CACTAGGT CGCTGGGC	TAAAAA ATTA ATGCCTT AAAAAT AAATAA ATATATC AAAAATTT AAAAAA AAAAAC GAGGAA CAATA AACACA AATCTG AAAGAT TTATATA ATTTAAA AWATAA ATCGAA AATAA TGTTGA AAACAA AAAAA ATAAT ATAAT ATAA AACACA ATAACA CTCACCC GGCTG CCCCAG AGGCAG GTAAC ATTACT GGCCAT ATGGCT TTTTTT	YAPGYEAAQSPCGRE PPRDHRRPRDACC SSHSEPCLTTPRLP ETVSAEPCDDSQRT RYASPHKQARTLHDA EPRDSREHAPSCAE PRCHRCQWTHWKD CCPHSTNTTDGPEGT DRCADTITSPATAAC PQRSPCLGSSNGCD ETAPEKRQPAADLVH TAPFAVLVRAGPFAD LVRAGPFADHHQDD DPLPHRSGSLGPLCSK QKDPKTHQHRHSG QAGNQTHTDIPRAA PSRRAAICLMAAA TREDLLRAATLSSEM AAANQPTRSPTGGG EPTSQRRGPQALA DAAKRIQQIYRTNIPR AMRKVLRTLLTAVFS ACLRTHVDPDLCKKS RTVLHKKGDRDLS NWRPLSMGDTIPKLF AAVMADRLTAFLTN GGRLSEKQKGLQHE GCHEHNFVLGQVLEE SRRQKDLVMGWL DLSNAFGSIPHATIM DAVAGMGIPSRITII HQLATGAATTAKTID GMSEIPIEAGVRQG CPASPLFNIAIERVLR KIKTVNAGYLLYSRI SPLAYADDLVLIASSP

277

NeS L	Utopia -1_Gav	.	Gavialis gangeti cus	CCCTGCCAAGTGGAGCTCCTCCAACCTAAAAAATTAATAATGCCTTAAAA ATAAATAATATATCAAAATTTAAAAAAGAGGAAACAAATAAAC ACAAATCTGAAAGATTATATATAATTTAAAWATAAATCGAAAAATAAT GTTGAAAAAAGAAAAAATAATAATAATAAAAAACAAATAACACT CACCGGCTGCCAGAGGAGGTAAACATTTACTGGCCATATGGCTT TTTTTTAA (SEQ ID NO: 1061)	CGCTGGAA AGACGGA GAACGCT TCTTTTCC TGCGCCG GCCTGGTA TTGCACCT CCTCCAGG ACCAAGCG CAACCTAG TCCGGCAG ACTGCCG AATAATAG CCTCAGAA AGAGAGCT GGCTAGCA GCCCTCTT TTCTTTCT CCGGTGCA GCGTGGG TTCTTGTC AGTCTGA TGGGCTAG GGAAGGC GGTGCCG CAGTACGT CCGAAAGA GCGCCGGT TGCGCGAG CGACCGCG CCGCTCAG GCGAGTA	TGAACC GCCCC CCTCGC GCCAGA CGGACC TTCACTT CACTCC GAGAGG ATTCTC GACCAC GGACGA CCCCGT CCACC GAAGAG GACCCC GCGAT GAGACT CTATAC GGACTG AGGCAC TTCCTC GAACCA CTTCTC CACCAT GCGGAC CAITGTA ACGGGT TTGTGT GTATCTA TCTCCT TCTCTC CAGCGT	MSGPRQAAADPRPS TDPRQRDSQSEPR LTRAASRRRTDPED APRTTAEHPERRTP PDRGPSATTAGPER RRPPDPGGPEDDPE GLPTLVEEPRTPPTD PPDGRPRRGCRGS AHVPPLPPCEAAVP DLPPAKAVQVAQRH EQPTALPPAAPSVLL LPLHRVRGPEAPEE PPQGMPPRGREET RHAGEVRRPTTAAAA RRPARPAAPPATPD QTSGRPTERPAPAT PPRSAPRDPDPVT PRDGGPPGPPGP DAPDPRIPEPPGP EPGALQLPSVPGSP GAETSAQQRMPTR QALWLELSRATAFE AFEASVARLTELAA ARPGQPRRGADNGP TTRDRPQQRPP RRQRYDPAASRIQK LYRANRPKAAREILEG PSAFCQVPRETLFNYF SRVFNPAEAAAPRP ATVEALTPVPAEGF EEAFTPREVEARLKRT RDTAPGRDGIRYGLL

AGGCCTCGGTGGCCGGCTACGGAGGAGCTCTCGGCGGCCGCCCGGCGC	GCCCAAGG	CGCGAA	KKRDPGCLVSLFN
CCGGCAGCCCCGAGGGGCGCGACAAACGACGACGACGACGACGACGACG	GTCTTACG	CCCCCTC	RCREFRTPAAWKR
GACCACAGCCGAGCCGAGAGGCGACCCAGGCGCGCGCGCTACGAC	GTTCCGCCG	CCCCACC	AMTVLIHKKGDPTDP
CCGGCGAGCCTCCGGATCCAGAAGCTGTACCGGGCCAAACCGCCCC	GACCCGAT	CCCCACC	GNWRPIALCSTVAKL
AAGGGCGAGAGAGATCTGGAGGACCTCGGCTTTCTGCCAGTCTC	AACGCGAA	CCCGGG	YASCLAARITDWAUT
CCCCGGAGACTCTGTTAACTATTTAGCAGGCTCTTCAACCCCCCGGC	AGCCCCGA	CTTAGTT	GGAVSRQKGFMT
AGAAGCCGCCGCCACGCCCGGACCGTGAAGCGCTGACCCCCCGTC	CTCGGCC	GGCTAA	EGCYHNFTLQMAL
CCCCGGCAGAGGGTTCGAGGAGGCTTACGCCGCGGGAAGTGA	AGTAGCCG	CATTGTA	DNARTRKQCAVA
AGCCGCTGAAGAGGACGAGGACACGCCGCCCGGAGGACGCGCAT	AAGACNT	TCTCTG	WLDISNAFGSVPHRR
CAGGTAGGTTCTCTNAAGAAAGTGAACCGGCTGCTCGTCTCTTCT	GGGCTCC	TAACTA	IFGTLRELGLPDGVID
GTCTCTTCAACAGGTGCAGAGAGTTCGGGCGCACGCCCGCCCTGGA	CTCCCCAG	GTCGCG	LVRELYHGCTTTVRA
AGAGGGCATGACGGTCTCTATCCACAAGAGGAGACCCGACCGAC	GTCGGAGT	TTCCCT	TDGETAEPIRSGVRQ
CGGGCAACTGGAGACCCATCGCCTGTGCTCCACCGTGGCCAAAGCTGTA	AGCGGAA	CCTCACC	GCPLSPIIFNLAMEPL
CGCAGTGCCTGGCGGCCGATCACCGACTGGCGGTGACCGGCGG	CGCCCGTG	CCCATCC	LRAVAGGPGGLDLY
GGCGTCAGCCGGAGCCAGAAGGGCTTCATGTCGACGGAGGGCTGCTA	CTCGGAGG	CTCTATT	GQKLSLAYADDLVL
CGAACAACTTACCCCTCAGATGGCCTGGACAACGCCCGGAGGACC	ACGGAAC	GTTAGT	LAPDATQLQQMLDV
AGGAAGCAGTGGCGGTGGCTGGCTGGACATCTCCAACGCTTCGGC	GTGGACAA	CCCTCGC	TSEARWMGLRFNV
TCCGTCCCCACCGCGCATCTTCGGCACCTCCGCGAGCTGGGCTAC	AACACCCC	TCGGGC	AKCASLHIDGRQKSR
CGGACGGCGTCATGACCTGGTGCAGAGCTCTACCAAGNITGACCA	CAGGTCCC	GATCTG	VLDSTLIQGGQAMRH
CGACCGTCGCGCCACCGAGAGAGACCGCGGAGATCCCCATCCGGT	AATGACGC	TATTTCC	LRDGEAYCHLGTPTG
CGGGGTGAGGACGGGTGCCCTCAGCCCCATCATCTTCAACCTGGC	CCTGATCC	CTATCG	HRAKQTPEETINGIV
CATGGAACCGCTCTTCGAGCCGTGGCGGGCGGCCCGGGGCTCGA	ACTGACAA	GCTTTGT	QDAHKLDSSLAPW
CCTGACGGCCAGAGTTGAGCGTCTGGCTACGCCGACGACCTCGTC	GAACGCTC	CATCTT	QKIDAAANTFLIPRVAF
CTCCTGCCCCGACGCCACCCAGCTGCAGCAGATGCTGGACGTGACGT	GAGGCAC	TTTCTGG	VLRSAPVPTPLKKA
CCGAGGGCCAGGTGGATGGCTCGCTTCAACGTGCCCAAGTGCG	NCCAGGA	ATTCGG	DAEIRLLKKWLHLPL
CCTCCTGCACATCGAGCGGAGGAGAGAGCCGCTCTGGACTCCAC	GACCCCCA	ATCCTAA	RASNEVLHIPYRQGG
CCTACGATCCAGGGCCAGCGATGAGGACCTGCGGACGGCGAGGC	GCTAGGGC	ACATTTA	ANVPRMGDLCDIAV
CTACTGCCACCTGGGACGCCACCGGCCACCGGGCCAAAGCAGCGCC	AGACCGCC	CTAATA	VTHAFRLITCPDATV
GGAGGAGACCATCAACGGGATCTGAGGAGCGGCCCAAGCTGGACTC	GACCACGG	AAAGTC	SIIAASALEETARKRIA
GTCCCTGTGCCCCCTGGCAGAAGATAGACGGGCGCAACACTTCCTC	GTGCGGGA	AATCTGT	RQPTGRDLATFLSGS
ATCCCCCGCTGCGTTCTGCTGAGAGGCTCGGCGGTCCCCAAGACCC	GGACCTC	TCITT	LEGEFRDGGDFASL
CCTCAAGAAAGCGGACGCCGAGATCCGGCGGTGCTCAAGAAAGTGGC	CCAGGAG	(SEQ ID	WSRARNATRLRKRI
TGCACCTGCCGCTGAGGGCCAGCAACGAGGTCTGACATCCCTACCG	GGTGAGCC	NO:	GCAWTWTEECRELG
GCAGGGAGGCGCAACGTCCTCCCGCATGGGAGACCTCTGCGACATCGC	AGCGAACC	1308)	VSLQAPHAADRVTVT
GGTGTCAACCCAGCCTTCGCTCTGACCTGCCCGGACCGCGACGCGTA	CGAGTCGG		PRTRFLERFLKDAVR
AGTATCATCGCCGCCAGCGCCTCGAGGAGACCGGCCGCAAGAGGATC	CGACGAAC		NKYAGDLRAKPDQG
GCGAGGCAGCCGACCGGACGNGACTTGGCCACCTTCTCAGCGGCTCG	CCCGACGC		KVFDVTSKWDSNH
CTGGAGGGCGAGTTCGGCCGAGACGGCGGGGACTTTGCCTCGCTGG	ACCCCCC		FMPSGSFRFADWR

NeS L	Utopia -1_LV	AGCV0 13581 06	Lytech i nus variegat us	AGCCGAGCCGCAACGCCACGCGCGCTCGGGAAGCGCATCGGCTGC GCCTGACCTGACCGAGGAGTCCGCGGAGCTGGGAGTCTCCCTGCAA CCAGCCCCGACGCCGACCGCTACCGTGACGCCCCGACGAGGACCT TCCTGAGAGGTTCTGAAGGACCGCTCCGAAACAAGTACGCCGCGG ACCTGAGGGCCAAACCCGACGAGGCAAGTCTTCGACGTACCTCGA AGTGGACGCTAGCAACCTTATGCCAGCGGGAGCTTACGCGCTT CGGGAAGTGGCTTCTCCACGCGCGCTCAACTGCCTGCCTCTGA ACGGGCGGTGCGCTTCGGCCACGGGACAAAGAGTGCCGACGCTGC GGTAGCGGACAGACCTCCCCACGTGCTGTGACGTGCAAGCCG CAGCCAGAGCTGACGCTCCGCCAACAACGCTGTCCAGGACCGCTG GTAGGGCCATCCGCGCGCGGGGGAGATCTCCGTGAACCGCAC GTCCGGGCTCGAGAGCCAGATGCCGCCGACATAGTCAATCAACAAC GAAGAGCCAAAGAGTGTGATCGTGACGTACCATCCCTTCGAG AACGGGCGCAAGCCTTACCGACGCCGCGCTCGCAAGCGGAGAA TAGCCCCGCTGGCGACACCTGAGGGCGCGGCTACGACGTGACG GTCGACGCGCTCATCGTGGGAACGCTGGAGCGCTGGACCCGACCAAC GAGAGCGTCTGCGTGCCTGCCGCTCTCCGCGCTACGCCAAGCTGA TGCGTGCCTCATGGTCCGACACCATCCGTTGGTCCCGGACATCTAC GTGGAACATACGGGCCACCGCAGTACTCCGACCCACCCAGACGA GCCGCCGCGGACCGACCGGAGGGACCGCTGAACCGCCCCCT CCGCGCAGAGGACCTTCACTTCACTCCGAGAGGATCTTCGACACG GACGACCCGCTCCACCGAAGAGACCCCGCGATGAGACTCTATACG GACTGAGCACTTCTCGAACCATCTCTCCACCATTCGGACCATGT AACGGTTTGTGTATCTCTCTCTCTCAGCGTCGCGAACCC CTCCCCACCCCGGCTTGTGGTAACTGATCTCTG TAACCTAGTCGCTTCCCTCTCAACCCCTCTATTTGATGCTCT CGCTCGGGGATCTGTATTCCTATCGGCTTGTCTCTTTCTGGAT TCCCGATCCTAAACATTTACTAATAAAGTCAATCTGTTCTTT (SEQ ID NO: 1062) ATCTACTATCATGTCTGTCCAAGAGAGGGAAAGCGATCACCTCGTCTCT GATCTGAGACACCGCTCCATCAGGGTTCTGACATCGGGTTACCA GTTCTGCTTCTGAGGTTCCCGCGGAAAGAGTCTCGCAACCAAGCTC CCGACACCAAGTTCCTGCCAGCGAGGCTTCGCCACCGCCAGCAGACT GCCGCCAACGAGTGCCAGGTGTGGATCTCTTCGCCACCTCCAGTG GACTCCGCGCCACATGGCAGGCTTCTCGAGTGCCTCTGCGGATCC TGAGGGTGTGCGCGGCTTCCATCACAGAGATTTTCGACTACCCCTTG CCTTCCCGGTGAAATGCTCGGCATGCTCGGAGAACTTTTCAACCAGC AGACCTCAAGCGACACCGACGAGGATCATCCAGCTACCACTTCGC	CGCG (SEQ ID NO: 1185)	FLHRARLNCLPLNGA VRFGHRDKRCRRCG YAAETLPHVLCCKP HARAWQLRHNAVQ DRLVRAIPAAAGEISV NRTVPGCESQMRPD IVITNEEAKKVIVDV TIPFENRRQAFTDAR ARKREKYAPLADTLR GRGYDVTVDALIVGT LGAWDPNESVLRA CRVSRRYAKLMRCL MVSDTIRWSRDYVE HITGHRQYSDPTRRA AAGPDEGTA (SEQ ID NO: 1430) TGAATA GCATTTA TATTGTG TTCCAAA CAACAT ACTCAT ATTATAT CTAAAC ATTTTT	MSPREGSDHLGPD PETPALHQGSDIRVT SSRLRGSRGKSSRQP SSRHQVPASEASATA QQTAA NECQVCGSS FATSSGLRRHMARLH RAASADPEGAAPASI TEIFDYPLPSRWKCSA CSENFENQQLKRH

GTATGCTTTCCGGTGTTCGTATGCGCGGTCCGAGTTCGACTCAGCACGG AGGGCTCGAAACCATTTGGCAGGTCCAAAGAGGAGCGATCTCAACTC TCTGGCACCAGCCCCAGGCTCTCCAAAGCCAGAGTTAGCATGGCTC ATTCTCTCACCTCTGCCAACACTTTTGGCGGAGCTCGCTCGAAT CCTGCCGAGATACCTTCTTCTGCTGGAGTCTCTCCCAAGAACCGCC CTCGTTGAGGAGTTCGGTTCGTCTGCGCAACTGATGTTACGATGATG TCTCAAAGTCTCCACCGCAGGTACAGTCTCTCTCCAGGGGTGCAA TCCTTTTACCCGCTGCCACTGCTCCAGTCTCTCCAGGGGTGCAA GGCAGCTGACCCCTCTACACAGACTAACACCCAGTACCCAGAGGC TCGCTGCAACCTGAAGCAGACGCTGTAACCTGAACTCCCTCTCAGTCA CCGAGCACCTGTGTGACGCTCAACACTGGGTGATGCTGTATCCTCT GCATCAGATTGGTCTGAGTTGAAGCAGTATGTGATCAA TTTGTCTATCCA CGCTGTGCTGTTCCCGTCCCAATCTTGCTCGACCCAGCAGCAAGATA GGCAGAGATCTGGTGACCACTCTCTAGACAGCAAGAGGTCAGCATC GACCAACCTTCGATGTCGTGAGGCAAGTAGAATCCAGAACTCTATCG TACCAGCAAGAAAGAGCCATCAGACACATACTGAAAGAGAGAAATCACC TTCCTTCTGTTCCGAGTCAGACGCTTAGACTTCTTCCCGAGGTGT ATTCTGTAAGAAAGTTGACGAGGAAGCAGTTGGTAACTAGCATCCTC GCTCTCGATGTCCTCAAGGTGATGACTCTGCGACATCTCTGTCTGTC CCAGCTCAGCAAGGAGATCGGAGCAAGGCTGCAAGGATGACAAACT CTGCCCCGGGAAGGATCGCTTGAGTACAGACACATTCGACGTGCGG ACGGTCTTCAGCATCTGAGGCCATCTTAACAAATGCCTGGCTGA AGGTCGGATCCAGCTCTTGAAGACAGCATCTACCATCTACTTCACA AGGCTGCCCCACGGATGATCCCGCCAACTTCGCCCAATCGCCTTACA GTATGCTCTACAAGCTTTTATGGCTGACTTGGGACCGGCTGACCA AGTGGCCTGTGAACACAGTACCTCAGCCCCGAGCAGAAAGTCCGCTC GCCCTGCGAGGGGTCTTCGAGCACTCTTCTCTCTCAGCTGCCCTG AAGGACTGCAAGGAAACCAAGACCATCTGCAATCGGTTGGTTGGAC CTTAGGAATGCAATTTGAAAGCATCTCTCATCTCTGATCAAGATCGTCT GTCAGTCTGGGTCTCCTGATTCGTTGTACCTCTCTCATGATGCTCT ACAAATGGTGCCTCAACCTCTGACGCTGACCGGGGGCCAGACCGACA CCGTACCCATCAGATCAGGGGTGAAGCAAGGCTGCCCGATGTCCCCAAT CCTCTCAACCTGGCCATCGAACTTATCATCAGGGCAGTCAAGAAAGAA GCATCAGACAACCATCTCGGAGTGACTGTCCAGGGCAAGAACCTCTCCA TCCTGGCCTATGCTGATGACCTAGTGTGCTCAGCCGAGACACTGAAGG CCTCCAATCCCTCTCAAGTTGTGGTCTTCTGATCTACCTCTCAGAT GCAGTTTAAGCCCCAGAGGTGCAACACTCACCTTGACTGCAAGCGT GGTACCAATGTTAGGCAGTCTGCTCACCATATCAAGGGGCTGCCATCC					TTTCTGT TCCTGAC AATCTAC GTAAAG TCTGCTA ACCAAC TGGCAT GATGAA ATAAGA TAAAT CCCTTA CACATTA ATTTCTT GTCACA TCATAAT GCTTTGT CAAAGC AATGTC CTACATA ATATCTC GATGTC ACCCCA ATTAATT TTACATC CTTCGG TAACCTT TATACC GTTGGA TCAACAT ATATGA TTTGTAA AACTGTT ATTTCTG AGTTTTT TCTATGC TAATAA A (SEQ	QTRHHPATTFAYAFR CSSCRSEFDSARRAA NHWQVHKHRSQLS GTEPQASSQARVSM AHSPPLPNTSWAEL ASNPAEIPSFVWESP PKNRPSVEEFGSSLPT DVTMMSQSPPPVQV QSSVPALPLSPAAT ASSPPGAARQLTPP TQNTTPVTQARLQ PEADVPELPPSVTE HPVSDAQHWVDAV SSASDWSEFEAVCD QFVIHAVAVSRPNLA RPQQQDRQRSGDH PPRQQRGQHRPTFD VREASRIQKLYRTSK RAIRHILKEKSPFSGS ESDVLDFREVVSAAKE VDEEAVGKLASSLFD VPQGGDSATSLSLPT SAKEIGARLSRMTNS APGKDRLEYRHIRRA DGSFSEIAIFNKCLA EGRIAPAPWKTASTILL HKAGPTDDPANFRPI ALQSCLYKLFMAVLA DRLTKWACENQYLSP EQKSARPCGECFEHS FLLSAAKDCRRNQK TICIGWDLRINAFESI PHPVIKIVLSSSLGVPD SLVTLLMDAYNGAST SFTLTGGQTDTPIRS GVKQGCMPSPILFNL AIELIIRAVKKNASDN HLGVTVQGGKNLSILA

				CATCGCTTACTGAAGAGGAGCGTTACCGCTACCTTGGGGTCCCCATTGG ACTCCAAAGGCTCACAAGTCTACAGGAGAGCTCTAGGAAGCTGTGCTCA GACATCGAGACCATTTTCATCTCCTCTAGCGCTTGGCAGAAACTGGA TGCAATCAAGACCTTTGTGATACCATCTTCACTATACCTTGGGGGCA CAGAATACCTCAAGTCGACTGAAACCCCTACGAGCAGCCATTATCAA GCAGTGAAGAAGATCTGCCATCTCCGGTAAGATCATCAACGCCTTC GTCTCGCTCCGACCTCCGGTGGTTGGCTTTTGTGACCCAGGAGT GGATGCCGACATCTTGTGTAACCAAGCCGTCCTGACCTGGCATCG GATGATGACACGGTGAGACCGTGTCTAGGACAGCTGACCTCAGTG GTCCACAGGACGGTCCATTTGACCATCAGATGATGATGATGATGATG TCTTGTCCGGAAGCAGTGAGGGCCCTCTGGCAAACTCAGGGAACAGTG GCCAGGCATCATCTTTTGTGTCAGTACAAAGGGCTGCATCCCGTCGCT GAAGATAAGGATCGTCGGAGCCCTCGAGCGGCGACATCAAGGTCGAGA GTGGAGCGCTGCCATCCCTCCAAAGAAAGTGACTGCAAGTCTCCGCTC CGACCATCAACAGGATGTCAGAGAAAGTGGGTCCTTCCAGACCCAG GGCAAGGTAGCCAGGGCACTCTCTTGGATAGTTTTCGCAACGCAACAT CATGGTGACATCCGGGAGTTTCATCCGATTTTGGACTGGAGGTTTCAT CCACCGTCTCGCTAAACTGTCTACCTACCAACGCTGTCTCCGGCGGT GGAACAGAAACGCAACCAAGTCCGAAAGGTGTGATCACCAGTTGG AGACGTTGCCCCACATTATTAACAACGCTCCCAACATGATCACCAGT CGCCGAGGCACAACTCCATCAAGAACGCTGTTGAAGGCGATTAC TACGGTGACATCTACAGGACCATGTCTCCCGGCGACCCGAACCCAA GGAGCGACCCGACATCACAGTGTGAGGGTAACAAGGTGACCATTG TCGACATCACCATCCCATTCGACAAATGTCGGATGCCCTCCACGGCT GCCAACGCCAAAGTCATGAATATGACACATCGCAGCAAGAGCTGGCC AGCCGTGGTATGGATGTCGAGGTCCAGCTTTTGTATCGGCTCCCTTG GGTGTGGCAGGGGACACGAGCGTGTCTGGGACGTCTAGGTATCT CTCGCCGATACAGAACTGATGAGGCGCTGTGCTGCTAGATGCGAT CAAAGGCGCGGGATATCTATCGAACATGTTACCGGTACCCGGCAG TATTGAATAGCATTTATTTGTGTTCCAAACAACATACTATTATATC TAAACATTTTTTTTCTGCTGACAATCTACGTAAGTCTGTAACCAA CTGGCATGATGAAATAAGATAAAATCCCTTACACATTAATTTCTTGCA CATCATAATGCTTTGCAAGCAATGCTCATAATAATCTCGATGTCAC CCCAATTAATTTTACATCTCGGTAACCTTTATACCGTTGGATCAACATA TATGATTTGTAAACGTTATTTCTGAGTTTTTCTATGCTAATAAAA (SEQ ID NO: 1063)		ID NO: 1309)	YADDLVLLSRDTGL QSLQVAGSSASTLQ MQFKPQKCATLTLDC KRGTVNRQSAHHIQ GAAIPSLTEERYRL GVPIGLPRLTSLQESS RKLSDDIETISSLLAP WQKLDIAIKTFVPIILQ YTLRATEYKSDLKPL RAAIKHVKKICHLPV RSSNAFVFASRPSGG LAFVDPGVDADILVW TQAVRTLASDDDTVR AVALGQLTSWHRV HSAPSDDDCIDKFLSGS SEGPLANSGNSGQAS SLWSRTRAASRLKIR IVGASSGDIKVESGGR AIPSKKVTAGLRSDH HNEMSEKLSLPDQ GKVARALSDSFANA TSWLTSGSFIRFCDW RFIHRARLNCPLTNA AVRRWKQNANTKCR RCDHQLTLPHIINNC RPNMVPIRRRHNSIQ ERLVKAIHYGDIYQD QHVPDPNPRERPD TVVEGNKVTIVDITIP FDNGPDALSTAANA KVMKYDTLRQELASR GMDVEVHAFVIGSL GSWHGDNERNVLRGL GISRRYRTLMRRLCCI DAIKGSRDYIEHVTG HRQY (SEQ ID NO: 1431)

NeS L	Utopia - 1_NVit	.	Nasonia vitripen nis	CCATTCCTTCGTACGGGTTTTCGTGCCGGCATAGCCGGGTGGGAGACTC GCGCGGGGAGGTTCATATCTCACCACCATCTCGAGCTTTTGTCTGCAC CTGATCCGGMTAACTGCCGAGAAGTTAAATGTATTACTGTTAATGTA CTCATGGAGCAACACACACACAAAGCGATCCGGAGGGGACTTC ATCGTATACTCTCCCGAAAGTACGACGAGACACTTTGTGTCCACT CTGGTGGACGGGACGTACTCCGCAAGACGCGGTTGAGTGTCT CAATAGACATGTCAAAGAGGTGCACCCAGACCTAACACACACTTTAGG TGCTGGGGTGCGGTTCCGGCTCCCGGGATAAGAAATATCCCGG AAGATCGTACCCAGACTGCGCCACCTGCGTCCCGAGGTTTCGTCCG CGCCTCGGCAGAGTGACGGAGAGAGGAGTTAACACCGCCGTC GGTTGGTATTCCCGGCAACCGAGGCCAGTCCCTGAGACGTACTC GTCGTAA CGCCTGGCTCACCTCCAGTCCGAGCAGAA CATCAGCCAAAG CGCGGCTCTCCGGAACCGCAAGGGTGCCACAAACATCCGAGATAGT TGCGCTGGAGAGTCAGCGGATGACGAAGTATTTCGATCTCCGGTGAA CTCTCCCCAAGAGACTGAGGGCTGCGGCGCCCAACAGCGGCTTC CTCCTCTCTCSGGGTTGCCGGTATCACTGCTGCAACCCGTCACAAC CCACGAGGACAGGGAATGGCTCGGCANACTCCATCTAGCCGAGCACC CGATACCTGCACCTCCGCAACCAACACACGAGGCGGAATGGTCGTGC CGACATACCCGTTCTGGCGGCTCCGCGCCCGGTCNACAAGCCGCC CGTAGAAGGGTCCACAGACAAACACGTCGCTTGTGGGACTCTC GAGCGGCCACTGTCGGGAGCAGCTTGGAGGCAACGACTCAAGC TATGCTATTCCTGGCCAGATTGACCGGGCGGCTGAACCCCGAAC GCCATCGCGCGGAMCGCGCAACGTCATCCAGCTCAGGGGGATGTC CAAGCGAGGCGCTGATCGGATTCASGAGGCCAAGAACTCCAGCGT CTTACCGAACATCCAAAGCGGGCAGTMCAGAAAGATTTGGCAGGT CCTNATGAAGTCCAAATCGACAAGATACCATCAGCCCACTTTGT GGAGTGGCTGCCCGCAGAGCGGTGGAGGATTGGCCCGATGCTT CGATCGGGAAGAGCCACGCGCGGCGAGCGGCGGCTTTGTACCCC TATCACAAGGGAGGAAGTTTCGACGGCTCAAGGGGAGGAACAACAC CTCACCTGGCCCGACGGAATCACCTACAGGACCTCGCCAAAGCCSAT CCAGGGCCCATGTGCTCGCCGCTCTATAACASCATCTGGCGCATTG AGCGACGCTGCTTGTGGGGGTGTCCAAACACMACCTKATATACA AGAAAGCGACGCCATGACATCTCCAACTGGAGACCGATCTCCTTGG GAGATACGGTCCCCAAGCTATTGCGCGGATCTGCGCGATAGAATCAA AAGTGGGCGGTGGCCAAAGCGGCTACAGTCCGCTCCAGAGGGCTT CTTGAGTTCGAGGGATGCTATGAACATAAATTCGTGCTCCAGGAGGCC ATTCGGGAGGCAAAAGGGGAGAAAGGAGTTGGTGGTGGCTGGCT GGACTTAGCCAGTGCCTTACATCCGTGCCCCCACTCCTCTATCCTGCAGG	CCATTCCT TCGTACGG GTTTTCTG GCCGGCAT AGCCGGGT GGGAGAC TCGCGCGG GGGAGGT CATATCTC ACCACCAT CCTCGAGC TTTTTGCT GCACCTGA (SEQ ID NO: 1187)	TAGTGG GGCCAT AACACC TAGGCC CCACAG TGTGGC GATGTC CATGTG TGTCGT CCTTACT TATTTAT TTATTTG TCCTGTC GAGCTG TTTAATC TATTCAT ATGTAT GTGTTG TGTGTG TCGTCC TCGTAG CAGTTT ATTCGT CCAACC AGAGGT CGACCA ATATTAA AATAAG CATGGC TTGAAG CAGGCC AAGCGC CGTGT CTAACCC CGTTTTA GGGAA GTTACTT AACCTA	SGXTGREVKCITNV LMEQPHTKAIREG DFIVLLPQSDDELIC CPLCVGRGRYSKTR VECLNRHVKEVHPDL TTTFCWGCFFAAP GDKYPRKIVTQHCA TCVPEVSSAPSGRVD GERRVNTRRRLGIAA ATEASPVRRTRRNL ASPPVEQNISQSAAP PEPARVPQHPEIVAL GESADDEVFRSPVNS PPRDWRAAAPQQA ASSPXAVPGITAATP SNTTRTNGSAXSILA EHIIPAPPPTNTEA NGRADIPRSGRAPPP GXQAARRRAPTTTEQ RRIVGLLEAATGREQL EATTOAMLFLARLT GRPEPRNAIRPGXR QRHPAQGDVQAQA PDRIEAKKLRLYRT SKRAVQKILAGPXM NCQIDKNTTIAHFVEL AARRDGGEDWPDVF DREPTAASGEALCT PITREEVFRRLKGRNN TSPGPDGITYRDLAK AXPGAHLAALYNXI WRIEATPALWGVSN TTLIYKGDAMDIN WRPISLGDTPKLF AILADRIKRWAVANG RYSASQKGFLEFGC YEHNFLQEAIREAK GGRKELVVAWLDLA

CGTCGAGGGACATGGCCTACCAAGCAAAGCGAGAAACATTATCTCCTC GCTCTACACAGGCATGACCAACCGCTTCCACACGGCGGAGGACCCACC GACCTATTCTATCCAGTCGGGGTCAGACAGGGTTGCCATTAAAGCC CCGACGTGTTCAACCTAACCTGGAAGTTGCTACGGGAAATACAAAG GACGGGGAGGGCTACACATCGAGGCGAGGCGTATTTCCCATCTCGC CTAGCCGATGATGTCGGATCTGGTGACTCCCGGCTGGAATGAGG AGGCTGTTTTGACAGAGAAAGGGAGCTCGGGCGTGGCCTCACC TTCAACCCAGCGAAGTGGCTACACTGCACATAGCGGACAGAGGKAG GAGGCAGTTAGACCGACTGAGTTCTCGTTCAAGGGACACCTGTCCGTG CCCTGCTTCGGGGAGGCTACGAACATCTTGGCATCCCAAGGCTA CCAAGTCAGGCAGACCCCAATCAACACCTTGGCGGATCTCTTGTGAC ATCGGCTCCATCGACCGGTCCTTACTGGGSCCTGGCAGAAGCTGGATG CAGTGGGAACCTTCCTACTGCCCCGGCTTGATTTACCATGCAGGGCGC CCATATCGACAAAGGGCTTCTTGACCGAGGCGGACAAAGATCATCAAAAA GCTGCAAAATCTTGGCTAAGCCTTCTCAGCGAGCGAGCGGGAGCTCG TGTTCTACCTCCCTCGAAAGGGGAGGGGTTTGCTCACGGTCGCTCA CAGCTATAAAATGCTGTACTCGAGCGATGTACCCGTGAGCACCATAGCG GGATCCACATGAGGAGGACCGTTTCGGAGAGGCTTAAAAAGCGGGCG TCTAACATCGATATCGCCCGTCTCTCCGGAGATCTTGATCTGCCCG GTCGACGAGCCCGTCGACATCTTGACCAAGTTTCGAGCGCGGCCCT GAGAAATTAAGACCAAACTCGGGCTCGCTGGAGCTGGTGTACAGGGG GGGAAGTACTCTGATGGCTGTGGGATCCACGTGCCCCGGGAACAA GAGTCTCGCCGACAGACAAACATCTGTACCACGAGTCTCGGGCGCTG CCTGAATCGGCACTATCGGAGAGTCTCTTGCCAAAAAGATCAAGGC AAGTCTTGAGGTTACCGGAGGAGCGGACATCCAACCACTTCCTCC GGAGTGGCAGCTTCACTCGCTTCTGCGACTGGCGCTTATCCACCGCGC CCGTCTGACGTTCTCCCTCAATGCTGCCAACGGTGGAACGAGGG ATGGACAAAGAGATGTCGAGATGGGGTCAGACTGGAGACTCTCCCG CATGTGCTGAGTCACTCGGCGCCGCACTCGGCGGAGAGACAAAGAGG CACAAACATCCAGGACAGTTGGTGAAGGAGCCTCGAGATGTCGG GGACCAATTAGTGTCAACCAACCGTTGTGGCGCTCGCGGCCCGGAC GCGGCATTAAAGACCKGACATTGTGCTTCGCGACGATGTCAACCGCGCG TCACCATCGTCGATGTGGCGCTGCCGTTGAAAAACCGCTCGAGGCGTT CGACGGCGTTGKGAGGCCAAGATTGCAAAATATACACCTTTGGCCAGG CAACTCAGGACAGCGGTTACACCGTAACGGTGGAGGCGCTTTGTAGTG GGCGCCCTCGGCGCTGGACCGCGGAAATGAACGGGCTCTGAGCCTC CTCAGCATTAGTCGTTATTATGCGATACTGATGAGGAGGCTCATGGTTA GTGACACAAATTCGCTGTTCCCGGACATCTACGTGGAAACATGTGTCGGG					AAAATA CAACTTT TCC (SEQ ID NO: 1310)	SAFTSVPHSSILQALE GHGLPSKARNISSLY TGMTTRFHTAEGPT DPILIQSGVRQGCPLS PDVFNLTLEVILREIQ RTGEGYTIEGRRISHL AYADDDAILADSPAG MRRLFAAERGARA VGLTFNPAKCATLHI AGRGEEAVRPTFSV QGTVPRALASGEAYE HLGIPTGYQVRQTPI NTRLRLADIGSIDRS LLAXWQKLDVAGTFL LPRLDFTMQGAHDK GFLTEADKIIKKAAS WLSLPQRASAEVLFL PPSQGGGGLLTVAS YKMLYSSDVTVSTIA GSTLRRTVSERIKKRA SNIDIARFLSGDLPL RSTSPSTFWTKVRS ALRIKTLGLRWSWC QGGEVLLMACGDP APGTRVSPQTKHLVT TSLRCLNRHYAESLL AKDQKGVFEVTRRS GQSNHFLRSGSFTFR CDWRFHRRARLDVLP LNAKRWRQGMMDK RCRRCGSDLETLPV LSHCGPHSAARQKR HNNIQDRLVKAASRC PGTISVNQTVWVGRG PDAALRPDIIVRDDV NRRVTIVDVAVPFEN RLEAFDGVREAKIAY TPLARQLTDSGYTVT

NeS L	Utopia -1_PCa	.	Phytophthora capsici	CATTGGCAGTACCGGGAATAGTGGGGCCATAACACCTAGGCCCCACA GTGTGGCGATGTCCATGTGTGCTGCTTACTTATTTATTTGTCCTG TCGAGCTGTTAATCTATTATCATATGATGTGTGTGTGTCGTCCTCGT AGCAGTTTATTCGTCCAACGAGAGGTCGACCAATATTAANAAGCAT GGCTTGAAGCAGGCCAAGCGCGTGTCTAACCCCGTTTTAGGGGAAGT TACTTAACCTAAAAATACAACCTTTTCC (SEQ ID NO: 1064)	GCCCGCGG GTGGAGTA GCCATGTT GGCCACCA CCGCCCAA GTCTCCGC CGCAGCTG CGACTGCT GCTGCTCA TCGAGTCG CAGTAGTC GCAGCTGC ACAAACAC CACCTCCT GCCGGAC GCCCGCC GTGGAGC ACCACGG CGCGTGA GCCGTACC AAGACCAA GGNTTCCA GGCTCGCG CGCGCTG GCGCTCC GAGCCAGC GACGCGTT CAGCAGCA GGGTCCG CTGCGGCA GCTCGACG CTGCACGG	TAGACG GCACAG TTCTGG CCACG TAGGCC GAAAGG GCCCA CCCATGT AGGAA CCGCCCT CGGAA ATCCATT CCGGTG TTCGACT GAAGAG ATGCTCC TTCGCT TGACGG AGGTAC ATCTCG ACAGTC GAACTT CAACTC GCAACA TATCG ATACAG TTACAA ACCACA GTTAGA TATCAG ATAGGA ACCTTC	VEAFVVGALGAWDP RNERVLSLSIRYAI LMRRLMVSDTIRWS RDIYVEHVSGIRQYRE (SEQ ID NO: 1432)
				CCGCGGCAAGCGCGCGCTGAACGACGCGGACGAGGACCAAGACATCCG GAGCTGGCGGAGCTTCTGCTATCGATGCGAGTCGAGTAGTCGAGCTGCAC AAGACCACCTCTGCCGACGCGCGCGTGGAGTAGCACCACGCGCGCG CTGAGCCGTACCAAGACCAAGNITTCAGGCTCGCGCGCGTGGMG CTCCAGCCAGCAGCGTTACGACGAGGTCGGCTCGGGCAGCTCG ACGCTGACGCGGACGCGAGCAGCGGACGACCAAGACCAACGACGCA CCAAGCCGTGCTGCCGTCCCATGGACGTGCACCAAGGTGCTCGGTGG CCGACAGGGATCAMACCCGCTGTGCCGCGACCGGAGCTCCGAGTC GGCGCAAGCGCGCGCTGAACGACGCGGACGACGAGACATCCG GAGTGGCCGAGCTTCTGCTGCCGACGAGGAGGAGCGCCGACGACCA AAACGAGCGCCAGGTTACCGCGACGAGCGCTCATCCGGCTCCGTC TCGCTGTGACGCGCAACGCGCAGCGCTTCAACTGACGTTGTGCGT GTACAGGCTGCCAGCTTCGCTGCTTTACGCGACACAGGGACTCTCGG CACCGCGCGTGACCTTCTGGACAGGTTCTCGGGGGTTGCGCGTGC GGCAACCTTTGCTCGAGGCTGCCGCGCAGCAAGACACGACCAAAACG TGCGCAGCTCAGCACCACTGTGTCGCGGTTTCGACGACAGGAGGA GCATCAAGCCACACTGCTCGGAGCAACACACCGCTGCCACGGCG GTACCGCGGAACCCCTGCTCCACCATCAAGCCTCGGAACCTCACTGT GCCCGCCACGTTGAGTTCGCCGACGTCGATGTGACGCTGCACAG TCCGCCACAGGAAGATCAGCAGGAGGACGCCACCCANACCCGGAGAG CAGCAACACCAACTCTGAGGCAACCGCTGGGTTGCGCGCTCGCG CCCACGCTGTTGCTCCAGGATTGCTCAGCGACTCGGCGAGCTGGAAC CTCCACGCTGGGGCCACCATTACCGCGCGGATGTTGTTTCGCGAAT TACAGCACGGCTTGAAGTACACTGCTCTCGATGGGATCCGCCACTC CCACGCGCGTCATCGCTCCCGCATTCGCCGACCGGCTTGTAACCTGCGA CGGACCGCTGCTCGCTCGGCGCTGAACGCTGCTATTCCTTCTCATCGCA GACACGGTGACGGAATGTTACCCAGGAGGACAGGCAACTGGAACCC CTGACGCGGCAGTGTGACGAAGAGACGAAGGACAGCGAGCTGCGCGG CAGGGCGCCGACTGTACTCGACGAGGAGTCCAAGGACAAATGACGCCAC TGCCGGTGAGTGGCTCCTCCGGTTCGACGGAGCTTGTGAGCGAACCCA GGCCAGGGGAGCAGGTGCCGTCTATTCAACCTAGCGGTGCGAGCG TCATGGACCTGCTCGCACTTTATGCCGGGGCGCCACGAGACAAACAACA			

CGCGGAGTACACGGCATTCCTCTAGGCGCCAGAGCTGCTGCAGATCA	GGACGGC	TTTAGG	RTDPHNRPAIAKARR
CGGTGCGACGCTACTTCGCGTCCAGGGGACAGCCAGCTCGTACTCAG	AGCGACG	AAGCTA	RVGRIRSAINQQLLR
GCAGGTCAAGGGCATTTACGGTGCAGAGACACCCGCTACGGAGACT	GCAGGCAC	ACGGGT	HKFDTDEKACVDGIL
ACGCGACGCTGTCGTCGGAATGGCTCGGTGGGGCAATTCCTCCTC	CAAGACCA	ACACTG	STARAEARAATPS
CACCACATCGACGAGGACATGCGACGACGACCGCTTGCGAAC	CGACGACC	GATGGT	PPASGAPTTTVSAPG
AGAGCGTCGACATGAAGAGCACACTAGTGAATGCGCAACGCCCC	AAGCCGTC	AAATAC	AIVTNDGTCPIPSDK
GGCGGAACGCTGACACGACACTGACCACTGCTGCGCTGAAT	GCTGCCGT	ACATAC	LWRHFDVNTPLRD
CACGGCATCACCCCGCGGTGGAGCTCGAGACACGCCGATGGCCG	CCCCATGG	ATTTC	FDAEAPGSAAFRAA
ATGCCGGCAAGAGCGCTCGAGACGTTGACGATGGTGAAGTTTACG	ACGTCGAC	(SEQ ID	MDHLPAATRLDLLK
CGGGATGCGGCTTGCCCGAGCGAAGTTCTGAACGACGGCCACGAC	CAAGGTGC	NO:	EAPSTDEIETQLQHV
TCCGCTCCGCCAGCTGAGCGACGAGGAACTGGAAGCAGCGAGCGAGA	TCGGTGGC	1311)	KASSPGLDGVGYDV
TGTTGGAGCGACTGGGCGCCGCTTATCGGCCAAGATCACAGACGCGG	CGACAGCG		YKRFTIQLLPVLRAAF
AAGACTGGGCGAGTGGGAGGGCTACATACAGGCCCTACCGTACATGC	GATCAMC		RCCWLYKKVPQSWK
TGTACGACAAGCTTCAGTCTGACTACAGGCACCGCTGGACCCCAACA	ACCCGCTG		LGVVRLHKKGPRED
ACCGTCTTGACCCGCTCCCGCGGCGACGATCGCCAGCTTCTCTCG	TCGCCGCC		PANWRPICLQQAIIK
GAGCCCAACGCATCATCCACGACTGGAGGGTGGCATCGGAGCACCAA	ACCGGAGC		LYTGILARRLTRWMD
CCGGCGCGCGCGACGACGAGGACGACGCAAGGGCGGACGACAGCG	TCCGAGTC		ANDRHAPGQKGFRA
CCGCAACCCGCGCGATCGGGACGAGGGAGCTACAGGTGGGCATCA	GGCGGCA		VNGCGEHNFLAATLI
ACAACACAAGAAACAAACACACGCCACCCCGGAGACGACGACATCAT	AGCGCCGC		DNARRKHRPLYEVW
CGGGAACACCGTCTCGACGAGGCTCTCGATGAGCTCCAGCCCTCGAGC	CGCTGAA		YDFRNAFGSVPFALL
GTACCGATCCTACAACCGGCCAGCAATTGCCAAGGCCCGCCGCGCT	CGACGCG		WDSLQRLGVPPDVV
CGTGCATCCGCTTGCCATCAACCAGCAGCTACTCCGCCACAAGTTC	GACGACG		DMCKGLYNQASFVV
GACACGACGAGAAAGGCTGTGTGACGGCATCTGTCTACAGCCCGC	AAGACATC		GNAVVGSTAPVEQR
GCCGAACGCGTGACGAGCAGCAACGCCGTGCGCCACCGCATCAGGA	CGCGAGCT		GVVFQGCPLSPQLFN
GGCCACACGACAGTCTCGGCGCCGGTGCCATCTGTGACCAACGATG	GGCCGAG		AAISPLLYALRRLPDT
ATGGCAGTGCCCCATCCGAGTGACAAGCTTTGGCGACACTTCGACGC	CTTCTGCT		GVQLSSVDRPGASAY
CGTGAAACACACCGCGCTTGACTTCGATGCGGAGGACACCCGGAGTGC	GCCCGACG		ADDLKIFSGTKAGITQ
GGCCTTCGAGCGCGATGGACCACTGCCCGCGGACACGGTTGCT	AGGAGGA		QHELVAFLRWGTGM
GGACCTGCTCAAAGAGGCGGACGACGAGCAGAGATCGAGACGCAACT	GGCCGAC		QANPAKCRSMGVRR
CCAGCACGTGAAGGCTCTCCAGCCCCGGCTCGACGGCGTCGGCTAC	GACCACAA		NTNGAVEADNVHLE
GACGTGTACAAGCGCTTCACGATACAGTTGCTGCTGTGCTGCGGGCGG	ACCAGCGC		LDDTPIPSMTHMQS
CGTTCCGGTGCTGCTGTATAGAAGGTTCTCAGAGCTGGAAGCT	CCAGGTTA		YTYLIGDGFHDVRR
CGCGGTGGTGGCTTGTCACAAGAAAGGGCCCGCTGAGGACCCCTGC	CCCGCGAC		RVELAPKLKTLKHDVT
GAACTGGAGACCCATCTGTCTCCAGCAAGCCATCTACAACTATACACG	CAGCGCTC		ALVESGLAPWQVVK
GGCATCTCGCGCGGCGACTAACCCGGTGGATGGACGCGAATGACCGC	ATCCGGCC		AVKVLYPRVEYALR
CACGCAACCGGTGAGAAGGATTCGAGCGGTGAACGGCTGCGGGGA	TCCGTCCT		HLRPDDQHLESFDLH
ACACAACTTCCTCGCAGCTACGCTCATCGACAACGCCCGCGCAAGCAT	CGCTGTGT		LRRGLRHLLRLPKNAT
CGGCCACTGTACGAAGTCTGGTACGACTTCCGTAACGCGTTCCGGCAGCG	ACGCGCAC		NEFFYAPVSRGGLGL

TGCGGTTGCGGTTGCTGTGGGATTGCTACAGCGCCTCGGTGTCGCCGC	AACGCGCA	LPLVELHAALQIAHG
CGACTACGTGGACATGTGCAAGGAGTCTACAACCAAGCTCGTTCGTC	GGGCTTCA	WQTLHSPDPAIRRVA
GTGCGGAACGAGTGTGAGTGGCAGACAGCGCGGTAGAGCAGCGCGT	ACTGCACG	REQLYQIADARHRLD
CGCGGTTTTCAAGGCTGCTCGCTAGCCCCAGCTCTTTAACGCCGCCA	TTGTGCGT	KDHWPHRREELCELL
TCAGCCCACTGCTCTATGCGTGGCGGCTACAGATACCGGCGTGCA	GTACACGG	LNGELGSLWVDVRK
GCTCTCAGTGTGACCGCCTGGTGCCTCAGCTACGCCGACGACCTT	CTGCCAGC	RNGDIGSLWVDVRK
AAGATTTTCAGCGGCAACGAGCGGATCAGCAGCAACACAGAGCTG	TTGCTGCT	NLKTGLKVATAPAN
GTGGCAGCTTCTGCGGTGACCGGATGACAGGCCAACCTGCGCAAG	TCTTACGC	QETGVPAQPLQLRVP
TGCAGGTCCATGGCGTGGCGGCAACACCAACGGMGCCGTCGAGGC	GACACAG	HHAEWLDHGNVLRH
AGACAAGTCCACTGAGCTGACGATACGCCGATCCCTCCATGACA	GGACTCTC	VKLHIKNLHWQTWC
CACATGCAATCTACAGTACCTGGGGATCGGGACGGCTTCGATCAG	GGCACCG	ALSDQGKTARVHGG
TGCGCCGCGCTCGAGCTGCTCCCAAGCTGAAGACTCTCAAGCACGA	GGCGGTGA	VGSFLTRPRGMWE
CGTGACGGCGCTCGTGGAGTCCGGCTTGGCGCATGGCAGGTGTGAA	CCTTCCTG	SDYRFAVAARLNVVD
GGCCGTGAAGGTGTACCTTACCTCGGGTAGAGTACGCCCTCCGGCAC	GACAGGTT	TVNTLSRRRLRAHDR
CTCCGACCCGATGATCAGCACCTGGAGAGCTTCCACCTCCACCTCGCC	CTCGGCGG	CRYPACRWKETLAH
GGGCGCTCGCCACCTGCTTGGCTTCCCAAGAACGCCACGAACGAATT	GTTGCGCG	VLNHCPGTMDAVRG
CTTCTACGCGCGGTGTCGCGGCTGTTGGGGCTGCTGCCGCTGGT	TGCGGCAA	RHDDALKEIEHLRAS
GAACTGCACGCAGCACTGCAATCGCGCAGGATGGCAGACGCTCCAC	ACCTTTTG	SGDRRELNVNQTV
TCACCCGACCCAGCCATCCGTGCGTGGCGCGGAGCAGCTCTACAGA	CCTCGAGG	GLPGPLRPDIQVYN
TTGCCGACGCGCGCACCGGCTGACAAGGACCACTGGCCACACCGGA	CTGGCCGC	HDKRTVAVVDLAVAF
GAGAGGAATTTGCGAGTCTCTCTCAACGGGGAGCTCGGGACGTCGG	AGCAAGAC	DEQPSDESSGLAK
CCCACGACCCGCCAAGCGACGCAACGGTGACATCGGCTCTCTGTGGT	ACGCACAA	AVQIKKAKYAGIKEHL
CGACGTCCGCAAGAACCTCAAGACGTTTGGACTGAAGGTGGCCACAGC	ACGTGCGC	ENQGWKVHLSAIV
GCCGGCAACACGAGACCGGTGTACAGCCCAAGCGCTGCAGCTCCG	CAGCTCA	GSLGSAASNHKVYT
CGTGCTCACACCGCGAGTGGTGGACCAAGCGCAACGTGTGAGGCA	GCACCACA	EHLGLLRDAKRLDR
CGTGAAGCTACACATCAAGAACTCCACTGGCAACCTGGTGGCCCTA	CTGGTCGC	QLSSACIQSSRIWN
AGTGACCAAGGGAAGACCGACGGGTGACGGCGGTGTGGCAGCGC	GGTTTCGA	FHCAKHRARQHEHQ
ATTCTGACCCGACTCGGGCATGTGGGAAAGCGACTACCGCTTCGCG	CGACAGG	APPSQATGRRRVTET
GTGGCGCTGCTCAAGTGGTGCACACCGTCAACACACACTGTGCGGCC	AGGAGCAT	GGNPSRTDRR (SEQ
GGCGACTCCGCGCCACGACCGGTGTGCGTACCGGCGCTGCAGGTGGA	CAAGCCAC	ID NO: 1433)
AGGAGACTCTGGCACACGCTCTCAACCACTGCCCGGGTACCATGGACGC	ACTGTCGT	
GGTACGTGGCGGCACGATGACGCGCTCAAGGAAATCGAACAACACGCT	CGGAGCCA	
CCGCGGTGTCGGGCGACCGCGCGAGCTGCGTGTAAACCAACACGGT	ACACCACC	
GCCTGGCTGCCCGGGCCACCCCTGCGCCCGACATCCAGGTCTACAAC	GTGCGCCAC	
CATGACAAACGACCGTCCGAGTTGTAGACCTGGCTGTGGCGTTCGATG	GGCGGTCA	
AGCAGCCGAGTGAGGACCCAGAGAGCTCGGGTTGGCCAAAGGCAGTC	CCGCCGAA	
CAGATCAAGAAAGCCAAAGTACGCGGGCATCAAGGAGCATCTCGAGAAC	CCCCCCT	
CAGGGGTGGAAGGTTCACTCTCGGCCATCGTGTACGGCTCGCTTGGCT	GCTCCACC	

288

CTCGGTATCGCCGGGTGTACACTCTCAGCCGCCACAGCTCGGGCCTTG	GCCGACTT	GCGACT	AGGMLPDPGLVEEPI
GCGGTCCGCCATTGGCCCTTGAGCTCGACAGCGACAGCAGCAGCAG	CGCGCTCT	ACCCAG	QATYAHDAAQFECA
GAGGACGCTCAAGACCCACCGCCGCCGCCAGAACCCCCAGAACGAC	GGCGCGG	GTAGGG	LCPVVAESMAVLVQ
GTGCGAGTGTCTTGCCACCCCGGCGGCGGCGAGCATGCTCGCC	CCACACG	AACCGC	HRRSAHRGTRFKDIF
GACCCAGCGCTGGACGCGGTCTGGCTCGGCTCGCGGCCCCCAT	CCGCGCC	CCTTAA	TSGCQCSLVFYARIVA
CGGACCCAGGACCAAGCCCGCTTCCCTGCTGGCCCGCGGCCA	GAGCTCC	AAACCC	ASHAVACARRNQRA
GAACCTCGCGCCGCCAGCCCGTGTGAAGTCCACACAGCTGTG	AAGCGCG	AGGAAG	VPPAPTPVAPTRPEA
GCACCCGCGAGGGCGGCGCATGTGCCCGACCCGGGGTGGTT	CCGTGGC	ACACAA	TPQPTGYLAAAMTA
GAAGAACTATCAAGCGACGTACGCGCAGCAGCTGCCAGTTCGAG	TTTCGCG	ACACCT	AAAASSDTVAAAT
TGTGCTGTGCCCCGTAGCTCGCTGAGAGCATGGCAGTGTCTGTTACG	ACGCAGG	CCACTTA	NMQSAVPAAAKTG
ACGACGTCGGCTACCGGGCACCCGCTTCAAGGACATCTTACACAG	GCTCGCG	GTGACA	LQLVPELEPALPQR
CGGTGCCAGTGCAGCTAGTGTTCGCGCGCATTTGTCGCGCAGC	CGACGGC	TACATAT	ASCHAGKRRRLNAD
CATGAGTGGCTGTGCCAGGGCAACCGAGCGCGCTGCCACCACT	CTCGCCG	TTTAGCC	EAVTCTPTARVSPQ
CCGACCCCTGTTGCCCCACGCGGCCGAGGCTACCCACAACTACTG	CCCCCAAG	TAGATT	TEVAMAPHDAPQD
GGTACCTGCGCGGCGATGACTGCGCTGCGGCGAGCGCGCATCAAGCG	ACCCCCCG	CAGTTA	DTVLQREAAEPQDP
ACACGTTAGTGGCAGCGCCACCAACATGACAGAGCGCGCTGCCAGCG	TACGATG	CGGAGA	AATPGAQVQRVEDT
CCGCAAGACCACTGGGCTTACGCTGTCCTCCCGGAGCTGGAGCCTGC	GGCACCA	GGTAC	TAAQDDTVQQDHD
CCTTCGCGAGCTGCTTGGCATGCAGGCAACCGCCGTCGGCTCAAT	CCGGCAG	TAACTG	ADTAQVSPRRTPTR
GCCGACGAGGGGTGACACCGTGTACGCTACGGCGGGGTGTCGCCG	GGCGGCC	GTAAT	WGRPSSTQEPSPM
CAACAGAAAGTGGCGATGGCGCTACGACGACCCCGAGGACGACAC	GGCAGGT	ACGAAC	TGEPAAATLAARRPLT
GTCTACAGCTGAGCGCGGAGCCACAGCCAGCCACGACGACGACG	GCCCGACT	ACATATT	PAATGTRATRWGPC
CCGGTGCCCAAGTGCAGGAGTGAAGACACAACTGCGGCGCAGGAT	CGGTATCG	CTGTCT	HRAIGAAAIALVTGL
GACACTGTCAGCAGGACACGACGCGGACACTGCCAGGTCTGCGCG	CCGGGTGC	AATCAG	STEPAQPRRQPPPP
CCCCGGCACACCCACACGCTGGGACCAAGACCAAGCTCAACTCAA	TACACTCT	TGTGAA	QEPSPQPEPLAAAT
GAGCCTAGCCCCATGACAGGCGGCGGCTGCCACCTGGCAGCGCGA	CAGCCGC	AACTGG	AAADIAATVAADIAA
AGACCACTGACTCTGCGCAACGGGCACTCTGTGCCAGCGGTGGGG	ACAGCTCG	TTTTCGC	AAANAAMDVDGGP
CCATGTACCGCGGATCGGGCGCGAGCCATCGCAGCACTTGTACCG	GGCCTGG	CTTTGG	AADETWLLRFDGAC
GCCTTCAACGGAGCCAGCAGCCTCAGCGACGGCAGCCACCGCCAC	CGGTCCG	CGGACT	RRNPGGGAGAALF
CACAAGAACTCCATCGCAACAGAGCCGCTAGCAGCAGCAGCAACAG	CATTGGCC	TTTTCAC	APSGAVVWTCSHVM
CCGACGACATAGCAGCAACAGTTGCAGCAGACATTTGCGGACGCG	CTTGAGC	TCGCATT	PSRSETNNTAEYFALL
CAGCGAACGACGCAATGGACGTGGATGTGGACCGCAGCCGACGAG	TCGACAGC	TTTGGG	LGVQSAVHHGASHL
ACATGGCTGTGCGCTTTGACGGCGCATGCCGACGGAACCCAGGACCT	GACAGCA	CAATCG	EVEGDSSLVIAQVKG
GGTGAGCAGGTGCCGCACTCTTTGCGCCAGCGGTGCCGTGTGTGG	GCGACGAC	TCTCGG	TFACRNARLRQLNR
ACGTGCTCGCACTACATGCCAAGCCGCTCCGAAACGAAACACCCGCG	GAGGACG	GCTAGC	VRHALRSVDTHKLRH
AGTACACGGCGTTACTGCTGGCGTACAGAGCGCAGTCCACCCAGGTG	CTCAAGAC	TTGCTA	IDRQANAHADRLAN
CCTCGCACCTGGAAGTCGAGGGCGACAGCAGCCTGGTAATCGCTCAGG	CCCCACGC	GCGCGG	RALDQRRRTSSECGTH
TAAAGGGCACCTTCGATGACGGAATGCGGGCTCCGCCAACTCCGCAA	CGCCGCC	GACGAG	GSCMDSCLAVPTALA
CCGAGTGGCTACGCGCTCCGCTCCGTGGACACGACAACTCCGGCAC	CAGAACCC	CGGTCT	AQETPPAAPPPSTSAT

ATCGACCGCCAGGCCAACCGCGCATGCAGATCGCCTTGCCAAACCGCGCGC TAGACCAGCGGAGGACAGCAGCGAGTGGGGACGCGATGGCTCGTGC ATGGACTCTGCTCGCCGTACCCACGGCCCTGGCGGCGCAAGAAACGC CACCGGACACACCCAGCAGCAGTCCACTCCAGCCGAAGTAATGC CATGGACGACATCGAGCCGAGATTGCTGCGCTGACGAGGGCGAGAC CTTCCAGTCTGCCAATTGGCCAGGCTAGCCCCAGAACCGCAACCC CGACTCCGTGCGCAACTTAGCGATGAAGAGCGCGACGACGCGGCA GATGCCCTGAGGAGCTGCCGATACCATGGCGAGTAAGATCGAGGAT GCTGACAGCTGACCTCGGCGAAGGATACATCAGCTCCATCCCGGAGC GTATCCGAGAGGTACTGACGCCATACGCTACGGCCCGCCGACGCCGG GACGCTCTCGCCGCAACAGCCGCTGACCCACCAACCGGTGACGCGCA ACCAGCGCAACATCGACTCGACGAAGCTTTGGACGACATGGCAGCCA CACAGCAAGCAGCCCCGCGAGACCAGCGGGCTGTGCGCCGTGCGCGGC GACGTGTGGCCGGGTGCGAGCCTCAATGGCACAACAGGAGCTGCGCC ACGAGTTGCCAAGGACGAGTCTAAGTGTGTGGCCAAAGATCCTGAAGA CAGCGTCGACAGAGACGGCAGCGGAGGAGCAGCACCCTGGAGACGTGC CCCATGATGACGTACGCTCCACGCCATTTACGGGAGTCAACGCGC CAAGGACGGACTCGACTACGATGCCAGCCACCGCGCGGGAGTTTCGAG CGCGATGAGCGACCTTCGCCAGCCACCGTGGAGATTGATGCGTTGG ACGCGGAGCTGACGATCGACGAGGTGGAGGACCAGCTCACACGCGCG GCCAAGACCTCAAGCCCCGCCACGACGGCATCGACTACGGCATCTACA GCCGTTCCCGCGCAGCTGTTCCGCTCTCCACGCGGTGTTCCAGTTT TGCTGGAGACACCGTCTGTGCCAAGACTGTGGAAGGTGGGCATCGTA CGCCTATCCACAAGAGGGCGACCCAGGAGCAGCCGACTAACTGGAGG CCGATCTGCTTACGCCACGATCTACAAGCTCTACAGCGGGTCTGCG CGCATCGGCTCTACGCTGGCTGGAAGGTAAACGACCGCTGCCGATGG CCAGAGGGGTTGAGGGCTTCAACGGGTGCCCGGAGCACCAGAAAGCTGA GGTACACAGTTGCTGGACAGACCCCGCGGAGCAGCCGAAAGCTGT CCAAGTGTGACGACCTTCGAACGCTTCGGCTCGCTTCTCTCAGCAG CTCATGTGCGCGTCTTCGCCACCTTGGCGTCGACTCAGGCTTCATCG ACCGCTGACGGGATATCTACAGGACTCGGCCTTCTGTTGGTGGCGAACG CGCGGACGTTGCGACGAGACCCGCTGCGCCAAAGAGGTGGCGGTATAC CAAGGCTGCCACTCAGCCGCTCCTGTTCTGTCGCGGCGCTGGTGCTC TCGTGCGCGTCTCGAGAACTCGACGCGCTTGGTGTGCCACTGGCGG ACGGTGTGCGGCCCTGCACACGGCTACGCCGATGACCTCAAGGTGTT CAGCGACAGCGCGCGGATCCGCAAGTGCCACGACACAGTGGCTGG CTTCTCGCGTGACAGGGCTGCGCGCTAACCTGGCAAGTGTGCCAGC CTCGCAGTGACGACGAACCGCCGGGCAACCCGACTAGAGACGACGCG	CCAGAAGA CGTCGCGA GTGTGCTT GCCCCACC CGGCAGG GC (SEQ ID NO: 1189)	CCGGGG GCGTTC ACCTTC CCCCGC GAGGCC AACTAC ACCGAT CTTCTCT ACACTTT TCTAATT CGCCTCC GTCTTC GGCTTT CGGTTC TCGGGC TTTTTC TTTTGA CCAATC AGAGCG CGCCAT GCGCT CTTCTG GCCAAT CAGAGA CCGGGC CCTGTCC TCGGAC AGCGAG GCCTCC ACGGCC AGCCAA TCGAGT CTCGGC AGCGAC GCGTCT TTCTATA GCGCAG CTGACG	PAEGNAMDDIAAEIA ARDEGETFVLPGP GSAPERQPRRLRQL SDEERDAADALQEL ADTMASKIEDADSW TSGEYSSIPERIEV LQYATAPPQGRSR RQRRRPPRVTRNQ REHRIDALDDMAA TQQATPRDQAVRR ARRVGRVRASMAQ QELRHEFAKDESKV AKILKTASTETADE HPETCPIDAATLHAH FTGVNAPRTDFDYDA TSGREFRAAMSLPP ATVEIDAFDAELTIDE VEDQLTRAAKTSSPG HDGIDYGIYSRFAAQ LVPLLHAVFQFCWRH RRVPRLWKVIVRLI HKKGDPQPTNWRP ICLQPTYKLYSGLLAH RLSRWLEGNDRPLM AQKGFRAFNGCHEH NFMATTLDDQTRRQ HRKLYQVWYDLRNA FGSLPQQLMWRVLR HLGVDSGFIDRCRDY RDSAFVVAADGA TDPVRQEVGVVQGC PLSPLLFVAALVPLVR RLEKLDGVGVPLADG VRPCTTAYADDLKV SDSAAGIRKCHDTVA GFLAWTGLRANPGK CASLAVTTNARGNPT RDSSMRLEVHDAIT

ATGCGTCTCGAGGTCCACGACGCCGCCATCACACGCTCAGCCTCCACG	AGGCCG	TLSLHESRYLGVGD
AAAGCTACCGCTACCTGGCGTGGGTGATGATATGACACGTCGCGC	ATCTGG	GYDHRHRLQLEPKL
ATCGTCTCAGCTCAGCCAAACTCAAGCAGCTCAAGCGGAGGCGG	CGGCC	KQLKREAVALLTSL
TGGCGTCTGACGTGGGTTTGGACCGTGGCAGGTGGTGGCGCGC	CCGATT	APWQVVRALKVVVY
TCAAGGTGACGTACCCCAAGGTGGAGTACGCGCTGCGCCACCTGG	GGTCCG	PKVEVALRHLRPLQS
CCCGTGCAGTCGAACTGACGGCTTCGACCGTGTGCTCTCAAAGGC	ACTTTCG	QLQAFDRVSKGLRH
TTGGACACCTGTGAGCTGCCTGCTCAGCCACCGAGGTGCTGT	GCCAAT	LLSLPRSATSEVLYAP
ACGCCGACGTCCAGCGGAGGCTGGGCTGACCCGCTGCTGAGC	CAGCGA	TSSGGLGLQLVELH
TGCACCGCCCTACAGCTGCCACGCGTGGCAGATGCTGCATTCCAA	CGACGA	RALQLAHAWQMLH
GGACCAAGCATCCAGCAGTGGCAAGAGCGCAGGCTTCCAGGTGGT	GGGGCG	SKDPAIQAVARAQAC
GCGCAAGCGTACCGCTGCAAGAAAGACCATGGCGCGGCGCGATGA	AGGGGT	QVVRKRYRLQEDHW
CGAGCTGTGCGCAGCTTCTGAACCTCGAGCTCGCGCTCGCCACAC	TTACACT	RGDDDELVRSLNSE
GCTGAGTCTCCGTGCAACGGTGACATTCGCTCGCTGTGTCGACG	TTTGCC	LAASPHAEVLRNRND
TGCAGCGTGGCTGCGCATCTACCACTGCGGTTGAGCACTGCGACGA	CCGTTTC	IASLWSDVQRWLRIY
GACAGAGGCCCATGGCCGCTCAGCTTCGCGGTGCGCACCAACAA	GACTTC	HLRFEHCDETEAHGP
GTGGCTTACGCACAAGACTGTGCTGCGGCACGTGAAGCTACACCTGAA	AACTTCA	LSFRVPHHNKWLTH
GATCCGACACAGACCCGATGGAAGGGCATGGTGACAGGGCAAGA	GGCCAA	KTVLRHVKLHLKIRH
CGGTGCGCTGACGGTGGGTGGCGGCCAAGTTTCATGACGACCGGA	AATGCG	QTRWKGMDVQDKT
GCGGCGCTGTGGACGACGACTACCGCTTCGGGGTGAAGGGCCGCTC	GATTGG	VRVHGGVGAKFMTT
AACCAAGTCGACACGAACTCGTGCTCAAGCGCAACGACTGCGGGCA	GACCTT	GAGLSDDDYRFGVK
CACAGACCTGACGGACCCAACTGCTCTCGGCTGAGACGCTGGCAC	CCACGC	GRLNQVDNLSVKRK
ATGTACTCAACCACTGCGAGTCCAACTGGACGCCATTCGCCAGCGTCA	GCCGTG	RLRAHTTCRDPTCSS
CGACGCGCCCTGGAGCAGATTGGGTCTAAGATCCGCGGAGCGCTCGA	CCACTG	AETLAHVLNHCESN
CCGCGCAAGTCTCCACAGAGCTGCGCTGAACAGACTGTACCCGAG	CTCGGC	MDAIRQRHDDALEQ
TACACAGGCGCAGCTACGCCCCAGACATTTGCTTAGAAACGTGGCCG	ACCGGC	IGSKIRGALDRAKSPT
CCAAGACGATGGTGATCGCCGACTTGGCGTCACGTTTGAAGACCAAG	GGCGAT	ELRLNQTVPEYTGAA
CGGCTCGGCTGCTACTCTCGTGAGCTGAGCTCAGCCACGACCAAGAC	TCAGCG	LRPDIVLRNVAAKTM
TCTCAAGTACCAACCAATCGTGGCGGAGCTGCAGCACAAGGGATGGCG	GGTGCA	VIALAVTFEDQAAR
AGTGCAGACTGTGCCATAGTGTACGCGACGCTGGGCTCGGTGCAGCC	ACTTCG	ARHSSLQLSHDHKTL
CAGCAACTTCAAGGCTACACTGAGAAAGTTCAAGTTTCAAGCGCGAG	GGCAGC	KYQPIVAELQHKGW
GCTAGGACGCTGGATCTCAGCTGTCTAGCCACTGCAATTCGCGCAGCC	TGTGCA	RVQTAIIWGTLSV
ACCGCATCTGGGCTGGACTGCCGACAGCACCGCGATCGGCAACGAA	ACACAT	QPSNFKAYTEKFKLH
GCGGCACTGCTCGGAGCGTGGGTGGGTCTGGGGGGCCCCCGCGCG	GCGCGC	KREARQLDLQLSSHCI
GCACATCGCAGGACCGGACCGGCGATAAGCTGGTTCATCATGACCGAC	CCCATTTG	RASHRIWGWHCRCQ
AGGGCACTACCCAGGTAGGGAACCGCCCTTAAAAAACCCAGGAAGACA	CACGCC	HRDRQRSGTASRASR
CAACACCCCTCCACTTAGTGACATACATATTTTAGCCTAGATTTCAGTTA	AAGCGG	GSGGAPRRTSQAPA
CGGAGAGGTTACTAACTGGTAAATACGAACACATATTCGTGTTCTAATCA	CATCGC	RR (SEQ ID NO:
GTGTGAAAACCTGGTTTCGCCCTTTTGGCGGACTTTTCACTCGCATTTT	GGGACG	1434)

				GGGCAATCGTCTGGGCTAGCTTGCTAGCGGGGACGAGCGGTCTCCG GGGGCGTTCACTTTCCCCCGGAGGCCAACTACACCGATCTTCTCTACA CTTTCTAATTCGCCTCGCTTCGGTCTCGGTGTCGGGCTTTTCTTT TTGACCAATCAGAGCGGCCATGCGCTCTTCTGGCCAATCAGAGACCG GGCCTGTCTCGGACAGCGAGGCTCCACGGCCAGCCAATCGAGTCTC GGCAGCGCGGTCTTTCTATAGCGAGCTGACGAGGCCGATCTGGCG GCCCCGATTGGTCCGACTTCGGCCAACTCAGCGACGAGGGGGCA GGGTTTACACTTTGCCCCGTTTCGACTTCACTTCAAGCCAAATGG CGATTGGACCTCCACGCGCGTGCCACTGCTCGGACCGGGCGCGAT TCAGCGGTGCAACTTCGGGACGTGTGCAACACATGACGCGCCATTG CAGCCAAAGCGGCATCGGGGACGAGCGCTCGGCCCCCAAGCGCAGC CCGCGCTTCCAGCACGACCTCGCGCGTTTGGCGGATCGCCATCAAGA CGTGCGAGAGCCAGCGGGTGGGCAAAATATACTTACTCTAAGTAT GCCGAATCCCTGCCCTCTCAGGCTGAACGCGGCCCATCTTGTATCTAA GTATGGGAGGATCCCTGGCCTCTCAGGCTGTACGCGAGACCC (SEQ ID NO: 1066)		ACGCT CGGCCG CCCAAG CGCAGC CCCGCC CTTCAG CACGAC CTCGCG CCGTTT GGCGGA TCGCCAT CAAGAC GTGCGA GAGCCA GGCGGG GTCGGG CAAAAT ATACTTA CTCTAA GTATGC CCGAAT CCCTGCC CTCTCAG GCTGAA CGCGGC CCCATAC TTGATCT AAGTAT GGGAGG ATCCCTG GCCTCTC AGGCTG TACGCG AGACCC (SEQ ID NO: 1312)	

NeS L	Utopia - 1_PMI	.	Patiria miniata	CTGATGTGGATACCTTGAATTACTCAACCGTGTGCGAGTCTTTTGCTT TTGCGCCCAACACCTCATGGATACCATGCTTGTGCTGGAGCGACGTTA CAAGGTGAGGGCGCCCTCCATGCCGACAGCTGGTCTGTGCCATGTG CTTGAAGTCTTCTCTCGACTTCGGGCTTGGGGGACACATGGCCCGG CTGCACGGCAGCCTTCTCCGACGCCCTCACGCCCTGACTATGACGG AAGTGTCCGTACCCCTTCCGAAGGTGTGCCGTGCGTGTGTGCCG CGAGAACTTCTATCACAATCAAACTTGAAGCGTACCAGAAAGAAATTTT ACCCAAATGTTGACTCACCACTCACCGTACCACTTCAATTTCAAAGTTTATCGG GGAGTTGTACCGGACGAGCAATCCCTACCCGTCTGCAATGCCCTCTCAC CGGATGTCAGCCAGCGCATCCCTACCCGTCTGGGAAACCCCTGCTCTAGCCCT GATGGATTGATCCGAGGCTGTGCGGAAACCCCTGCTCTAGCCCT GCCGTACACCGCCACCACTTGCCCGATACATCTCCCAAGCGCCCGGTC TTCTTTTTCACCACTGGAATCCCTCCCAAGCGCCCGTGTGCCACCC ACTGCCAGGCCACTCACTCCACCGCCACGATCACCGAGCCCGCGGCC CGGTCCCAACACCACTCCCGGTCACCTCACCGCCCGGTCACCGAGCC CGAGCCCGCGTCCCGCTCACCGCCCGCCCGTCACTCCCGCCCGGTCAC TCCGCCCGCGTCCCGCTCACCGCCCGCCCGTCACTCCCGCCCGGTCAC GGTACCGAGCCCGCCACCGCCCGCCCGTCACTCCCGCCCGGTCAC AGACTCACTACCGGTCACTACATCTCATGCTCGATCTCTCTCACGGC CAGCCAGCCCTTCTCTGTGACGCCCCGACATCTCGTCCGTCTACG CTGCTCACCGGCTAGACAATGCAGTTCACCATCGCAGCAGTCTG CAATTGACGTGTGGTGGCGAGTGTACGCTCGGCTGACTTCGAGTC TTTTGAGGATGTGTGATCGGTTTGTGGAGTTTCCCGCCGTGAGGGT CGAAACAAGGACGCCCGCCCGCCAGCGCACCAACCGCCCGGTGAC CGCGGAACAGGACCTCGCCCCAACGTCCTCCCTGTCCTCCACCGCC CCGGCTTGGCCGAGTTTGTACAGAGGCGCATCAGGACTATCTTGACCG AGCTGATCGGACCAAGAAAGCGCGCATCAGGACTATCTTGACCG GCTGATGTCCGATACAGCGGTACCGAGGTCTCACTAGTGGAATTC TTTCACTTGAAGTATTTCCGCTCGTGAAGTCCGCGACACCGCATTTGCC ACGGATACGAGCTCTGTTCCGAATCCCGCCAGCGTCTCTCCCTCGTCTT GCCAAGACTGCTTCCGCGATCACCGAGCGGAAAGCGGTGGAGTACCGT GGCCGGATGTCCAACCTCCGACAGGAAAGCGGTGGAGTACCGT CACATCCGCCAGGTTGATGGCGCATTCGCGGTCACTCCCTGGAATCTTCA ACCGGTGCTGCTGAAAGCGCGGTCCCATCTCTGGAAGACAGCGA CCACCGTGTGATCCAAAGAGGGGATGCCACCGACCCGGCCAACT TCCGCCCATCGCCCTCCAGTCTCTCTACAAGCTGTGATGGCGATC TTGTCCGACCGTGTACCAACATGGGCCCTGGACAACGATCTCATCAGCA GCTCGAGAAAGAGCGCCCGCGGTGAAGTTGCTACGAAACACAT	CTGATGTG GATACCTT GGAATTAC TCAACCGT GTGGAGT CTTTGTCT TTTGGCC CAACACCT CATGGATA CCATGCTT GTGCTGG AGCGACGT TACAAGCG TGAGGGC GCCCTCCA TGCCGGAC AGCTGGTC TGTGCC (SEQ ID NO: 1190)	GATTAG CGAACA CTAATAT CCTGCC MTAGAC GTGATT GCTAAT CCGCAA ACCAAC CGGATC TACAGC CTGAAC ACTGAA CTTTAAT CTTCACC CATGTC ACATCT GGACAC TAGGTTT TTGCTCT GTTTGT GTTTTCC TGCCTTT WCMTT GGAWCT TTCGCM CTGGAA TTTATT GTCGCT TGGATT ATTTTT TTCTCAC AATTTG GATCTAT TTTCGTT TGTTAC TTCGAA CTCTAGC	MCLKSFSTSLRRH MARLHRQSPDAST PSTMTVEFPYPLKV WPCVVCRENFYHNQ TLKRHQKNFHPNVDL TTVYQCSVCGQEFVT GRKASHFKVHRRM SASAIPTSAMPPSP MDLIRGLVGEPLPPS PARTPPPLARYISPAP RSFSPPWNPPPPR SPTPLRPLTPPPRSP SPPRSPTPPPVTLT TAPVTEPAVPVALTT AQVTEPAPAVHTAA PVTLSNAPVTEPATP ATDPATPVTRLHSPV THISCSIFTASHAPYS CAAPTSVSVACSPR RRQCSSTIAAVCNSE ASSGNPCLLALPVHR HHL PDTSPQRPGLLF HHPIPPHRPGRPPH CHGSLHRHDHRRAR RPGRQHHLPRSPSP PRSPSPSRSPSPSP SPSPSPRSTPPRSP SPTPRSPSPRPPILP PRSPDTHRSLTSHA RSPSRPATPIPVQPR HLRSTPAHPGVND AVPPSQQAIDVWL AELRSADFESFEDVC DRFVEFAAEGRNN GRPARPAHQPPRDR GNQGRPQRP RP RPGLGPEFDAQEASR LQKLYRTSKKRAITIL

294

295

296

297

NeS L	Utopia -1_PS	.	Phytop hthora sojae	AGACGAGCAACGCGCTGGGGCCCAAGACCTGGCACGAACGACACTGCC CAGGCTGTGAACGACGAGCAGCTGCTAAACCGGCGAGCGCACCGGCT CTGGGCTCGCTGCACCGTTACCGGTGCAACACGCTGGGGTCCACGTC ACGGCGGATCGGGGCGAGCGGCTGTAGCTCGCTGCTCACAGGCCTAC CCACGGCACGACCAACGACTACGCGCGGCGCATGCTCGGCTCGGCGCTG CCAGACCGCGCGCTCCCCCGGCGGCGCATGCGGCGGCGGCGGCGGCG ATGGACGTGACGAGGCGGCGGCTATGCGGAGCGGTGGGTGCTCCG TTCGACGGCGCTGCCGCGGAACCGGACCTGGTGGAGTGGTGCC GCCTTTCAAGCGGTGCGGACCGTGGCTGGACGTGCTCGCACTACA TGCCAAACAGCAGCAGACCAACCAACACCGCTGAATACAGCGCGCTGC TGCTGGCGTGACAGTCCGTCCACCGGCGCTCGCACCTGGAGAT CGAAGGTGACAGCACTCGTGTAGCGAGGTCAAGGGCACCTTCG CTGAGGAACCGCGCTCCGCGAGCTGCGCAACCGCTCGGCGCACGCG CTCCGCGGTGACGTCTTAACGCTCAAGCATCGACCGGAAGGCGA ACGGCAGCGGACAGACTGGCAACCGGCGCTGGACCTGAAGCGTT CGCTGGCTGAGTGGCGGAGCAACCAAGGGCGATGGAGAGCTGCCTCC ACATGAATCCTGTGCAACCGCAGCGCGAGCAACCGCACCTCCTGC CCGCCCGCTGCGCACCGCAGCGCGCGAGTGTGCGACGATCATGAC GAGGACATCGACGACAGATTGCCGCGCGACCGCGCGGCGAGCGGT CCCCACGTGCCTATCGGGCCGCGGAGCTCCAGCGCGGACGCGCG CCTTCGACTGCGGAGCTGACCGGAGGAGCTAGCGTGAAGATCGAAGAGC GGCCCTGCAGGCCATGGCGGAGGAGCTAGCGTGAAGATCGAAGAGC CAGACAGCTGGACCTCGGGCGATGGATACATCAGCGCGCATCCCCAGTC GCATCCGGCAGTTGCTGACGCCCTTACAGCAGCGCGCGCGCGCGCTCG GCCTCCACTGCAACAGCAGCAGCAGCGCGCGCGCGCGCGCGCGCGC TCAGCGCGAGCATCGACTCGACGAGCGCTCGACGAGATGGCGGCGGT GCAGCAGGAGCGACCCACCGAGATCGGCGCTCGGCGGTGGCGGCG	AGACGAG CAACGCGC TGGGGCCC AAGACCTG GCACGAAC GACACTGC CCAGGCTG TGAACGAC GAGCACGC TGCTAACC CGGCAGC GCACCGGC CTCTGGGC TCCGCTGC ACCGGTTA CCGGTGCA ACACGGTG GGGTCCAC GTCAACGC GCGATCGG GGCAGCG GCTGTAGC TCGCGTGC TCACAGGC CTACCCAC GGCACCG CACCAGCT	TGAAGC TGCACA AGCGCG AGGCTC GACAGC TGGACC TTCAGCT GTCGAG CCACTG CATCCG CGCCAG CCACCG CACCTG GGGCTG GTACTG CCGGCG CCACCG CGAAGG ACAACG GAGCGG CAACGC CTCGCG AGCGCC GCGTGG GTCTGG GGGGGG CCCCGG	MDVDGGPAMPEPW VLRFDGACRRNPGP GGAGAALFKPCGT AWTCSHYMIPNSSET NNTAEYTAALLGVQS AVHHGASHLEIGDS HLVVAQVKGTFCR NPRRLQLNRVRHAL RAVTSILKHIDRKAN AHADRLANRALDLR SLAECGEHQGAMES CLHMNPAARQRE QPAPPARPACAPTRA ESASDHDEDIDAEIA RDGGEAFPTLPIGP TAPARQPRRLRLQLT EDEQEAASALQAM AEELACKIEDADSWT SGDGYISAPSRQLL QPFTAAQPHPRPLQ QQRQPRPVTRTQR EHLRDEALDEMAAV QQRPTSAVRRAR RRVGRIRASMRQQ LRHDFARNESKVED ILRAASAEATAEEHPE TCPIDSGTLHEHTAV

CCGCTGCGCGGATCCGCGCTCGATGCGCAACAGCAGCTCCGCCAC GACTTCGCTCGAACGAGAGCAAGTGCCTGAGGACATTCGCGTGCT GCGTCGCGGAGACCGCAGGAGGAGACCCGGAGACGTGCCCAT CGACTCGGCGACGCTCCAGAGCACTTACGCGGCTGAATCGCCCGC ATCAACTTCTACCATGAAAGTTCGCGGCTGTTCCGCGAGCGGA TGCGGACGTCGGGACACCGAGGAGCGGGAGCGGCTCAGGAC GAGCTGACGATGAGAGGTGAGGACAGCTCATGCAAGCGGCCAC GAATCCAGCCCCGCGACGAGCGGTGGGTACGACATCTAAGAA GTTGCGGCGAGTGTCCCTCTCCACGAGCCTTCAGTCTCTGT GGCGCACACCGGTGCTGCTGTGGAAGTGGGTTCGTCGCGC TCATCCACAAGAGGCGACCCCAACGACCCAGCAACTGGCGCGCA TCTGCTGCAGACCGGATCTACAAGCTCTACAGCGGTCTGCTGGCGCG TCGCTGTGCGCGTACCTGAGGCGAACGGGCTGCTCATGGCGCA GAAGGGTTTCAGGGGTACAATGGTTGCCACGAGCACAACTTCGTGGC AACCACGCTCTGGACCAACCCGACGCTGCGCGCAGGCTGTACCG GTCTGATGCTGCGCTGCTGCGGCTGAGCGTGGTTCGTCGAGCG CTGCGACGATCTACGAGGACTCTACTTCTGCTGCGCAACGCGCC GACGAGCGACGGAACAGTGGCGGAGGAGTGGCGGTGTACCAAG CTGTCCGCTCAGCCCTCTGTTTATCGCGCGTGGTGGCGTGTCTC GAGCATGGAAGCTCGACGAGTGGCGTGGCGTGGCGGACGGG GTGCGGCTGACCAACCGCTACCGGACGACCTCAAGGTTCAGCG ACAGCGCGCGGCTACCGGCTGCCAGCTGCTGTGAGAGTCTCT GGAGTGACGCTGCTGAGGCAATCCAGGCAAGTGGCGCTTCTGCG CGTGACACGCAACGCGCGGCAACCCCGCCACGACAGGACATGAA GCTGCTGCTCCAGACGAGGAGTCTCAAGCATCAAGCTGCACGACAG CTACCGCTACCTGCGGTGGCGACGCTTCGATCAGTGGCGCACCGT CTCAGCTAGAGCCCAAGCTACAGAGATCAAGCGGAGGCGGTGGCG CTGATGAGTGGGGTGGCGGCTGAGGAGTGGTGAAGCGCTCAAG ACCTACGCTACCCCAAGTGCAGTACGCGCTCGCCACCTCCGACCGC TCCAGTCCCAACTGACGGCTTCGACCGTGTGTAGCCAAAGGCGCTCG GCACCTGCTACGCTGCGCGCTGCGCAACCAAGAGGTGCTCTACGCA CCAACGTCGAGCGGGGCTTGGGGCTCCAGCGCTGGTGGAATGCAC CGCGGCTGAGATCGCCCATGCTGCGGAGATGCTGCACTCGAAGGAC CCAGGATCCGGGAAGTGGCGCGCACAGGTGGCGAGGTGGCGCG CAAGCGTACCGGCTCCGGAAGAACTGGCGGAGCGTGACGACGA GCTGTCGGTGTCTTCACTCCGAGCTCGCGGCTCGCGGCTCGCGG GAGGCGCTGCGCGGACCGGACATCGGGTCACTGTGTCGACGTA	ACGCGCGG GCTGCTT CGGCTCGG CGTGCGCC AGACCGCG CCGCTCCC CCGCGGCG CAGCAGCA CAGCCCCC GATGCG (SEQ ID NO: 1191)	GCGCAC ATCGCA GGCTCG GGCAGC GCGGTA AGCTGG TCATTTG ACCAAC AGGCA CTACCCA GGTAGG GAACCG CCCTTCA AAAAAC CAGGAA GACACA AACACC CTCCCTT TAGTGA CATGCA TATTTA GCCTAC ATTCAG TTACGG AGAGGT TACTAAC TGATA ACACGA ACACAC AT (SEQ ID NO: 1314)	NSPRINFLPDEACGA LFREAMADVGTPE RRSALTDLTMDEVE DQLMQAATNSSPGH DGVGYDIYKFAAQL VPLHAAAFQSCWRH HRVPALWKVGFVRLI HKGDNDPANWR PICLTAIYKLYSGLLA RRLSAYLEANGLLM AQKGFAYNGCHEH NFVATLLDQTRRM RRRLYQVWYDLRNA FGSVHQDMLWVVL LLGVERAFVERCDDIY EDSYFVVGNAADGA TEPVRQEVGVYQGC PLSPLLFAALVPLRA LEKLDGVGVADAGV RPCTTAYADDLKVES DSAAGITRCHAVEK FLEWTVLQANPGKC AFLAVTRNARGNPA HDKDMKLSLHDEEVS SIKLHDSYRYLGVGD GFDHVRHRLQLEPKL QQIKREAVALMQSG LAPWQVVKALTYVY PKVEYALRHLRPLQS QLQGFDRVAKGLR HLLRPRSATNEVLYA PTSSGGLGLQLVEM HRAHQIAHAWQML HAKDPAIREVARAQV WQVARKRHLREEH WRERDELVRCLFNS ELAASPHAEALRRHG DIGSLWSDVQRWLRI

				CAGCGGTGGCTGGCGCATCTACCACTCAGCCTGGTGGTGCGAGCAGCAC CGCAACGGCCTGGACCCGCTCGGCCCTCGCGCTGCCGACACACGCCAAGT GGCTGACCAACAAGAGCGTCTGCGGCACGTCAAAGTTGCACCTCAAGA TCCGCCACAGACCCGATGGAAGGTTCTGCGGACCAAGGCAAGACG GTGCGGCCACGGCGTGTGCGGTGCAAGTTCTATGTCGACATGGCG GGCTGTGCGACGACACTACCGTTCTGGCGTGAAGCCCGCTCAACC AGATCGACACGAACCGCTGCTCAAGCGCAAGCGGCTCGGTTCGCACA AGACCTGCAGGGACCGACCTGCTCTCGGCCGAGACGCTGGCCCATGT CCTCAACCATTCGAGTCCAACATGGACGCAATCCGCCAGCGTCATGAC GACGCTTGAGCAGATCGGTTCAAAGATCCGCAACGCGCTCAAGCGC GGCAAGTCGACCGGAGTCGCTGAACCAAGCATACCGGTACCCGAGTAC ACGGGCGAGCTCTGCGGCCAGACATCGTACTGCGGATCTGGCGGCC AAGAAGATGGTATCGCCGACTTGGCGTCACTTGAAGAGCACGCG GCCGGCGCCCACTCTCGTGCAGCTCAGCCACGACCAACAAGACAC TCAAGTACCAACCGATCGTGGGGAGCTGCAGCTCAAGGGTGGCAGG TGAGACCGCGGGATCGTGACGGTCTGCTGGCTCGGTGCGAGCCCA GCAACTCAACACTACGGAAGATTGAAGCTGCACAAGCGCGAGGC TCGACAGTGGACCTCAGCTGTGAGCCACTGATCCGCGCCAGCCAC CGCACTGGGGTGTACTCGCGCGCCACCGCAAGGACAAACGGAGC GGCAACGCTCGCAGCGCGCTGGTCTTGGGGGGGGCCCGCGCGG CACATCGCAGGCTCGGCGACGGCGGTAAAGCTGTCATTTGACCAACAG GGCACTACCCAGGTAGGGAACCGCCCTTCAAAAAACCAAGGAAGACACA AACACCTCCCTTTAGTGACATGCATATTTAGCTACATTTTCAGTTACG GAGAGGTTACTAACTGGTAAACACGAACACACAT (SEQ ID NO: 1068)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
NeS L	Utopia -1_PU	ADOS0 10013 21	Pythiu m ultimu m	GCGGTGACGCGCAACCGCGGCTCTTCGAGTGCACTGTCGCGCG ACACCGCGGGGATCTCGCGCGCTCCAGCAGCATCGGCGCTCCGCGC ACCGCAGCTCCGCTTTGTGGATCACTTCCACAGCGGATGCGCGTGCGG CGTGAGCTTCCACTCGCTGCGGCGGCAACCAAGCACGCGCGCGAAATG TCCAGAGAGCGGTTCTCGGTGCGCGCGCGCGCACTGCAGCGGC CAACACCGCAGGTATGCTCTCGGCGCGGACGAACGCAACCACTCGCG TCCGTGCGGCTTGATGATGTTGCATCCCCGCTTTTGGCAACATTTT GCCGTTGCTTTCGACCGCGGACAGCATCAAGCGCCACCGTGAT CGCAGACGACCCATGCAGCACAGTGTGTGCCCTCTGCTGCCGCCAA TCCCTCGCGGTGCGCACGTCCCCCGTGGCGCGCCACCAACGACAC CATCCGCGCTGCGGATTGGTGGCAACGCCCGCCCTGAACGACGACG GCGACAACGAAACAGCGACGCGCGCAGCCGACATCGAGATGCGC GCCGACGACACCGACGACGCGCCGACCAACCCCGCACCAAGTGGC GCTGCAACGCCCGCGCACAGCACCAACGACCAACCGGATGCCGCGACC																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						

GGCGCCGCGCCACACGCGATGGGGACCCAGAGTGCTCGTGAACGTG	TTGTGGAT	TAGTGA	SQVRGTFACNRRLR
GGCGCCTTCGCGCCGCACTACACAGCTGGGGCCCGCACTCCGCC	CACTTCCA	CATGCA	KLRNRVQAALRELGD
GGACCAACAGCATGCAGCAAACTGAGAACTCGAATCAGTGTCTGG	CAGCGGAT	TATTCAA	YRLVHIDRQANAHA
TGGCCGACGGACAATGCAGCGATGGCGGGAACGCGAGCGCGACG	GCGCGTGC	ACCGAT	DRLANRALDLRKTIV
CCGAGCGGACATCCGCGCAACGAGCGGACTGCAACGGCGACGGCG	GGCGTGA	ACATATT	DCGPHATTTDACVQ
ACGGGACGACCGGACAGCGACGACGCGACGCGCAACGACGCGCGA	GCTTCCAC	CGTTAG	PAEILAPTARLSSSSS
CGACGACGACGACGCGAGCGACGCGGCGACAAACGACGACGCGC	TCGCGTGC	GAGAGG	SSSSSDEPMPGLEEP
GCCGACGACGACGACGACGACGACGACGACGACGACGACGACGAC	GGCGGCA	TTACTAA	AADDETDADAEDIA
GACGCGGCGACGACGACTCCCTGGTCTTGCGCTTTGACGGCGCA	ACCAAGCA	CTGGTA	MRDGGEIFPTLQIGP
TGTGCGCGAAACCGAGCGCGCGCGCGCGCTGCTGCGCTTTCGAGC	CGCGCGCG	ATATATC	GSAPAQQPRLRLRL
CAGTGGCGCGTGTGTGACGGTGTCTCACTACCTCCCGGGCGAGTG	AATGTCCA	ACCATTT	SDDESEAAARTLEHF
AGACCAAAACACCGCGAGTACTCGGCGATGCTGCTCGGGGTGCGCA	GAGAGCG	C (SEQ ID	ANDMASKIADADD
GCGCCATCCACACGCGGACCCGACTCCGCGTGAAGGGCGACAGCC	CGTTCTCG	NO:	WRSGEYISAIPVRLR
ACCTCGCGCTCTCGAGGTGCGCGGCACTTTCGCTGACGAACCGGCG	GTCGCCGC	1315)	ELLAPYAVPIRSPRN
TCTGCGCAAGCTGCGCAACGAGTCAAGCGGCACTGCGCGAGCTCGG	CGCCGCGC		ASSRPPRPQSRPPRP
CGACTACCGGCTCGTGACATCGACCGGCAAGCAACGCGCATGCCGA	GCACTGCA		PRVTRHQREHRLDEA
CGGCTCGGAACCGGCTCGACCTCCGGAAGCAACGAGTGCAGTGC	GCGGCCAA		LDDLAQAQRSTSDQ
GGGCCCCAACGACACCGACCGACGCGTGTGCGAGCGCGGGAATC	CACCGCAG		RSIRNARRRVGRIATA
CTGGCGCAACAGCCCGCTGTCCAGTCCAGTCCAGTCCAGTCCAGTCCA	GATATGCT		QAQSDLRSQFATNE
GCTCAGCGACGAGCCGATGCCGGACTCGAGGAGCCAGTGCAGAC	CTCGCGCG		RACVESILRAAKPDG
GACGAGACGACGCTGACGCGGAAGCCGACATCGCAATGCGCGATGGA	CGACGAAC		TEPQASAGTCPIDRA
GGCGAGATTTTCCGACGCTCCAGATCGGCCCGGCTCGGCGCCAGCAC	GCGACCAC		TLHAHFAGVNTPRER
AACAAACCGGCTCGCTCCGCGAGTCTCGGACGACGAGTCCGAGG	CTCGCGTC		FDFFDALGADFRAAL
CAGCCGCGCAGCTCGAGCACTTCGCCAACGACATGGCGAGCAAGA	CGTCGGCG		DVLPPDQQAADAF
TTGCGGACGCTGACGACTGGCGTTCGGGGAGGGCTACATCAGTGCCA	CCTTGCA		DELSGEVEDQLDRV
TCCCTGTGCGCTCGTGAATCTCTGCACTATGCGGTGCCAATCAG	GATGTTGC		VASSPGLDGVGYDV
GAGCCCTCAGGAAACGAGTCCCGTCCCGCGCCCGCAATCTCGT	ATCCCCCG		FKRFRLLPLHAAAY
CCGCTCGCCCGCGGAGTGAATCGACACCGAGCGGCAACACCGACTG	CTTTTGGC		QCCWRHRRVPATW
GATGAGGCGTCAACGACTGGCGTGGCGAGCGCTCGACATCGAC	AACATTTT		KVGLVRLHKKGDRA
GACGAGCGTCCATCCGCAACGCGCGGCTCGAGTGGCGCGCATCCGCA	GCGCGTTG		EPNNWRPICLQQAII
CGGCAAAAGCCCAATCCGACTCCGAGCGCAATTCGCCACGAACGAGCG	CGTTTGGC		KVYSGLLARLSRWL
CGCTGCGTCGAGAGCATCTCGCGCGCCCAACCCGACGCGCACCCGA	ACCGCGGC		EANERFTTAQKGFRE
GCCTCAAGCAAGCGCGGCGACGTGCCCATCGACCGTGGACGCTCCA	AGACGCA		FNGCHEHNFVASSLL
CGCGCATTTTCGCGGCGTCAACACGCGCGCGAGCGCTTCGACTTTGAC	CAAGCGCC		DQTRRLHRKLYAVW
GACGCGTGGGAGCCGACTTTCGCGCGCGCTCGACGCTTCGCCACCGC	ACCGTGAT		YDLRNAFGSMPQPL
CGGACCAAGCCCGGATGCGTTGCGCGACGAGCTCTCGTGGCGGAGG	CGCAGACG		MWRVLRARLGVDTAF
TCGAGGACGAGCTCGACCGCGTGTGCGCTCTCCAGTCCCGGCGCTGGA	CAGCCATG		LQRCEDIYADSEFFW
CGGAGTGGGCTACGACGTTTCAAGCGCTTTCGGCTGCGAGCTCCTGCCG	CAGCACAG		GNAADGATDPVRQE

CTCTCCACGCTGCATACCAAGTCTGCTGGCGCCACCGCGCGTGCGCTG	TGCTGTGC	VGWVQGCPLSPLLFIS
CAACGTGGAAGGTGGGATTGGTGGCGCTGCTGCACAAGAAAGCGGATC	CCTCTGCT	ALIPLLRALQRLPGVG
GCGCTGAGCCAAACAACACTGGCGCGGATCTCTTGACGAGGCCATCTA	GCGGCCCA	VPLADGVRPCTTAYA
CAAGGTCTACTGGGCTGCTGGCGCGGACTGTCGCGGTGGCTCGA	ATCCCTC	DDLKVFSDSAAGIQQ
GGCGAACGAACGTTCAACACCGCCAGAAAGTTTCGGGAGTTCAA	GGCGTGC	CHGTVARFLRWTL
CGGTGCCAGAGACAATTTCTGGGATCTCTGCTCTCGACCAAGCG	GCACGTCC	RANASKCALPVTTT
CGTCGGCTCCATCGGAAGCTGACGCTGCTGTGACGACTTGGGAACG	CCCCGTG	ARGNPAIDDTLQLEL
CCTTCGGTTCAATGCCGAGCCCTGATGTGGCGCTCTCGCTCGCCTC	CCGCGCGC	HGDAIARLTQSSYAY
GGAGTCGACACTGCTCTCCAGCGGTGCGAGGACATCTACGCGGAC	CACCACGA	LGVDGDFHVQHVRV
TGCTTTTCTGCTGGCAACGACCGACGGTGCCACGGACCCCGTCC	CACCATCC	QLAPKLAELKRDAVA
GACAGGAGGTGGGCTTACCAAGGATGCCGCTGAGCCCGTGTCTGT	GGCTGTGC	LLRSGLAPWQV/LKAI
TCATCTAGCAATTGATCTCTCTGCGGCTCTGCAGCGGCTACCAAGC	GATTGGTG	KVLYPRIEYALRHLR
GTGCGGTTCCCTGGGACGAGTCCGTCCGTGTACGACCGCGTAC	GCAACAGC	PLQSQLEGFDRVAK
GCGGACGACCTCAAGGTGTTCAAGCACAGTCCCGCGGCATCCAGCAA	CGCCGCTT	GFRHLLRLPANATNE
TGCCACGGCACTGCTGCTGATTCTCCGATGGACGGGCTGCGCGCGA	GAACGAC	LLYAPVSSGGGLLPL
ACGCGAGCAAGTGTGCGTGTGCGAGTGACAAACCACTGCGCGCGCA	GACGGCG	VELHKALQIAHWQ
ACCCGCCATCGACGACACGCTGCAGCTGCAGTGCACGGCGACGCCAT	ACAAACGA	MLHSKDAAVQAIAR
CGCCGCTCACCTCCAATCGAGCTATGCTACCTTGGGTTGGTGAT	AACAGCGA	AQVRQVWQKRYTLD
GGCTTCGACCACGTGACGACCGCGTGACGCTCGCGCCCAAGCTCGCA	CGCGCGCG	ADHWQGRDDDELVQ
GAGCTGAAGCGGACGCTGTGGCACTGCTGCGGTGCGGTTGGCGCT	ACGCCGAC	LFLNSELAASPHATIK
TGGCAGGTACTGAAGCCATCAAGGTCTATCTGACCCGCGCATCGAGT	ATCGAGAT	RRNGDIGSLWSDVQ
ACGCGTGGCCACCTCCGACCGCTGCAGTCGACGCTCGAGGGTTTGA	GCGCGCGG	RHLKTLQLRLETREPT
TCGCGCGGTAGCAAGGGGTTCCGACACCTCTCCGTACCCCGGAAC	ACGACACC	ADAPDSPNGLLHLRV
GCCACCAACGAGTGTCTACGGCCTGTGTCGAGCGGCGGCTGGGG	GACGCACC	PHHRKWLSHKTVLR
CTTCTGCCACTGCTGAGCTCCACAAAGCTCTGCAGATTGCCACGGGT	AGCGCCGA	HMKLHRLCHKHKW
GGCAGATGCTCCACTCGAAAGACCGCGCGTGCAGGCGATCGCGCGG	CCAACCCC	QSMSDQGRTVRAH
CACAGTCCGCCAAGTGGTCCAAAGCGCTACACGCTGGACGCGGACC	GCGACCAAG	GQAGSHFVSRGVGL
ACTGGCAGGGACGAGACGAGCTCGTGCACTCTTCTCAACAGTG	TGCGGCTG	WDADYRFALQARLN
AGCTCGCGCTCGCCACGCGACCATCAAGCGCGCAACGGAGACA	CAACGCCC	QLDNTSTLKRRTQRT
TTGGTTGCTCTGGAGCGATGTCAGCGCACTGAAGACGCTGCAGCT	GCGCGCAC	NATCRAPNCSRTETL
CCGGCTCGAGACGCGGAGCCGACTGCGGACGCGCTGACTCCCCCAA	AGCACCAAG	AHVLNHCETNMDVI
CGGCTCTCTCACTCCGTGTGCGGCAACACCGCAATGGCTCAGCCAC	CACCAACG	RQRHDSGALEQIGAAI
AAAAAGGTGTTGCGGCACATGAAGCTCCACATCCGCTGTGCCACAAGC	GATGCGCG	NAAIKGRRTDTEVRL
ACAAAGTGCGAGAGCATGAGCGACCAAGGGGCGCACGGTGCGAGCGCAC	GACCGCGC	NQTVPEFNGPAWRP
GGACAGGCTGGCAGCCACTTGTGTGCGGCGGGTGGGCTATGGGAC	GCCGCGCG	DIQVRDARSKTMVIA
GCGGACTACCGCTTCGCGCTCCAGGCCGCGCAACCAATGCCACCTGCGCGC	CACACGCG	DLAITFEDQPNDQSA
ACTCGACACTCAAGCGTCCCGCGCAGCGCACCAATGCCACCTGCGCGC	(SEQ ID	SSSLQHSREHIAKYQ
GCCGAACTGCTCGCGCACCGAGACACTGGCGCACGCTCCTCAATCATTCG	NO: 1192)	PIAAALERQGWVRVH

303

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 303

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 303

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHIER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

What is claimed is:

1. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain, wherein one or both of (i) or (ii) are derived from an avian retrotransposase; and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence.

2. The system of claim 1, wherein the avian retrotransposase is a retrotransposase from *Taeniopygia guttata*, *Geospiza fortis*, *Zonotrichia albicollis*, or *Tinamus guttatus*.

3. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain, wherein at least one of (i) or (ii) is heterologous, and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide and (ii) a heterologous object sequence.

4. A system for modifying DNA comprising:

(a) a polypeptide or a nucleic acid encoding a polypeptide, wherein the polypeptide comprises (i) a reverse transcriptase domain and (ii) an endonuclease domain; and

(b) a template RNA (or DNA encoding the template RNA) comprising (i) a sequence that binds the polypeptide, (ii) a heterologous object sequence, (iii) a first homology domain having at least 10 bases of 100% identity to a target DNA strand, at the 5' end of the template RNA, and (iv) a second homology domain having at least 10 bases of 100% identity to a target DNA strand, 5' end of the template RNA.

5. The system of any of the preceding claims, wherein (a) comprises RNA and (b) comprises RNA.

6. The system of any of the preceding claims, wherein (a) and (b) are part of the same nucleic acid.
7. The system of any of claims 1-5, wherein (a) and (b) are separate nucleic acids.
8. The system of any of the preceding claims, which comprises only RNA, or which comprises more RNA than DNA by an RNA:DNA ratio of at least 10:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, or 100:1.
9. The system of any of claims 1-8, wherein the heterologous object sequence comprises an open reading frame in a 5' to 3' orientation on the template RNA.
10. The system of any of claims 1-8, wherein the heterologous object sequence comprises an open reading frame in a 3' to 5' orientation on the template RNA.
11. The system of any of the preceding claims, wherein sequence that binds the polypeptide is a 3' untranslated sequence.
12. The system of claim 11, wherein the template RNA further comprises a 5' untranslated sequence.
13. The system of any of the preceding claims, wherein the template RNA further comprises a promoter operably linked to the heterologous object sequence.
14. The system of claim 13, wherein the promoter is disposed between the 5' untranslated sequence and the heterologous object sequence.
15. The system of claim 13, wherein the promoter is disposed between the 3' untranslated sequence that binds the polypeptide and the heterologous object sequence.

16. The system of any of claims 11-15, wherein the 5' untranslated sequence is a sequence of column 6 of Table 3, or a sequence having at least 80% identity thereto.
17. The system of any of claims 11-16, wherein the 3' untranslated sequence is a sequence of column 7 of Table 3, or a sequence having at least 80% identity thereto.
18. The system of any of the preceding claims, wherein the heterologous object sequence comprises an enzyme, a membrane protein, a blood factor, an intracellular protein, an extracellular protein, a structural protein, a signaling protein, a regulatory protein, a transport protein, a sensory protein, a motor protein, a defense protein, or a storage protein.
19. The system of any of the preceding claims, wherein the template RNA comprises at least 10 bases of 100% identity to a target DNA strand, at the 5' end of the template RNA.
20. The system of any of the preceding claims, wherein the template RNA comprises at least 10 bases of 100% identity to a target DNA strand, at the 3' end of the template RNA.
21. A method of modifying a target DNA strand in a cell, comprising administering the system of any of claims 1-20 to the cell, thereby modifying the target DNA strand.
22. The method of claim 21, which results in the addition of at least 5 base pairs of exogenous DNA sequence to the genome of the cell.
23. The method of claim 21, which results in the addition of at least 100 base pairs of exogenous DNA sequence to the genome of the cell.
24. The method of any of claims 21-23, which results in insertion of the heterologous object sequence into the target DNA at an average copy number of at least 0.01, 0.05, or 0.5 copies per genome.

25. The method of any of claims 21-24, which results in about 50-100% of insertions of the heterologous object sequence into the target DNA being non-truncated.
26. The method of any of claims 21-25, wherein the nucleic acid of (a) is not integrated into the genome of the cell.
27. The method of any of claims 21-26, wherein the template RNA comprises at least 10 bases of 100% identity to the target DNA strand, at the 5' end of the template RNA.
28. The method of any of claims 21-27, wherein the template RNA comprises at least 10 bases of 100% identity to the target DNA strand, at the 3' end of the template RNA.

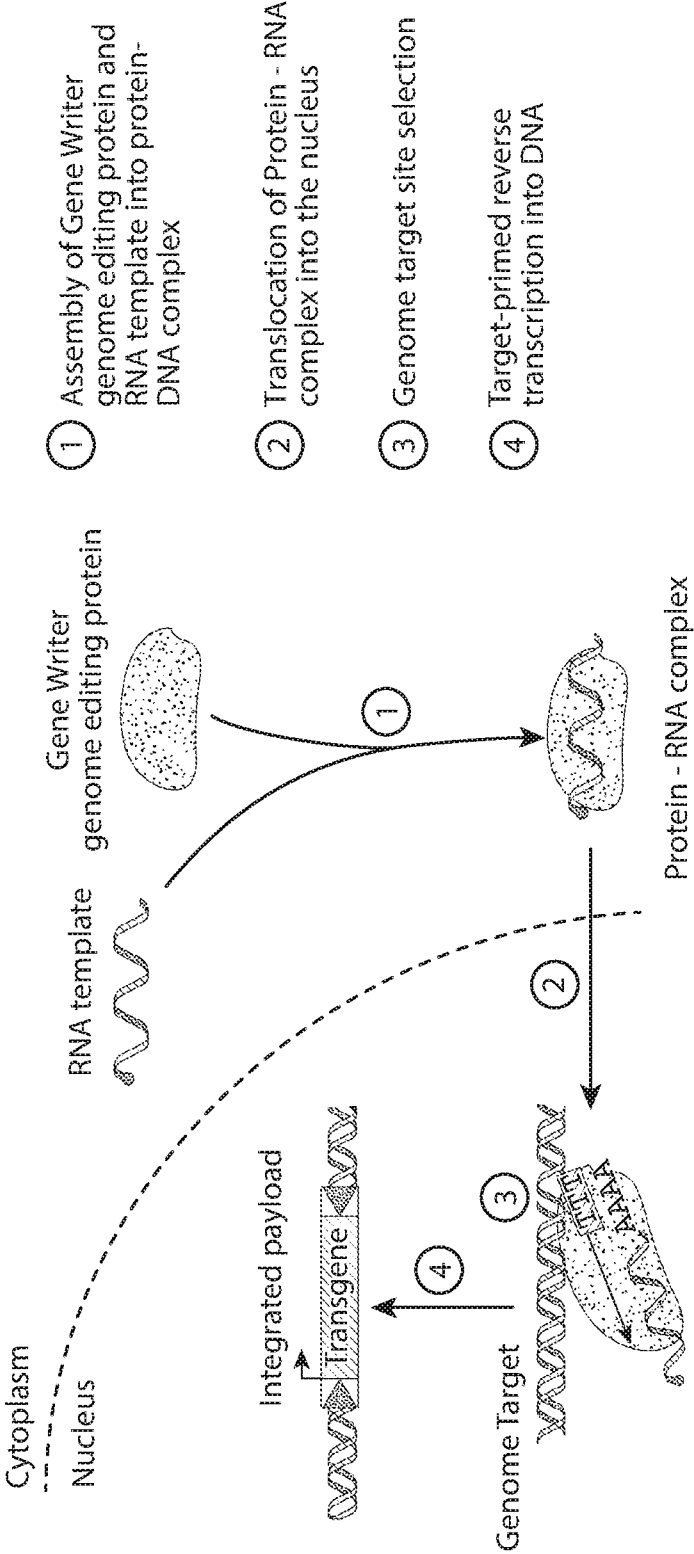


Figure 1

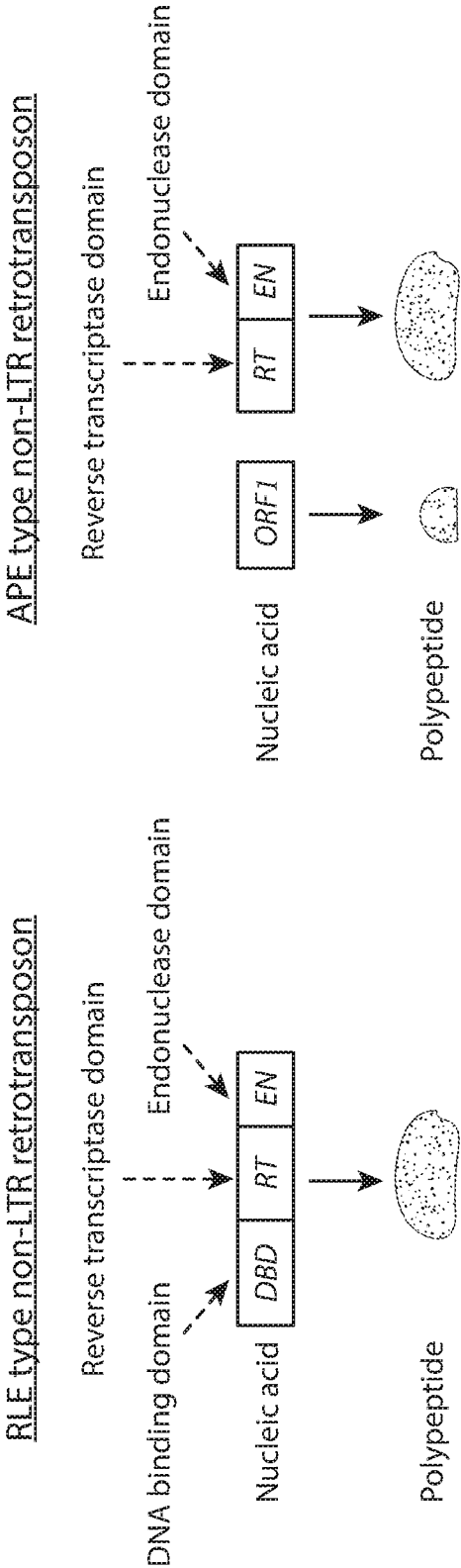


Figure 2

Gene Writer genome editor proteins with different DNA binding domains target different DNA sites for Gene Writing genome editing

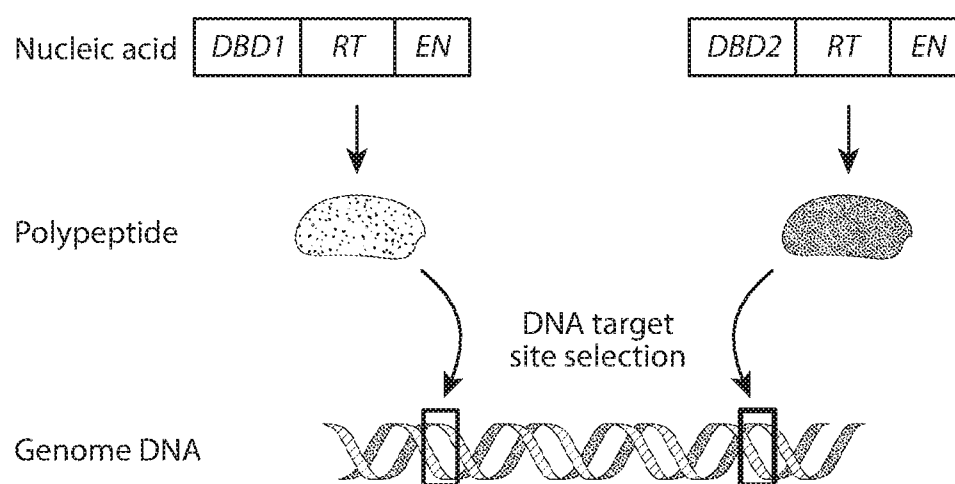


Figure 3

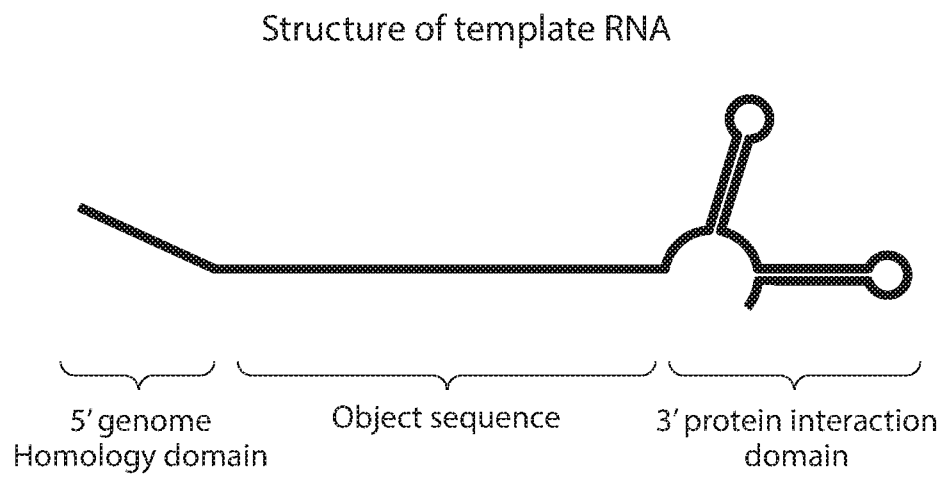


Figure 4

5/24

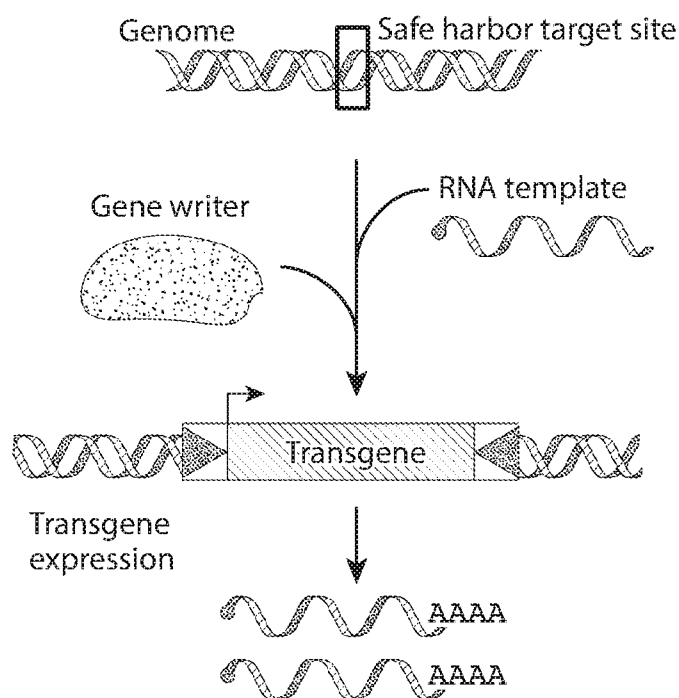


Figure 5

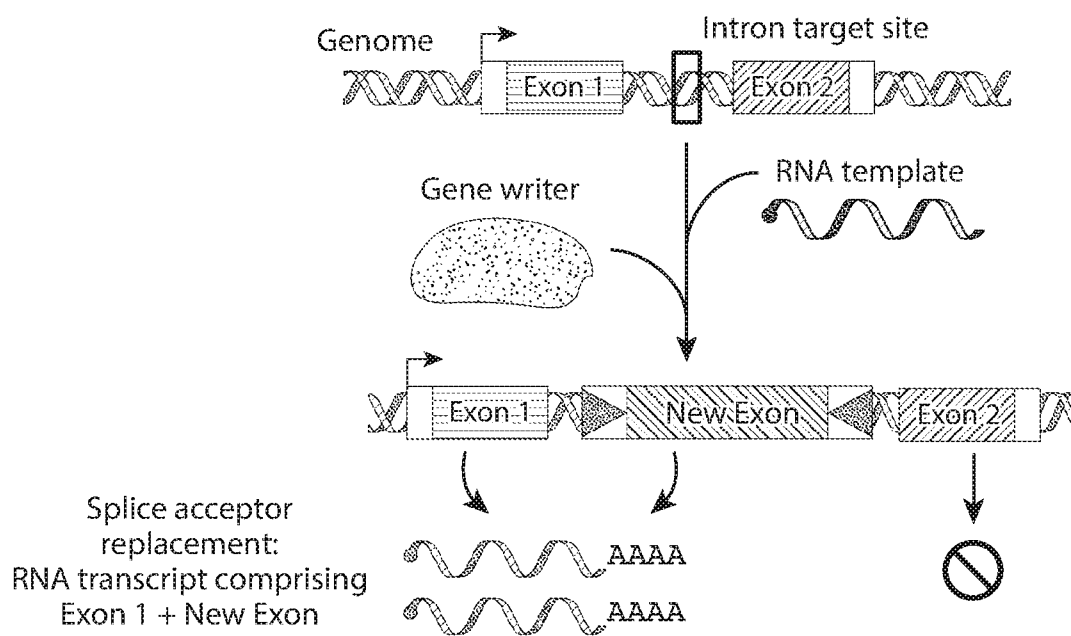


Figure 6

6/24

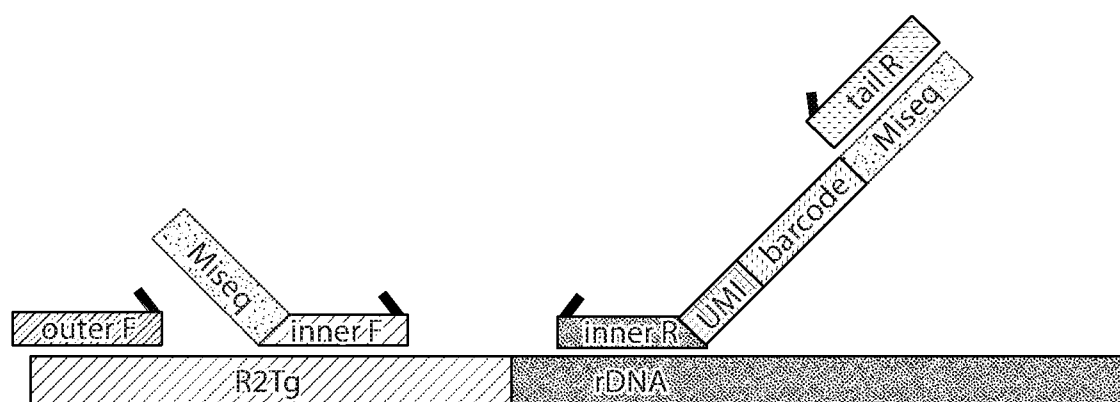


Figure 7

7/24

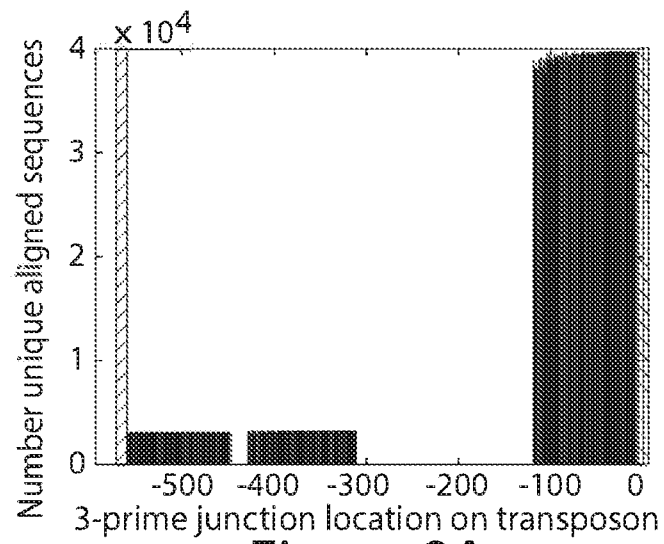


Figure 8A

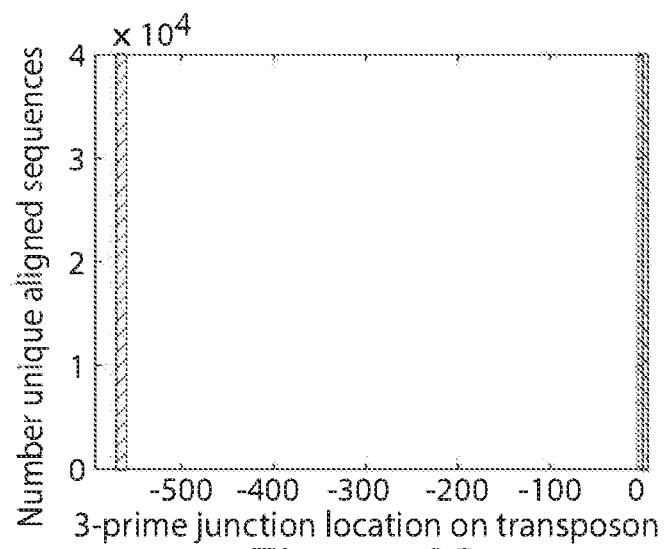


Figure 8B

8/24

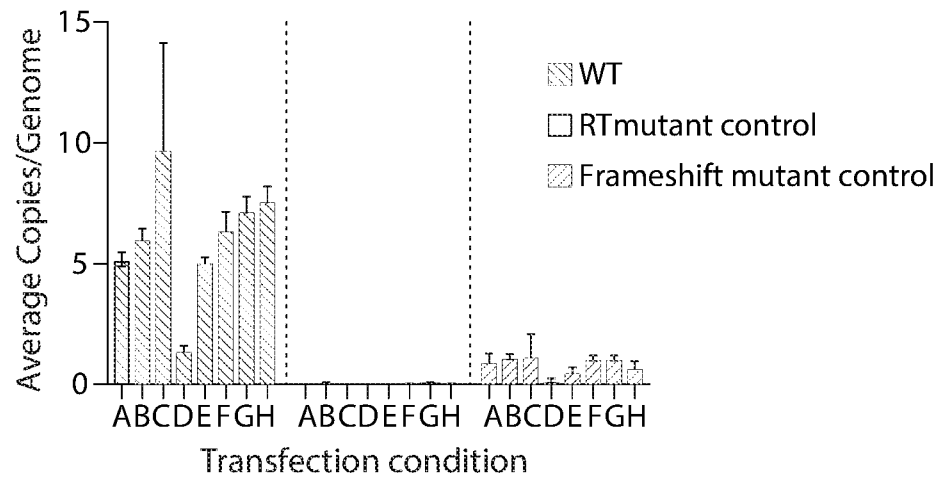


Figure 9

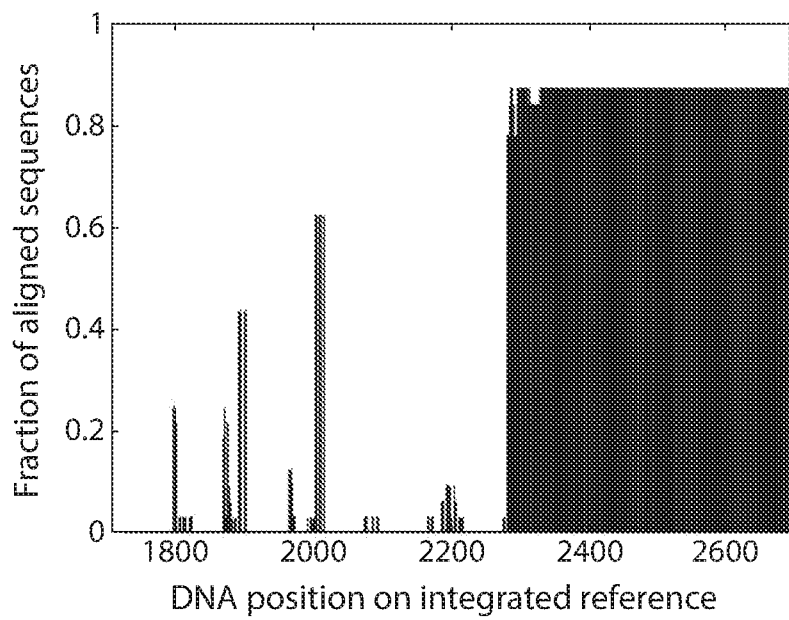


Figure 10

9/24

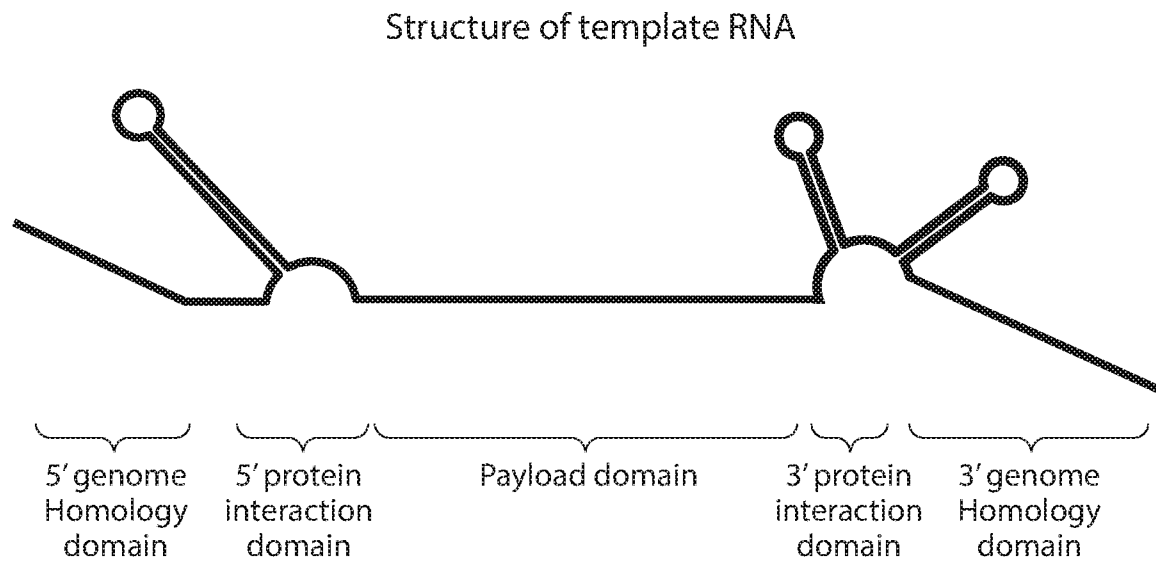


Figure 11

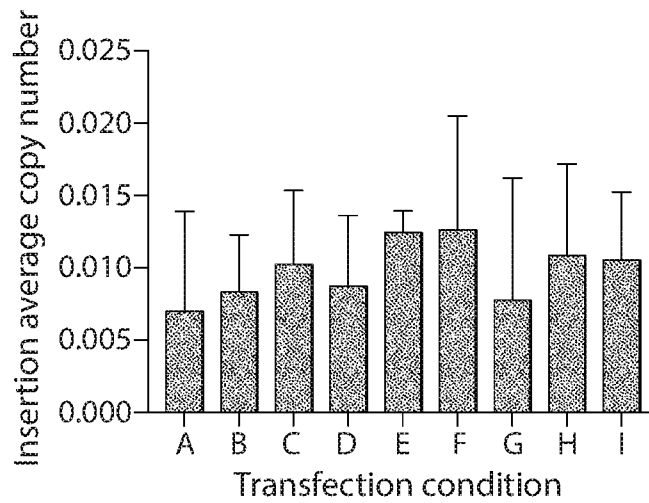


Figure 12

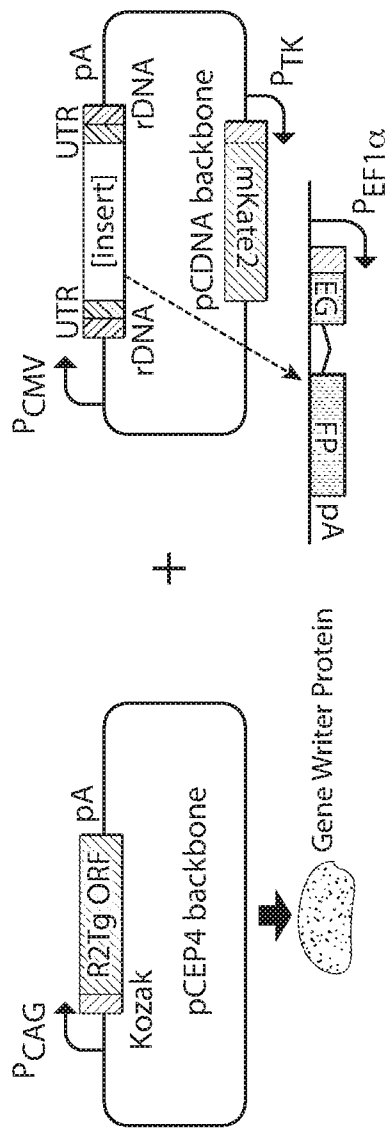


Figure 13

11/24

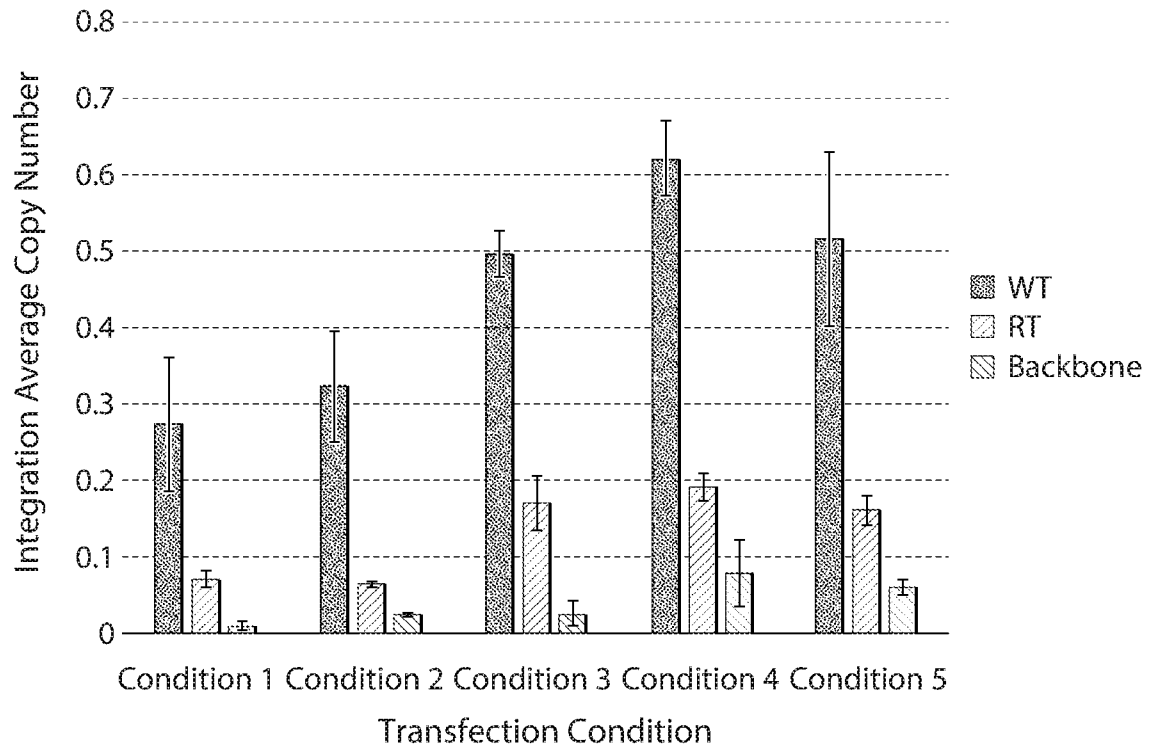


Figure 14

12/24

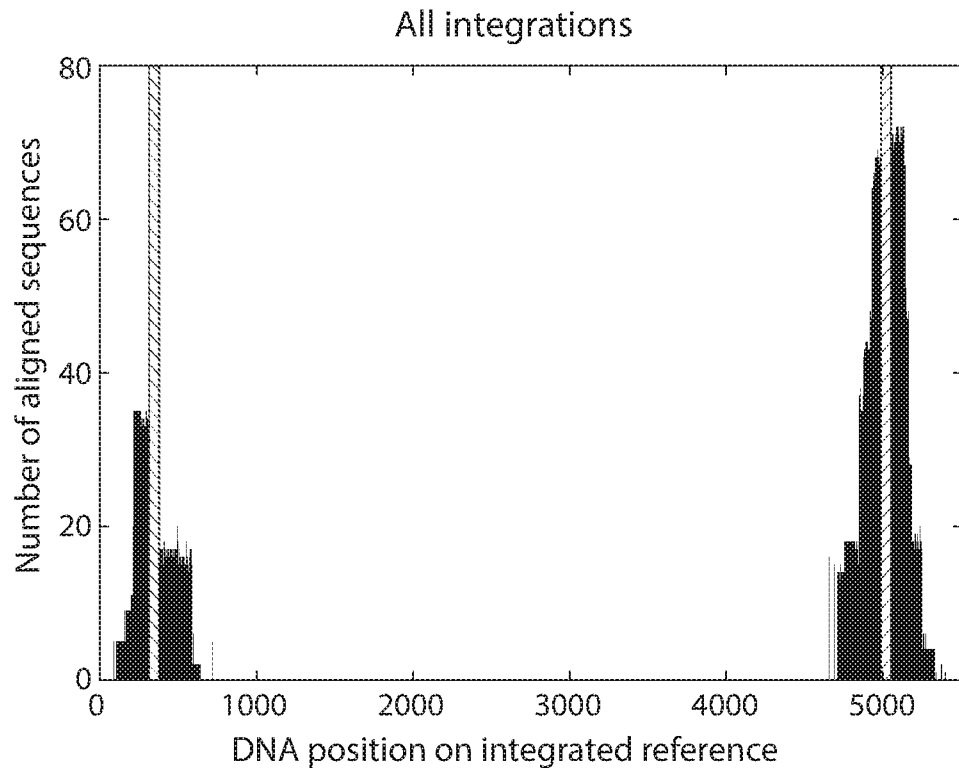


Figure 15A

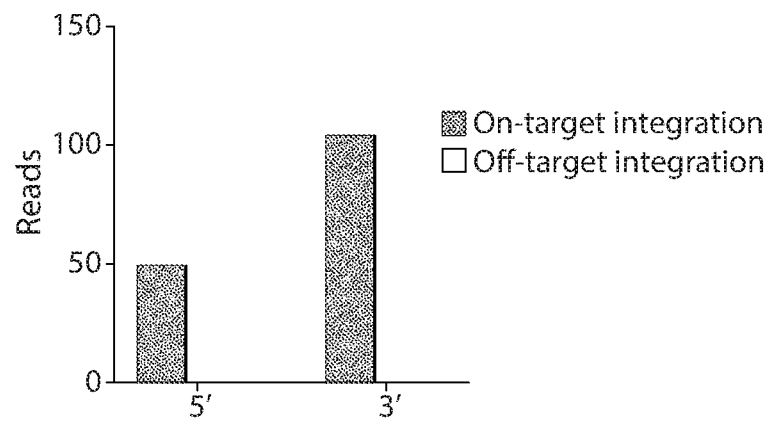


Figure 15B

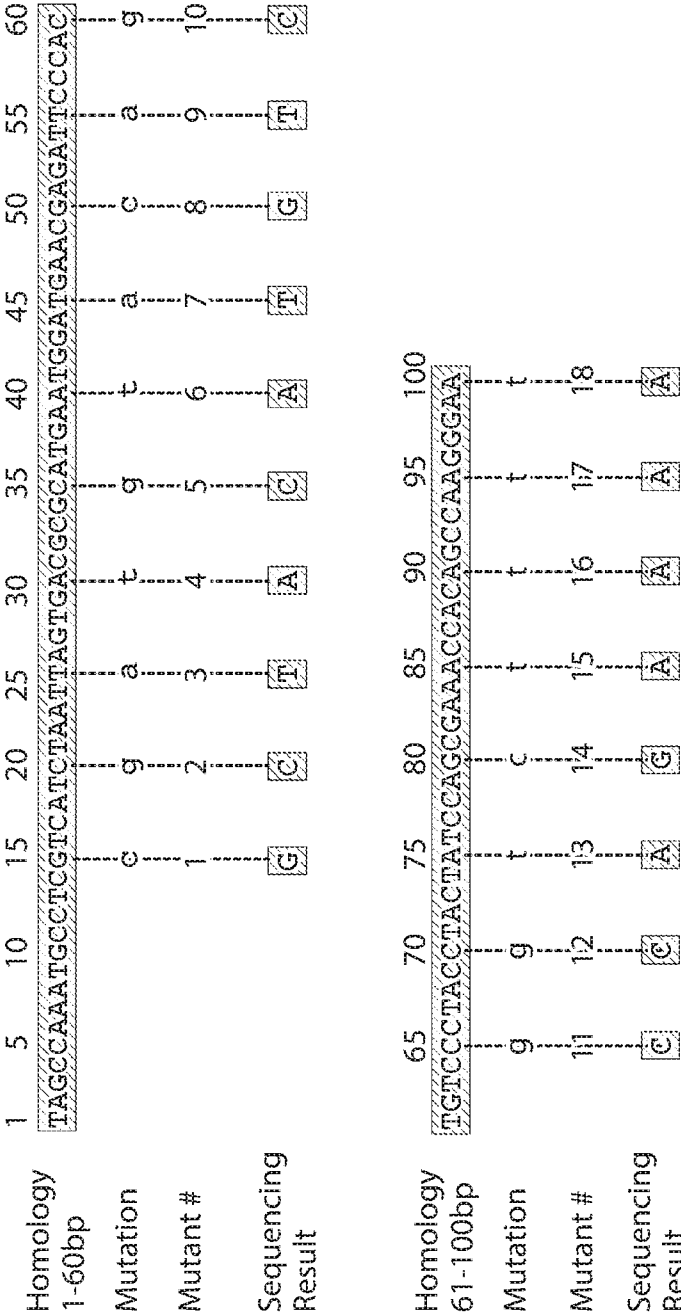


Figure 16

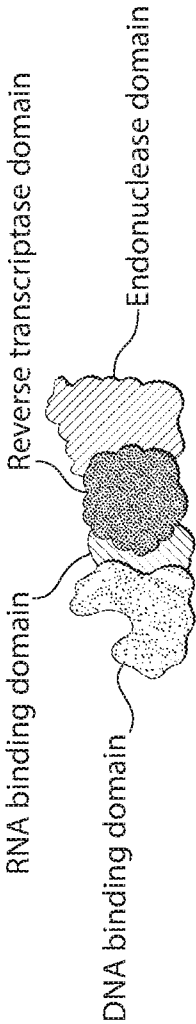


Figure 17A



Figure 17B

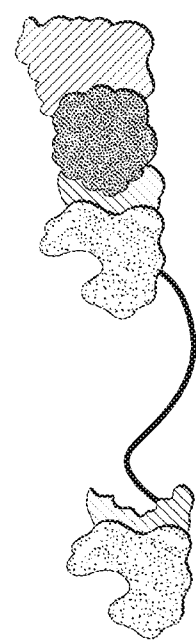


Figure 17C



Figure 17D



Figure 17E

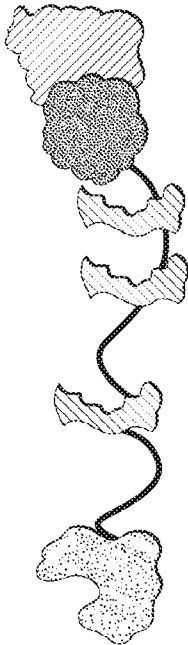


Figure 17F

15/24

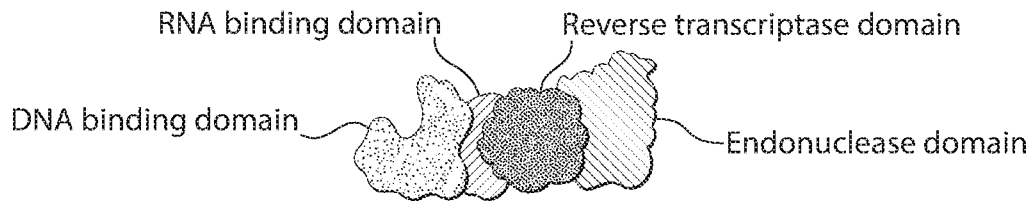
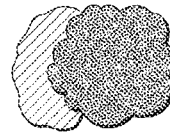


Figure 18A



DNA binding
Zinc fingers
Cas9
Transcription factors

Figure 18B



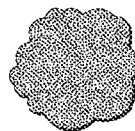
Full-length Reverse transcriptase
Different full-length RT from other R2
Different full-length RT combined with additional RNA binding domain

Figure 18C



RNA binding
B-box protein
MS2 coat protein
dCas protein
UTR binding protein

Figure 18D



RT without RNA binding
Truncated RT from R2
HIV RT
AMV RT

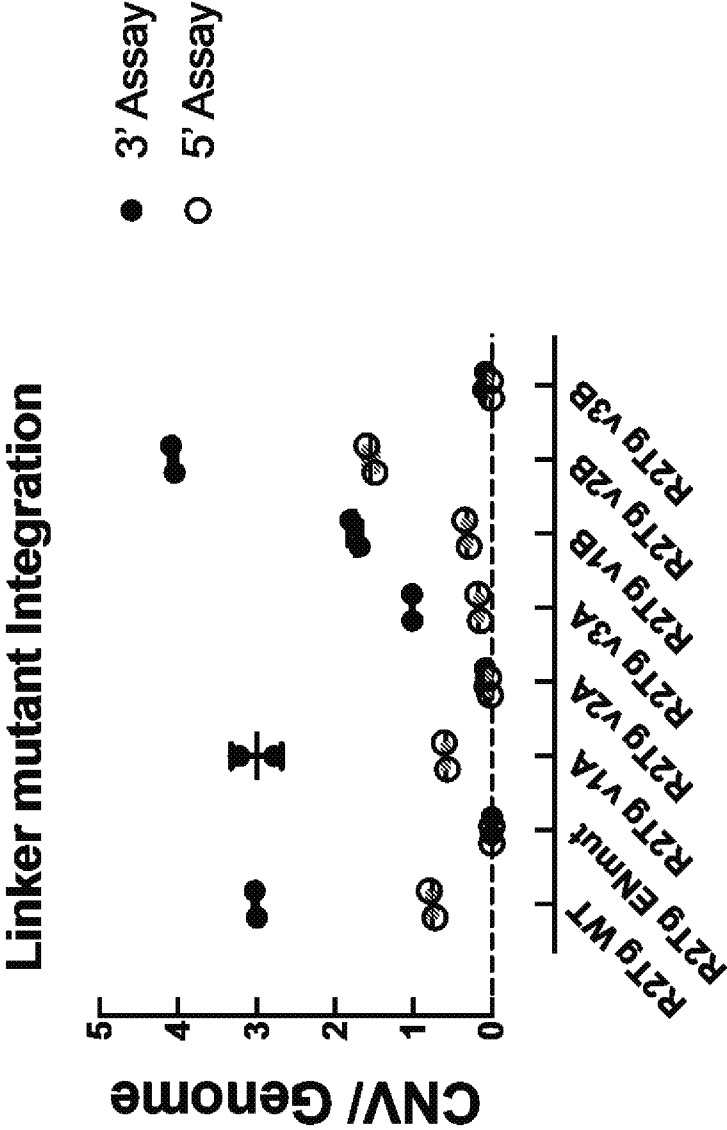
Figure 18E



Endonuclease
Cas9 nickase
Cas ortholog
FokI
Restriction enzyme

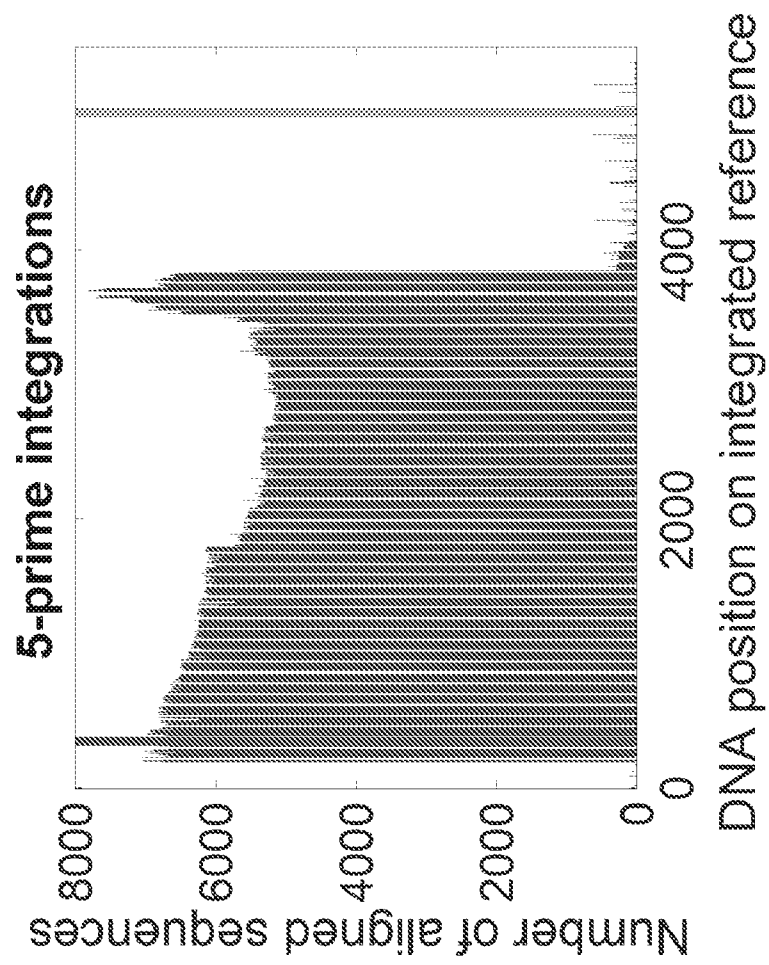
Figure 18F

FIG. 19



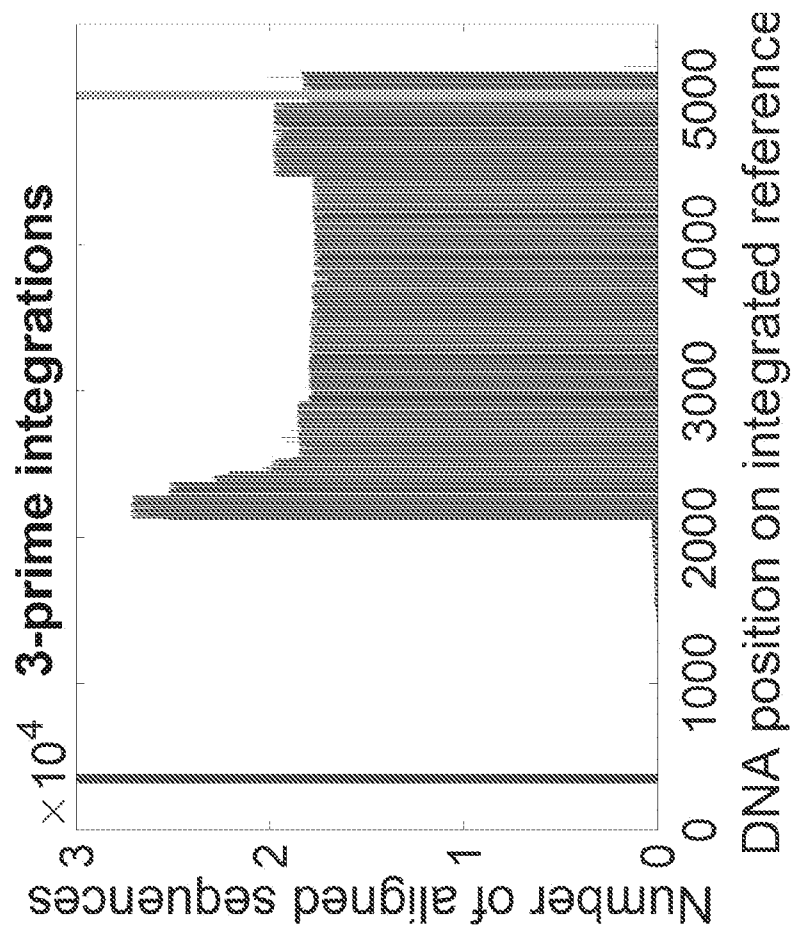
17/24

FIG. 20A

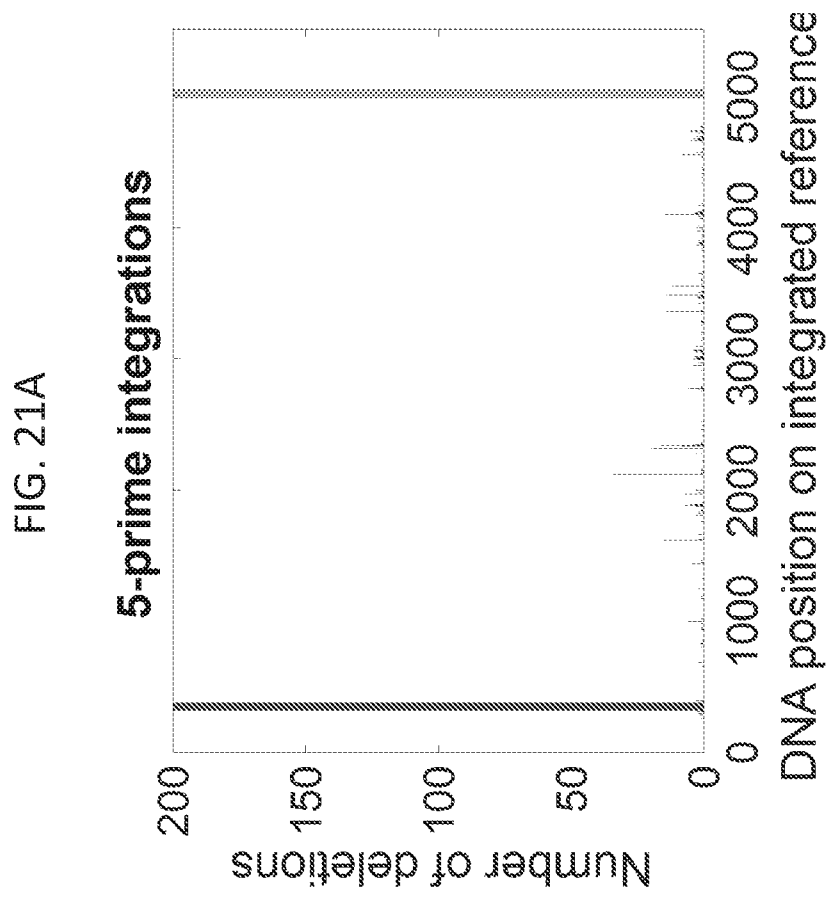


18/24

FIG. 20B

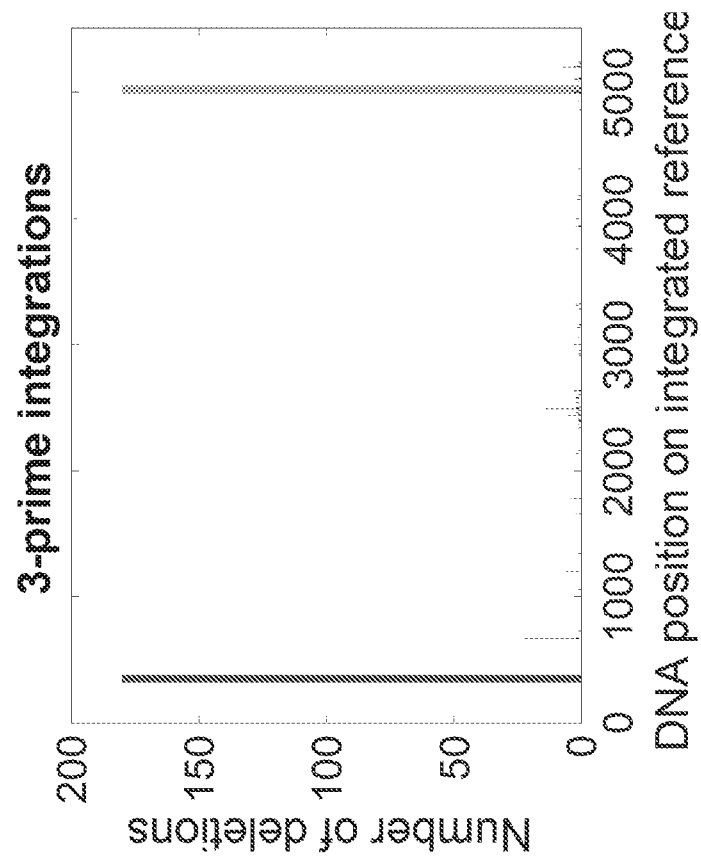


19/24

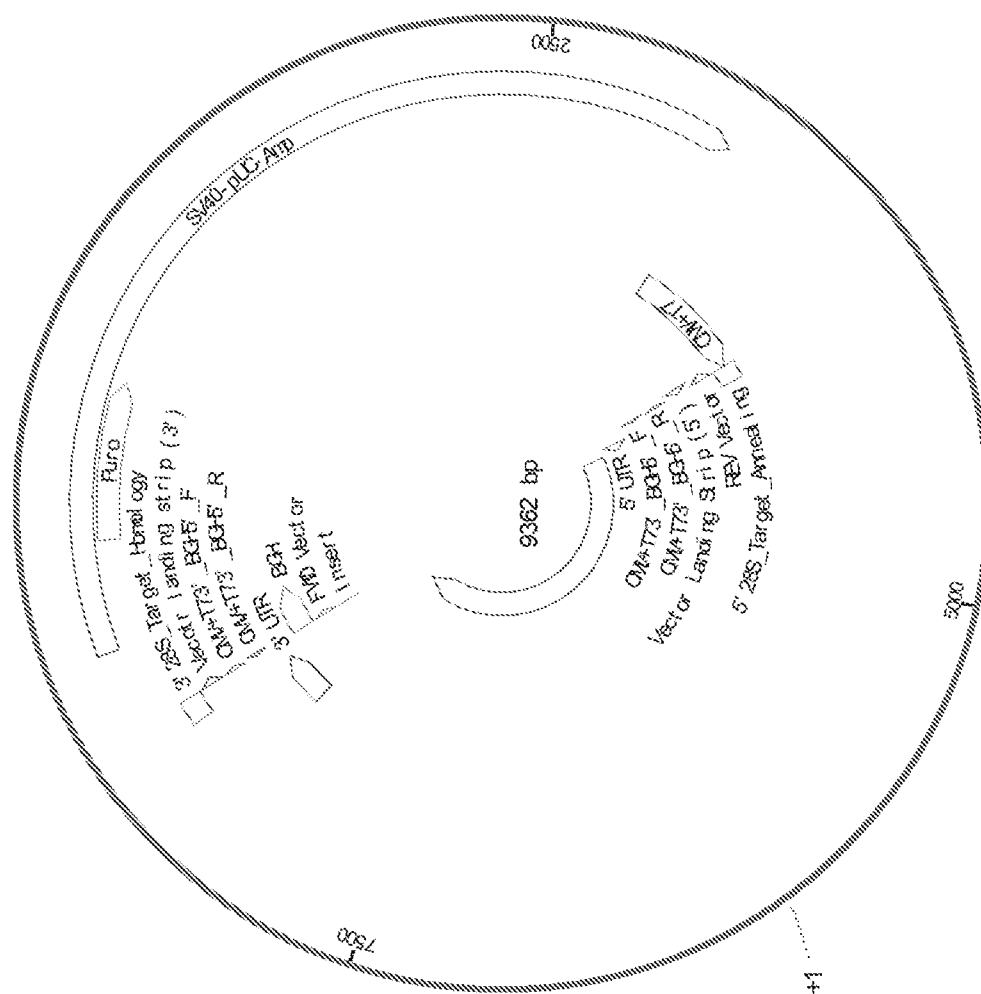


20/24

FIG. 21B

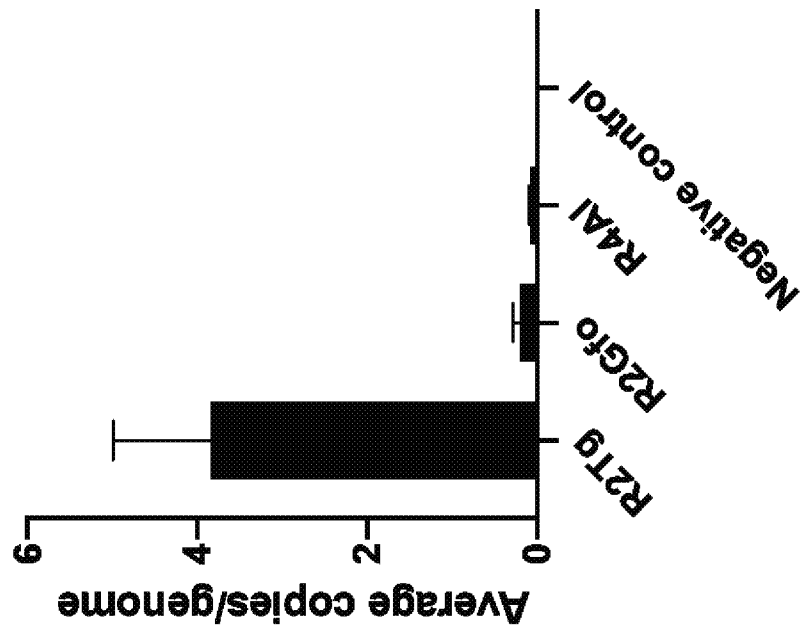


pR21Gfo_WT_28S_Annealing (9362 bp)



22/24

FIG. 23



23/24

FIG. 24

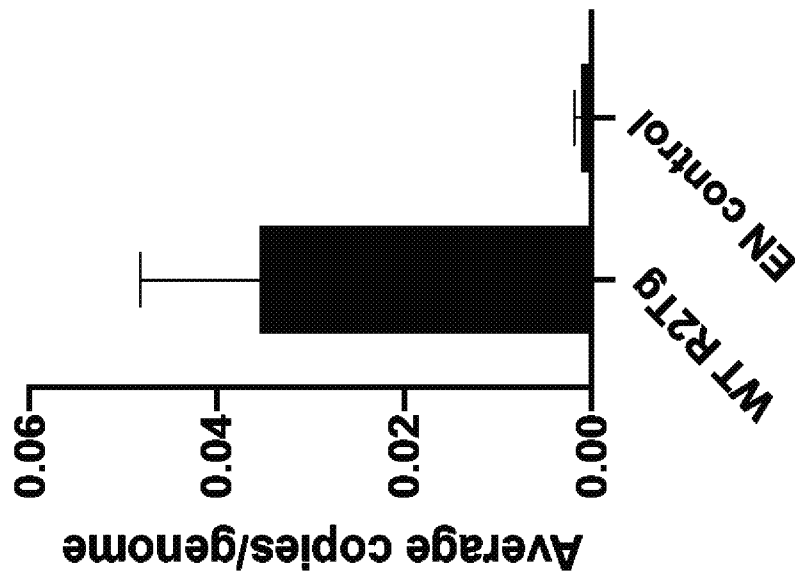
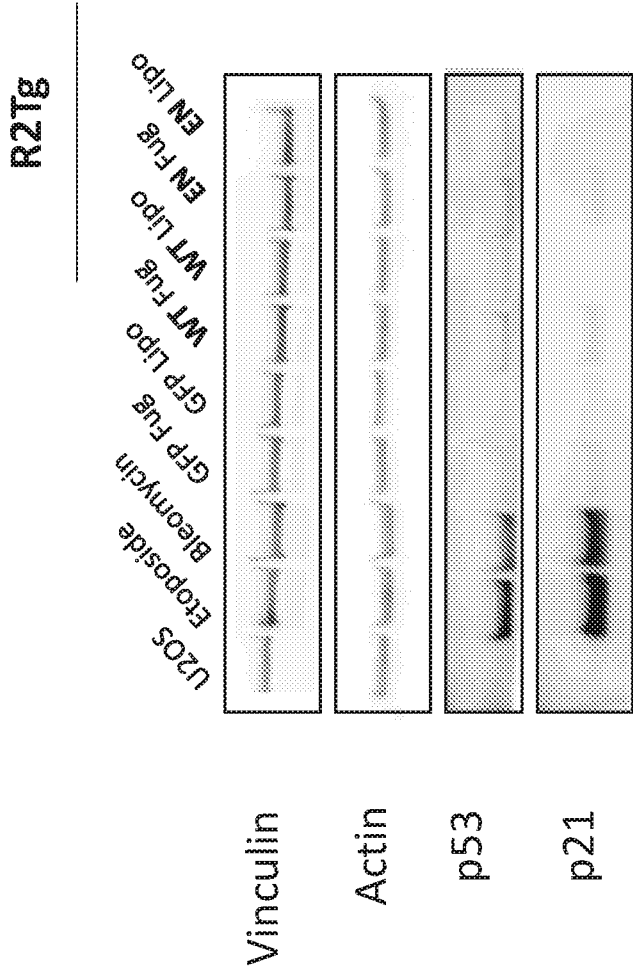


FIG. 25



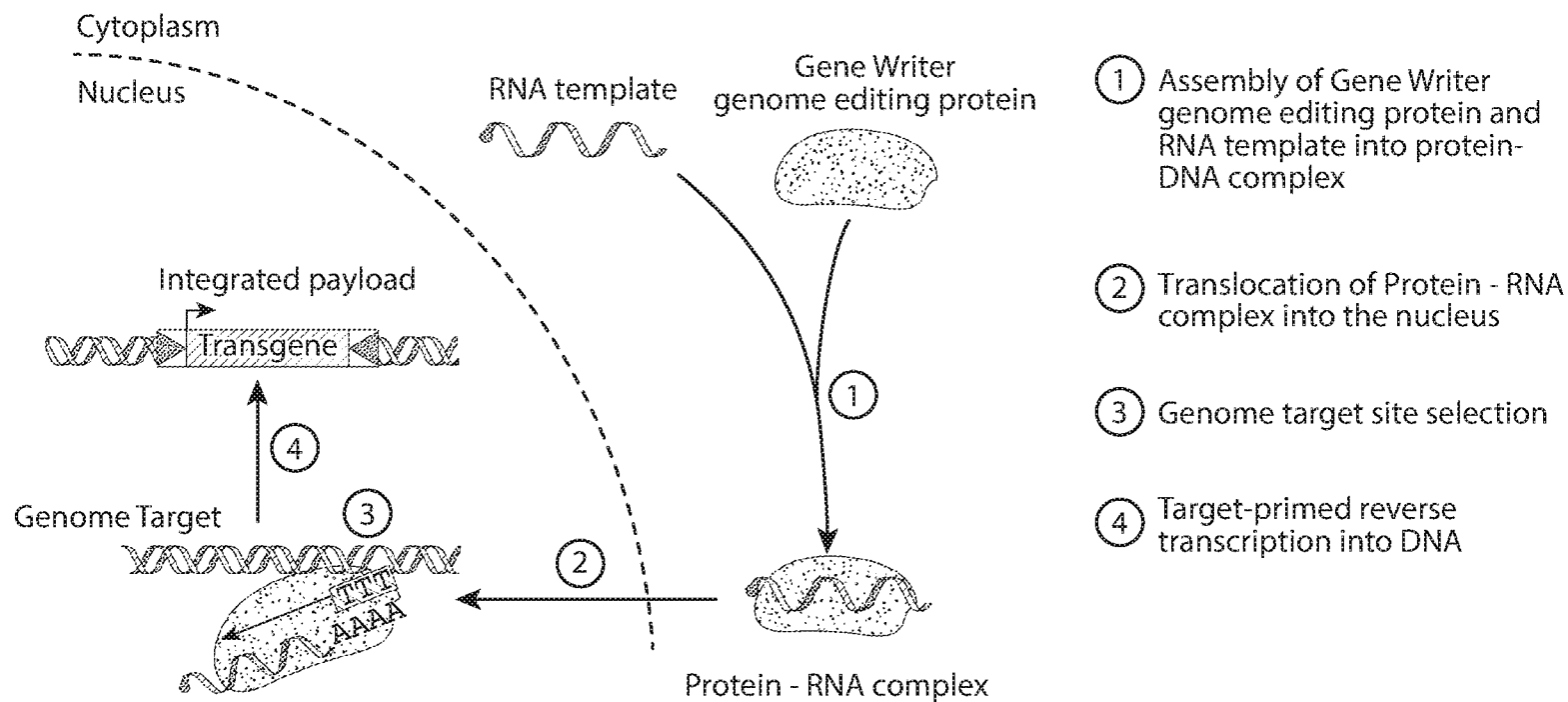


Figure 1