

DOCUMENT MADE AVAILABLE UNDER THE PATENT COOPERATION TREATY (PCT)

International application number:	PCT/US2022/073508
International filing date:	07 July 2022 (07.07.2022)
Document type:	Certified copy of priority document
Document details:	Country/Office: US
	Number: 63/219,291
	Filing date: 07 July 2021 (07.07.2021)
Date of receipt at the International Bureau:	29 July 2022 (29.07.2022)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a),(b) or (b-bis)

1218352

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

July 28, 2022

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 63/219,291

FILING DATE: *July 07, 2021*

RELATED PCT APPLICATION NUMBER: *PCT/US22/73508*

THE COUNTRY CODE AND NUMBER OF YOUR PRIORITY APPLICATION, TO BE USED FOR FILING ABROAD UNDER THE PARIS CONVENTION, IS *US63/219,291*



Certified by

Kathi

Performing the Functions and Duties of the
Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

ABSTRACT

Embodiments of the present disclosure are directed to a plurality of nucleotide sequences encoding a first intein fused to at least a portion of a first transcription activator-like effector (TALE) and encoding a second intein fused to at least a portion of a rare-cutting nuclease.

CLAIMS

1. A plurality of nucleotide sequences encoding a first intein fused to at least a portion of a first transcription activator-like effector (TALE) and encoding a second intein fused to at least a portion of a rare-cutting nuclease.
2. The plurality of nucleotide sequences of any of the claims provided herein, wherein the plurality of nucleotide sequences further encode the first intein fused to at least a portion of the second TALE.
3. The plurality of nucleotide sequences of any of the claims provided herein, wherein the plurality of nucleotide sequences include:
 - a first nucleotide sequence encoding the first intein fused to the first TALE;
 - a second nucleotide sequence encoding the first intein fused to at least a portion of the second TALE; and
 - a third nucleotide sequence encoding the second intein fused to the at least portion of the rare-cutting nuclease.
4. The plurality of nucleotide sequences, wherein the first nucleotide sequence encodes the first intein fused to a first portion of the rare-cutting nuclease and the first TALE, and the third nucleotide sequence encodes the second intein fused to a second portion of the rare-cutting nuclease.
5. The plurality of nucleotide sequence, wherein the third nucleotide sequence encodes the second intein fused to the rare-cutting nuclease.
6. The plurality of nucleotide sequences of any of the claims provided herein, wherein the plurality of nucleotide sequences each include a separate vector and/or are on a single expression construct.
7. The plurality of nucleotide sequences of any of the claims provided herein, wherein the first intein and the second intein are configured to self-splice when in contact

and to form a first half transcription activator-like effector nuclease (TALEN) including the first TALE bound to the rare-cutting nuclease.

8. The plurality of nucleotide sequences of any of the claims provided herein, wherein the first intein and the second intein are configured to self-splice when in contact and to form a second half TALEN including the second TALE bound to the rare-cutting nuclease.

9. The plurality of nucleotide sequences of any of the claims provided herein, wherein the first intein and the second intein are configured to self-splice when in contact and to form a fusion protein including the first intein bound to the second intein.

10. The plurality of nucleotide sequences of any of the claims provided herein, wherein each of the plurality of nucleotide sequences further encode a promoter and a terminator.

11. A method comprising:
contacting a cell with:
a first nucleotide sequence encoding a first intein fused to at least a portion of a first transcription activator-like effector (TALE);
a second nucleotide sequence encoding the first intein fused to at least a portion of a second TALE; and
a third nucleotide sequence encoding a second intein fused to at least a portion of a rare-cutting nuclease; and
in response to contacting the cell, splicing the first TALE, the second TALE, and the rare-cutting nuclease by the first intein and the second intein to form:
a first half transcription activator-like effector nuclease (TALEN) including the first TALE and the rare-cutting nuclease; and
a second half TALEN including the second TALE and the rare-cutting nuclease.

12. The method of any of the claims provided herein, further comprising translating the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence to form the first intein fused to the first TALE, the first intein fused to the second TALE, and the second intein fused to the rare-cutting nuclease.
13. The method of any of the claims provided herein, further comprising transforming the cell using the first half TALEN and the second half TALEN.
14. The method of any of the claims provided herein, wherein the first nucleotide sequence encodes an N-terminal intein fused to a C-terminal of the first TALE, the second nucleotide sequence encodes the N-terminal intein fused to a C-terminal of the second TALE, and the third nucleotide sequence encodes a C-terminal intein fused to an N-terminal of the rare-cutting nuclease or the C-terminal intein fused between portions of the rare-cutting nuclease.
15. The method of any of the claims provided herein, wherein the first nucleotide sequence encodes a C-terminal intein fused to an N-terminal of the first TALE, the second nucleotide sequence encodes the C-terminal intein fused to an N-terminal of the second TALE, and the third nucleotide sequence encodes an N-terminal intein fused to a C-terminal of the rare-cutting nuclease or the N-terminal intein fused between portions of the rare-cutting nuclease.
16. The method of any of the claims provided herein, wherein splicing comprises binding the first intein to the second intein to form:
- a first intermediate including the first intein fused to the second intein, wherein the first intein is further fused to the first TALE and the second intein is further fused to the rare-cutting nuclease; and
 - a second intermediate including the first intein fused to the second intein, wherein the first intein is further fused to the second TALE and the second intein is further fused to the rare-cutting nuclease.
17. The method of any of the claims provided herein, wherein splicing comprises:

binding the first intein to the second intein;
cutting splice sites associated with the first intein and the second intein; and
binding the first TALE to the rare-cutting nuclease and binding the second TALE
to the rare-cutting nuclease to form the first half TALEN and the second half TALEN.

18. The method of any of the claims provided herein, wherein the splice sites are
between:

the first intein and the first TALE;
the first intein and the second TALE; and
the second intein and the rare-cutting nuclease or portions thereof.

19. A system or expression construct comprising:
a first nucleotide sequence encoding a first intein fused to at least a portion of a
first transcription activator-like effector (TALE);
a second nucleotide sequence encoding the first intein fused to at least a portion of
a second TALE; and
a third nucleotide sequence encoding a second intein fused to at least a portion of
a rare-cutting nuclease.

20. The system or expression construct of any of the claims herein, wherein the first
nucleotide sequence encodes an N-terminal intein fused to a C-terminal of the first
TALE, the second nucleotide sequence encodes the N-terminal intein fused to a C-
terminal of the second TALE, and the third nucleotide sequence encodes a C-terminal
intein fused to an N-terminal of the rare-cutting nuclease or fused between portions of the
rare-cutting nuclease.

21. The system or expression construct of any of the claims herein, wherein the first
nucleotide sequence encodes a C-terminal intein fused to an N-terminal of the first
TALE, the second nucleotide sequence encodes the C-terminal intein fused to an N-
terminal of the second TALE, and the third nucleotide sequence encodes an N-terminal
intein fused to a C-terminal of the rare-cutting nuclease or fused between portions of the
rare-cutting nuclease.

22. The system or expression construct of any of the claims herein, wherein in response to translation of the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence by a cell, the first intein and second intein are configured to bind to one another and self-splice to form:

- a first half transcription activator-like effector nuclease (TALEN) including the first TALE bound to the rare-cutting nuclease;

- a second half TALEN including the second TALE bound to the rare-cutting nuclease; and

- a fusion protein including the first intein bound to the second intein.

23. The system or expression construct of any of the claims herein, wherein the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence include separate vectors or are on a single expression construct.

24. The system or expression construct of any of the claims herein, wherein:

- the first TALE comprises a first plurality of TALE repeat sequences that, in combination, bind to a first nucleotide sequence in a target DNA sequence; and

- the second TALE comprises a second plurality of TALE repeat sequences that, in combination, bind to a second nucleotide sequence in the target DNA sequence.

25. The system or expression construct of any of the claims herein, wherein each of the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence further encode a promoter and a terminator.

26. A host cell and/or organism transformed by the methods, vectors, expression constructs, nucleotide sequences, and/or systems described herein.

27. A method of forming any of the nucleotide sequences and/or systems claimed herein.

TRANSCRIPTION ACTIVATOR-LIKE EFFECTORS FUSED TO INTEINS

BACKGROUND

[0001] Genome editing technologies using rare-cutting nuclease, such as Transcription activator-like effector nucleases (TALEN), zinc finger nucleases (ZFNs), Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and related CRISPR associated protein 9 (Cas9) or Cpf1 systems, have accelerated basic biology research, biotechnology, breeding, and gene therapy. The rare-cutting nucleases can be used to generate deletions, insertions, and initiate homologous recombination. TALENs include a rare-cutting nuclease and a TALE that can target a specific sequence and generate a precise break in deoxyribonucleic acid (DNA). One TALEN design includes a left-half TALEN that includes a left TALE fused to the rare-cutting nuclease that recognizes a first binding site followed by a spacer region and a right-half TALEN that includes a right TALE fused to the rare-cutting nuclease that recognizes a second binding site.

SUMMARY

[0002] The present disclosure is directed to overcoming the above-mentioned challenges and needs related to TALEs.

[0003] Various embodiments are directed to a plurality of nucleotide sequences encoding a first intein fused to at least a portion of a first transcription activator-like effector (TALE) and encoding a second intein fused to at least a portion of a rare-cutting nuclease. In some embodiments, the plurality of nucleotide sequences further encode the first intein fused to at least a portion of a second TALE.

[0004] In various embodiments, the plurality of nucleotide sequences include: a first nucleotide sequence encoding the first intein fused to at least a portion of the first TALE, a second nucleotide sequence encoding the first intein fused to at least a portion of the second TALE, and a third nucleotide sequence encoding the second intein fused to the at least portion of the rare-cutting nuclease. In some embodiments, the first nucleotide sequence encodes the first intein fused to the first TALE (e.g., as a single sequence).

[0005] In some embodiments, the first nucleotide sequence encodes the first intein fused to a first portion of the rare-cutting nuclease and the first TALE, and the third nucleotide sequence encodes the second intein fused to a second portion of the rare-cutting nuclease. In some embodiments, the second nucleotide sequence encodes the first intein fused to the first portion of the rare-cutting nuclease and the second TALE.

[0006] In some embodiments, the third nucleotide sequence encodes the second intein fused to the rare-cutting nuclease.

[0007] The plurality of nucleotide sequences can each include a separate vector and/or are on a single expression construct.

[0008] In some embodiments, the first intein and the second intein are configured to self-splice when in contact and to form a first half TALEN including the first TALE bound to the rare-cutting nuclease.

[0009] In some embodiments, the first intein and the second intein are configured to self-splice when in contact and to form a second half TALEN including the second TALE bound to the rare-cutting nuclease.

[0010] In some embodiments, the first intein and the second intein are configured to self-splice when in contact and to form a fusion protein including the first intein bound to the second intein.

[0011] In various embodiments, each of the plurality of nucleotide sequences further encode a promoter and a terminator.

[0012] A number of embodiments are directed to a method comprising contacting a cell with: a first nucleotide sequence encoding a first intein fused to at least a portion of a TALE, a second nucleotide sequence encoding the first intein fused to at least a portion of a second TALE, and a third nucleotide sequence encoding a second intein fused to at least a portion of a rare-cutting nuclease. The method further comprises, in response to contacting the cell, splicing the first TALE, the second TALE, and the rare-cutting nuclease by the first intein and the second intein to form: a first half TALEN including the first TALE and the rare-cutting nuclease, and a second half TALEN including the second TALE and the rare-cutting nuclease.

[0013] In some embodiments, the method further comprises translating the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence

to form the first intein fused to the first TALE, the first intein fused to the second TALE, and the second intein fused to the rare-cutting nuclease.

[0014] In some embodiments, the method further comprises transforming the cell using the first half TALEN and the second half TALEN.

[0015] In various embodiments, the first nucleotide sequence encodes an N-terminal intein fused to a C-terminal of the first TALE. The second nucleotide sequence encodes the N-terminal intein fused to a C-terminal of the second TALE. And, the third nucleotide sequence encodes a C-terminal intein fused to an N-terminal of the rare-cutting nuclease or the C-terminal intein fused between portions of the rare-cutting nuclease.

[0016] In various embodiments, the first nucleotide sequence encodes a C-terminal intein fused to an N-terminal of the first TALE. The second nucleotide sequence encodes the C-terminal intein fused to an N-terminal of the second TALE. And, the third nucleotide sequence encodes an N-terminal intein fused to a C-terminal of the rare-cutting nuclease or the N-terminal intein fused between portions of the rare-cutting nuclease.

[0017] In some embodiments, splicing comprises binding the first intein to the second intein to form a first intermediate and a second intermediate. The first intermediate includes the first intein fused to the second intein, wherein the first intein is further fused to the first TALE and the second intein is further fused to the rare-cutting nuclease. The second intermediate includes the first intein fused to the second intein, wherein the first intein is further fused to the second TALE and the second intein is further fused to the rare-cutting nuclease.

[0018] In some embodiments, splicing comprises binding the first intein to the second intein, cutting splice sites associated with the first intein and the second intein, and binding the first TALE to the rare-cutting nuclease and binding the second TALE to the rare-cutting nuclease to form the first half TALEN and the second half TALEN.

[0019] The splice sites can be between: the first intein and the first TALE, the first intein and the second TALE, and the second intein and the rare-cutting nuclease or portions thereof.

[0020] Some embodiments are directed to a system or expression construct comprising a first nucleotide sequence, a second nucleotide sequence, and a third nucleotide sequence. The first nucleotide sequence encodes a first intein fused to at least

a portion of a first TALE. The second nucleotide sequence encodes the first intein fused to at least a portion of a second TALE. And, the third nucleotide sequence encodes a second intein fused to at least a portion of a rare-cutting nuclease.

[0021] In some embodiments, the first nucleotide sequence encodes an N-terminal intein fused to a C-terminal of the first TALE. The second nucleotide sequence encodes the N-terminal intein fused to a C-terminal of the second TALE. And, the third nucleotide sequence encodes a C-terminal intein fused to an N-terminal of the rare-cutting nuclease or fused between portions of the rare-cutting nuclease.

[0022] In some embodiments, the first nucleotide sequence encodes a C-terminal intein fused to an N-terminal of the first TALE. The second nucleotide sequence encodes the C-terminal intein fused to an N-terminal of the second TALE. And, the third nucleotide sequence encodes an N-terminal intein fused to a C-terminal of the rare-cutting nuclease or fused between portions of the rare-cutting nuclease.

[0023] In various embodiments, in response to translation of the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence by a cell, the first intein and second intein are configured to bind to one another and self-splice to form: a first half TALEN including the first TALE bound to the rare-cutting nuclease, a second half TALEN including the second TALE bound to the rare-cutting nuclease, and a fusion protein including the first intein bound to the second intein.

[0024] In some embodiments, the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence include separate vectors or are on a single expression construct.

[0025] In some embodiments, the first TALE comprises a first plurality of TALE repeat sequences that, in combination, bind to a first nucleotide sequence in a target DNA sequence. And, the second TALE comprises a second plurality of TALE repeat sequences that, in combination, bind to a second nucleotide sequence in the target DNA sequence.

[0026] In various embodiments, each of the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence further encode a promoter and a terminator.

[0027] Various embodiments are directed to host cell and/or organism transformed by the methods, vectors, expression constructs, nucleotide sequences, and/or systems described herein.

[0028] Various embodiments are directed to a method of forming any of the nucleotide sequences and/or systems claimed herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0029] Various example embodiments can be more completely understood in consideration of the following detailed description in connection with the accompanying drawings, in which:

[0030] FIGs. 1A-1D are diagrams illustrating example nucleotide sequences that encode a first intein with a TALE and encode a second intein with a rare-cutting nuclease, consistent with the present disclosure.

[0031] FIG. 2 is a flow diagram illustrating an example method for generating TALENs using example nucleotide sequences, consistent with the present disclosure.

[0032] FIGs. 3A-3B are diagrams illustrating example expression constructs, consistent with the present disclosure.

[0033] FIG. 4 is a flow diagram illustrating an example of splicing by inteins to form a TALEN, consistent with the present disclosure.

[0034] FIG. 5 are images showing resulting genotyping from editing a cell using nucleotide sequences, consistent with the present disclosure.

[0035] FIGs. 6A-6C illustrate full images of the genomic regions from an YFP negative control, a TALEN positive control, and inteins, consistent with the present disclosure.

DETAILED DESCRIPTION

[0036] Aspects of the present disclosure are directed to a variety of methods, nucleotide sequences, systems, expression constructs, and host cells and/or organisms transformed using the nucleotide sequences. While the present invention is not necessarily limited to such applications, various aspects of the invention may be appreciated through a discussion of various embodiments using this context.

[0037] Accordingly, in the following description various specific details are set forth to describe specific embodiments presented herein. It should be apparent to one skilled in the art, however, that one or more other examples and/or variations of these embodiments can be practiced without all the specific details given below. In other

instances, well known features have not been described in detail so as not to obscure the description of the embodiments herein. For ease of illustration, the same reference numerals can be used in different diagrams to refer to the same elements or additional instances of the same element.

[0038] Rare-cutting nucleases, such as rare-cutting endonucleases, can be used to target specific genes or target multiple genes. Non-limiting examples include TALENs, engineered homing endonucleases, ZFNs, meganucleases, and CRISPR. Rare-cutting nucleases can be natural or engineered proteins having endonuclease activity directed to a nucleic acid sequence with a recognition site, sometimes referred to as a target sequence, of 12-40 base pairs in length or longer. Typically, rare-cutting nucleases cause cleavage inside the recognition site. In some instances, the rare-cutting nuclease is a fusion protein that contains a binding domain fused to a nuclease domain with cleavage activity. The binding domain can be configured to bind to a target sequence. The rare-cutting endonuclease domain can be configured to induce a mutation at a target genomic locus associated with the target location. TALENs and ZFNs are examples of fusion proteins of the binding domains with the nuclease domain, such as an endonuclease of FokI. TALENs can be genetically engineered to have specificity to a target sequence via the binding domains fused to the nuclease domain, resulting in chimeric nucleases targeted to specific, selected DNA sequences, and leading to cutting of DNA at or near the target sequence. Such DNA cuts (double-stranded breaks) can induce mutations, such as knocking out or otherwise altering gene function with precision and efficiency.

[0039] Embodiments in accordance with the present disclosure are directed to use of inteins in nucleotide sequences containing TALEs and nucleases. In some embodiments, the half TALEs and nuclease are separate from one another, and the inteins can splice and ligate the respective half TALEs to the nuclease after translation. By using complementary inteins for the nucleotide sequences containing the half TALEs and the nucleases, the vectors delivered to and used to transform the cells of an organism can have a lower plasmid size, sometimes referred to as “cargo size”, as compared to nucleotide sequences containing the full half TALENs. Decreased plasmid size can increase the expression frequency of the nucleotide sequences. Additionally, use of inteins in nucleotide sequences containing TALEs and nucleases can add flexibility to delivery of gene editing material to the cell.

[0040] Turning now the figures, FIGs. 1A-1D are diagram illustrating example nucleotide sequences that encode a first intein with a TALE and encode a second intein with a rare-cutting nuclease, consistent with the present disclosure.

[0041] As shown by FIG. 1A, the plurality of nucleotide sequences 100 encode the first intein 104 fused to at least a portion of the first TALE 102 and encode the second intein 106 fused to at least a portion of the rare-cutting nuclease 108. As further illustrated by FIG. 1B, the first TALE 102 can include a TALE N-terminal 103, a first plurality of TALE repeat sequences 105 (illustrated as a binding domain), and a TALE C-terminal 107. The first plurality of TALE repeat sequences 105 can, in combination, bind to a first nucleotide sequence in a target DNA sequence, and can be referred to as a binding domain.

[0042] As shown by FIG. 1B, another plurality of nucleotide sequences 101 can encode the first intein 104 fused to at least a portion of the first TALE 102, the second intein 106 fused to at least a portion of the rare-cutting nuclease 108, and a first intein 112 fused to at least a portion of a second TALE 110. The first inteins 104, 112 can include sequences encoding the same intein (e.g., the intein sequence occurs twice). The second TALE 110 can include a TALE N-terminal 109, a second plurality of TALE repeat sequences 111 (illustrated as a binding domain), and a TALE C-terminal 113. The second plurality of TALE repeat sequences 111 can, in combination, bind to a second nucleotide sequence in a target DNA sequence, and can be referred to as a binding domain.

[0043] As shown by FIG. 1B, the plurality of nucleotide sequences 101 can include a first nucleotide sequence encoding the first intein 104 fused to the first TALE 102, a second nucleotide sequence encoding the first intein 112 fused to the second TALE 110, and a third nucleotide sequence encoding the second intein 106 fused to the rare-cutting nuclease 108. However, examples are not so limited and the first and/or second inteins can be fused to portions of the first TALE 102, second TALE 110, and/or rare-cutting nuclease 108, as further illustrated by FIG. 1C.

[0044] In some embodiments, the plurality of nucleotide sequences 100, 101 can each form a separate vector. In some embodiments, the plurality of nucleotide sequences 100, 101 can be formed on a single expression construct, such as illustrated by FIGs. 3A-3B.

[0045] The first intein(s) 104, 112 and the second intein 106 can be configured to self-splice when in contact. In response to the self-splicing, a first half TALEN can be

formed including the first TALE 102 bound to the rare-cutting nuclease 108. Further, a second half TALEN can be formed including the second TALE 110 bound to the rare-cutting nuclease 108. Additionally, a fusion protein including the first intein(s) 104, 112 bound to the second intein 106 can be formed.

[0046] As used herein, inteins are internal protein fragments or elements that self-excise from other protein(s) that the inteins are bound to and the inteins catalyze ligation of flanking components, sometimes referred to as exteins, with a peptide bond. Intein excision is a posttranslational process that may not require auxiliary enzymes or cofactors. This self-excision process is called “protein-splicing” by analogy to the splicing of RNA introns from pre-mRNA (Perler F et al., Nucl Acids Res. 22:1125-1127 (1994)). The first inteins 104, 112 and second intein 106 can respectively include an N-terminal intein and C-terminal intein having affinity for one another, and which bind together when in contact. The first inteins 104, 112 and second intein 106 can be referred to as trans-splicing inteins. With trans-splicing, one intein is an N-terminal intein (e.g., a fragment) which is bound to an N-extein and the other intein is a C-terminal intein which is bound to a C-extein. The N-terminal intein and C-terminal intein bind together, self-splice, and catalyze ligation of the N-extein and C-extein. In some examples, the intein sequences are derived from *Synechocystis* sp, *Saccharomyces cerevisiae*, *Pyrococcus horikoshii*, *Mycobacterium xenopi*, *Thermococcus kodakarensis*, *Methanocaldococcus jannaschii*, and *Nostoc punctiforme*, among others. Example inteins include Tfu pol-1 intein, DNA polymerase (DnaE) inteins (e.g., Ssp DnaE, Npu DnaE), and Mxc GyrA, Mru RecA, MTU RecA, Tli Pol-2, Sce VMA, and Ssp DNA helicase (Dna B), among others.

[0047] In some embodiments, the first and second inteins 104, 112, 106 can be orthogonal inteins. For example, an N-terminal intein from *Synechocystis* can bind to a C-terminal intein from *Nostoc*, and/or two different N-terminal inteins can bind to the same C-terminal intein.

[0048] As further illustrated herein, such as by FIG. 1D, the plurality of nucleotide sequences 100, 101 can each further encode additional elements, such as a promoter and/or a terminator.

[0049] FIG. 1C illustrates an example of a plurality of nucleotide sequences 115 that encode an intein fused between a portion of the rare-cutting nuclease and a TALE. For

example, the first nucleotide sequence encodes the first intein 104 fused to a first portion of the rare-cutting nuclease 108-1 fused to the first TALE 102. The second nucleotide sequence encodes the second intein 106 fused to a second portion of the rare-cutting nuclease 108-2. As shown, the first portion of the rare-cutting nuclease 108-1 is fused between the first TALE 102 and the first intein 102. Similar to the plurality of nucleotide sequences 100, 101, the first intein 104 and the second intein 106 can be configured to self-splice when in contact. In response to the splicing, the first and second portions of the rare-cutting nuclease 108-1, 108-2 can be bound together to form a first half TALEN. Although not illustrated by FIG. 1C, in some embodiments, a third nucleotide sequence can encode the first intein 104 fused to a first portion of the rare-cutting nuclease 108-1 fused to the second TALE, and, in response to the splicing, the first and second portions of the rare-cutting nuclease 108-1, 108-2 can be bound together to form a second half TALEN.

[0050] FIG. 1D illustrates an example of a plurality of nucleotide sequences 117 which each additionally include a promoter and a terminator. For example, the first nucleotide sequence encodes a first promoter 116, the first TALE 102, the first intein 104, a first terminator 118, and optionally, a portion of the rare-cutting nuclease 108-2. The second nucleotide sequence encodes a second promoter 124, the second TALE 110, the first intein 112, a second terminator 126, and optionally, a portion of the rare-cutting nuclease 108-3. The third nucleotide sequence encodes a third promoter 120, a second intein 106, at least a portion of the rare-cutting nuclease 108-2, and a third terminator 122. In the particular embodiment, the second intein 106 is fused to a second portion of the rare-cutting nuclease 108-2 and the first inteins 104, 112 are respectively fused to a first portion of the rare-cutting nuclease 108-1, 108-3 and the first TALE 102 or the second TALE 110. However, embodiments are not so limited and in various embodiments the first intein 104, 112 can be fused to a first portion of the first TALE 102 and/or a first portion of the second TALE 110, and/or the second intein 106 is fused to the second portion of the first TALE 102 and/or the second portion of the second TALE 110 which is fused to the rare-cutting nuclease as a single sequence. In other embodiments, the second intein 106 is fused to the rare-cutting nuclease as a single sequence.

[0051] As non-limiting examples, the promoters can include a nopaline synthase promoter (NosPro) or a T7 promoter, among others. Other example promoters can

include Sp6 promoter, a T3 promoter, Ubi promoter, a cauliflower mosaic virus (CaMV)35S promoter, an ADHI promoter, and ADH1 promoter, a GDS promoter, a TEF1 promoter, a Gal1 promoter, a CaMKIIa promoter, a T7lac promoter, an araBAD promoter, a trp promoter, a lac promoter, a Ptac promoter, among others.

[0052] As non-limiting examples, the terminators can include Nos terminator (NosTerm), CaMV terminator, t7S, tE9, tmas, tocs, tTr9, tpinIII, tORF25, ttml, among others.

[0053] FIG. 2 is a flow diagram illustrating an example method for generating TALENs using example nucleotide sequences, consistent with the present disclosure. The nucleotide sequences used in the method 230 can include the sequences 100, 101, 115, 117 of any of FIGs. 1A-1D.

[0054] At 232, the method 230 includes contacting a cell with a first nucleotide sequence, a second nucleotide sequence, and a third nucleotide sequence. The first nucleotide sequence encodes a first intein fused to at least a portion of a first TALE. The second nucleotide sequence encodes the first intein fused to at least a portion of a second TALE. The third nucleotide sequence encodes a second intein fused to at least a portion of a rare-cutting nuclease.

[0055] At 234, in response to contacting the cell, the method 230 includes splicing the first TALE, the second TALE, and the rare-cutting nuclease by the first intein and the second intein to form a first half TALEN including the first TALE and the rare-cutting nuclease, and a second half TALEN including the second TALE and the rare-cutting nuclease. The first and second half TALENs can include left and right-half TALENs.

[0056] For example and in response to contacting the cell with the nucleotide sequences, the method 230 can include transcribing and/or translating the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence by the cell. In response to the transcription and/or translation, the first intein fused to the first TALE, the first intein fused to the second TALE, and the second intein fused to the rare-cutting nuclease can be formed or expressed. In various embodiments, a plurality of copies of each of the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence can be transcribed and/or translated, resulting in a plurality of copies of each of the first intein fused to the first TALE, the first intein fused to the second TALE, and the second intein fused to the rare-cutting nuclease.

[0057] The first intein and second intein can self-splice when in contact. The splicing process can result in intermediates being formed prior to forming the first and second half TALENs. For example, the first intein and the second intein can bind to form a first intermediate and a second intermediate, as further illustrated by FIG. 4. The first intermediate can include the first intein fused to the second intein, wherein the first intein is further fused to the first TALE and the second intein is further fused to the rare-cutting nuclease. The second intermediate can include the first intein fused to the second intein, wherein the first intein is further fused to the second TALE and the second intein is further fused to the rare-cutting nuclease.

[0058] As described above, splicing includes binding between the inteins, cutting at splice sites, and binding between components. For example, splicing comprises binding the first intein to the second intein, cutting at splice sites associated with the first intein and the second intein, and binding the first TALE to the rare-cutting nuclease and the second TALE to the rare-cutting nuclease to form the first half TALEN and the second half TALEN. In some examples, splicing can include binding first and second portions of the rare-cutting nuclease to one another to form the full rare-cutting nuclease.

[0059] The splice sites can be between components of the plurality of nucleotide sequences. For example, splice sites can be between the first intein and the first TALE, between the first intein and the second TALE, and between the second intein and the rare-cutting nuclease or portions thereof. In some embodiments, splice sites can be between the second intein and each of the first and second portions of the rare-cutting nuclease.

[0060] In some embodiments, the method 230 can further include transforming the cell using the first half TALEN and the second half TALEN. For example, the first half TALEN and the second half TALEN can bind to a target DNA sequence via the binding domains and, in response, the endonucleases of the first half TALEN and the second half TALEN can cause a double stranded break in or near the target sequence.

[0061] Various embodiments are directed to systems that include the plurality of nucleotide sequences. In some embodiments, a single expression construct can include each of the plurality of nucleotide sequences. In other embodiments and/or in addition, each nucleotide sequence can form an individual vector.

[0062] FIGs. 3A-3B are diagrams illustrating example expression constructs, such as expression constructs including the nucleotide sequences illustrated by FIGs. 1A-1D,

consistent with the present disclosure. Although FIGs. 3A-3B illustrate a single expression construct 340, 360, embodiments are not so limited and the respective sequences can be on separate vectors forming system

[0063] The system or expression construct 340, 360 comprise the above described first nucleotide sequence 341-A, 341-B, second nucleotide sequence 343-A, 343-B, and third nucleotide sequence 345-A, 345-B. The first nucleotide sequence 341-A, 341-B encodes a first intein 304, 362 fused to at least a portion of a first TALE 303, 305, 307. The first TALE 303, 305, 307 can include a TALE N-terminal 303, a first plurality of TALE repeat sequences, e.g., the first binding domain (BD1) 305, and a TALE C-terminal 307. The first plurality of TALE repeat sequences bind to a first nucleotide sequence in a target DNA sequence. The first nucleotide sequence 341-A, 341-B further includes a first promoter 342 and a first terminator 344. The first promoter 342 can be upstream of the first TALE 303, 305, 307 and the first intein 304, 362 and the first terminator 344 can be downstream of the first TALE 303, 305, 307 and the first intein 304, 362.

[0064] The second nucleotide sequence 343-A, 343-B encodes the first intein 312, 364 fused to at least a portion of a second TALE 309, 311, 313. The second TALE 309, 311, 313 can include a TALE N-terminal 309, a second plurality of TALE repeat sequences, e.g., the second binding domain (BD2) 311, and a TALE C-terminal 313. The second plurality of TALE repeat sequences bind to a second nucleotide sequence in the target DNA sequence. The second nucleotide sequence 343-A, 343-B further includes a second promoter 346 and a second terminator 348. The second promoter 346 can be upstream of the second TALE 309, 311, 313 and the first intein 312, 364 and the second terminator 348 can be downstream of the second TALE 309, 311, 313 and the first intein 312, 364.

[0065] The third nucleotide sequence 345-A, 345-B encodes a second intein 306, 366 fused to at least a portion of a rare-cutting nuclease 308. The third nucleotide sequence 345-A, 345-B further includes a third promoter 350 and a third terminator 352. The third promoter 350 can be upstream of the rare-cutting nuclease 308 and the second intein 306, 366 and the third terminator 352 can be downstream of the rare-cutting nuclease 308 and the second intein 306, 366.

[0066] As shown by the expression construct 340 of FIG. 3A, in some embodiments the first nucleotide sequence 341-A encodes an N-terminal intein 304 fused to a C-terminal 307 of the first TALE 303, 305, 307. The second nucleotide sequence 343-A encodes the N-terminal intein 312 fused to a C-terminal 313 of the second TALE 309, 311, 313. And, the third nucleotide sequence 345-A encodes a C-terminal intein 306 fused to an N-terminal of the rare-cutting nuclease 308 or fused between portions of the rare-cutting nuclease 308. In such embodiments, the first TALE 303, 305, 307 and second TALE 309, 311, 313 are respectively upstream from the first inteins 304, 312. The second intein 306 is upstream from at least a portion of the rare-cutting nuclease 308.

[0067] As shown by the expression construct 360 of FIG. 3B, in some embodiments the first nucleotide sequence 341-B encodes a C-terminal intein 362 fused to an N-terminal 303 of the first TALE 303, 305, 307. The second nucleotide sequence 343-B encodes the C-terminal intein 364 fused to an N-terminal 309 of the second TALE 309, 311, 313. And, the third nucleotide sequence 345-B encodes an N-terminal intein 366 fused to a C-terminal of the rare-cutting nuclease 308 or fused between portions of the rare-cutting nuclease 308. In such embodiments, the first TALE 303, 305, 307 and second TALE 309, 311, 313 are respectively downstream from the first inteins 362, 364. The second intein 366 is downstream from at least a portion of the rare-cutting nuclease 308.

[0068] As used herein, an expression construct includes and/or refers a nucleic acid sequence (e.g., DNA sequence) including one or more vectors or binary vectors carrying genome editing reagents. The genome editing reagents can include or encode a nuclease and/or a TALE. In some embodiments, the expression construct can include a variety of nucleotide sequences, selected and arranged to facilitate transport of genome editing reagents in the cells. For example, the expression construct can include the above-described first sequence, second sequence, and third sequence. The rare-cutting nuclease can include a FokI protein, among other nucleases. In some embodiments, the expression construct and/or vectors can include other components, such as a detectable label, a promoter, and a terminator. The detectable label can include a fluorescent protein, a fluorophore, or nucleotide bound to a fluorophore, among other types of labels.

[0069] A vector or binary vector includes or refers to a nucleic acid sequence that includes one or more transgenes, sometimes referred to as “inserts”, and a backbone. The

binary vector can include an expression cassette that includes the transgene and a regulatory sequence to be expressed by a transformed cell.

[0070] As used herein, a domain includes and/or refers to a conserved part of a protein sequence and tertiary structure of the protein that can form a three-dimensional structure. The domains can be encoded by the expression constructs.

[0071] As further illustrated herein, in response to transcription and/or translation of the expression constructs 340, 360 respectively comprising the first nucleotide sequence 341-A, 341-B, the second nucleotide sequence 343-A, 343-B, and the third nucleotide sequence 345-A, 345-B, a plurality of copies of each of the first intein fused to the first TALE, the first intein fused to the second TALE, and the second intein fused to the rare-cutting nuclease can be formed. Further, respective inteins 304, 312, 306, 362, 364, 366 can bind to one another and self-splice to form first half TALENs and second half TALENS.

[0072] In some embodiments, the first promoter 342, the second promoter 346, and the third promoter 350 can include the same promoter. In other embodiments, the first promoter 342, the second promoter 346, and the third promoter 350 can each include different promoters. In further embodiments, the first promoter 342 and the second promoter 346 can be the same promoter, and the third promoter 350 can be a different promoter from the first promoter 342 and the second promoter 346. For example, the third promoter 350 can be a stronger promoter than the first promoter 342 and the second promoter 346, such that additional copies of the second intein fused to the rare-cutting nuclease are formed as compared to the number of copies of the first intein fused to the first TALE and the first intein fused to the second TALE after transcription and/or translation.

[0073] FIG. 4 is a flow diagram illustrating an example process of splicing by inteins to form TALENs, consistent with the present disclosure.

[0074] After contacting a population of cells with the plurality of nucleotide sequences, the plurality of nucleotide sequences are transcribed and/or translated by the cells to form the components 471, 473, and 475 including the first intein 478-A fused to the first TALE 474, the first intein 478-B fused to the second TALE 482, and the second intein 481 fused to the rare-cutting nuclease 483. Although a single copy of the

components 471, 473, and 475 are illustrated by FIG. 4, a plurality of copies of each of the components 471, 473, and 475 can be formed.

[0075] The component 471, 473, and 475 can come in contact with one another, and in response, the first inteins 478-A, 478-B respectively bind to copies of the second intein 481. In response to the binding, a first intermediate 477 and a second intermediate 479 are formed. The first intermediate 477 includes the first intein fused 478-A to the second intein 481, wherein the first intein 478-A is further fused to the first TALE 474 and the second intein 481 is further fused to the rare-cutting nuclease 483. The second intermediate 479 includes the first intein 478-B fused to the second intein 481, wherein the first intein 478-B is further fused to the second TALE 482 and the second intein 481 is further fused to the rare-cutting nuclease 483.

[0076] After the first inteins 478-A, 478-B respectively bind to copies of the second intein 481, the inteins 478-A, 478-B, 481 self-splice to form a first half TALEN 485, a second half TALEN 489, and a fusion protein 487. For example, the inteins 478-A, 478-B, 481 can cut at splice sites associated with the first inteins 478-A, 478-B and the second intein 481, bind the first TALE 474 to the rare-cutting nuclease 483, and bind the second TALE 482 to the rare-cutting nuclease 483. The first half TALEN 485 can include the first TALE 474 bound to the rare-cutting nuclease 483 proximal to the C-terminal of the first TALE 474. The second half TALEN 489 can include the second TALE 482 bound to the rare-cutting nuclease 483 proximal to the C-terminal of the second TALE 482. The fusion protein 487 can include the first intein 478 (c.g., 478-A or 478-B) bound to the second intein 481.

[0077] As previously described, the first half TALEN 485 and the second half TALEN 489 can transform cells. For example, various embodiments are directed to a host cell and/or organism transformed by the methods, vectors, expression constructs, nucleotide sequences, and/or systems described herein. In some embodiments, the cells can be transformed and/or organism transformed or regenerated as described by US Patent 8440431, issued on May 14, 2013, entitled "TAL effector-mediated DNA medication", which is incorporated herein in its entirety for its teaching.

[0078] As used herein, contacting the population of cells with the plurality of nucleic acid sequences can include delivering an expression construct into the population of cells. The expression construct can be delivered into the cells via different approaches

including, but not limited to, PEGmediated transformation, Agrobacterium infection, electroporation, particle bombardment, or microinjection mediated protoplast transformation, as well as combinations thereof.

[0079] In various embodiments, prior to contacting a population of cells with the plurality of nucleic acid sequences, such as an expression construct comprising the plurality of nucleic acid sequences. The plurality of nucleic acid sequences can be generated using standard molecular techniques.

[0080] In some examples, the population of cells can be screened to identify target cells that are genetically transformed by the plurality of nucleic acid sequences and/or expression construct. Target cells, as used herein, include and/or refer to cells that express the plurality of nucleic acid sequences and/or that otherwise exhibit or express the gene modification. The target cells can include the intended mutation at the target genomic locus. In some embodiments, the population of cells can be screened and target cells can be selected for expression of the expression construct via a detectable label. Screening the population of cells for the detectable label can include isolating target cells that have the detectable label from a remainder of the population of cells. Various embodiments include fluorescence activated cell sorting (FACS) based selection of transformed cells.

[0081] Accordingly, a number of embodiments are directed to the combination of DNA-mediated gene editing of cells, along with the selection of target cells receiving both half TALENs using FACS and fluorescent proteins or fluorophore labelling of the two TALENs. Organisms regenerated from FACS selected cells can be enriched for the intended gene edits, thus reducing the screening efforts typically required with transient gene expression.

[0082] Various embodiments of the present disclosure are directed to a non-naturally occurring host cell and/or organisms generated by the method 230 described by FIG. 2 and/or using the plurality of nucleic acid sequences or components illustrated by FIGs. 1A-1D, 3A-3B, and 4. For example, the method 230 can further include culturing the identified target cells that are transformed with the plurality of nucleic acid sequences, and regenerating an organism from the cultured target cells, where the regenerated organisms express the target modification. The plurality of nucleic acid sequences and/or resulting components (e.g., half TALENs and fusion proteins) can be removed (e.g., crossed away) from the regenerated organism.

[0083] In some embodiments and consistent with method 230, a non-naturally occurring organism can be generated by a genomic editing technique that includes using the plurality of nucleic acid sequences. The plurality of nucleic acid sequences can be separate vectors and/or formed on a single expression construct. The genomic editing technique can include contacting a population of cells with the plurality of nucleic acid sequences, screening the population of cells to identify target cells that are transformed with the plurality of nucleic acid sequences, and, optionally, regenerating a non-naturally occurring organism from the identified target cells.

[0084] The cell and/or cell population, as used herein, can be from a variety of different types of organisms. Examples cells can be from mammals, birds, reptiles, amphibians, fish, insects, crustaceans, arachnids, echinoderms, worms, mollusks, sponges, plants, fungi, algae, bacteria, among others.

[0085] As used herein, upstream can include a location proximal to and/or closer to the 5' end of the nucleic acid sequence as compared to the referenced sequence. Conversely, downstream can include a location proximal to and/or closer to the 3' end of the nucleic acid sequence as compared to the referenced sequence. As used herein, a sequence with adjectives listed in front, such as the rare-cutting nuclease sequence, intein sequence, or TALE sequence, includes or refers to a nucleotide sequence that encodes or is the adjectives (e.g., encodes or is the nuclease).

[0086] Different example approaches for enriching and/or screening the cells for the intended gene edit(s) are now described. Enriching and/or screening the cells can increase the representation of cells likely to contain the intended genomic edit.

[0087] The plurality of nucleotide sequence can be delivered into cells or other tissues using a variety of known methods such as PEG-mediated transformation, electroporation, bombardment, or microinjection mediated transformation. For larger tissues with cell walls such as embryos, bombardment (or biolistics) with gold particles coated with DNA can be used as delivery methods. Following delivery of the nucleotide sequences, FACS can be used to select fluorescent colored positive cells.

[0088] For particle bombardment transformation, the expression constructs can be coated onto particles, such as gold particles. To coat the nucleic acid on the gold particles, different volumes of nucleic acid solution are mixed with a fixed amount of gold suspension by pipetting.

[0089] Although embodiments are not so limited, and various particle bombardment transformation protocols can be used.

[0090] For convenience, certain terms employed in the specification, examples, and appended claims are provided here. The definitions are provided to aid in describing particular embodiments and are not intended to limit the claimed invention, as the scope of the invention is limited only by the claims.

[0091] The use of the term "or" in the claims and specification is used to mean "and/or" unless explicitly indicated to refer to alternatives only or the alternatives are mutually exclusive, although the disclosure supports a definition that refers to only alternatives and "and/or."

[0092] Unless the context clearly requires otherwise, throughout the description and the claims, the words "include", "including", "comprise," "comprising," and the like, are to be construed in an open and inclusive sense as opposed to a closed, exclusive or exhaustive sense. For example, the term "comprising" can be read to indicate "including, but not limited to." Words using the singular or plural number also include the plural and singular number, respectively. The words "a" and "an," when used in conjunction with the word "comprising" or "including" in the claims or specification, denotes one or more, unless specifically noted.

[0093] As used herein, the term "polypeptide" or "protein" refers to a polymer in which the monomers are amino acid residues that are joined together through amide bonds. When the amino acids are alpha-amino acids, either the L-optical isomer or the D-optical isomer can be used, the L-isomers being typical. The term polypeptide or protein as used herein encompasses any amino acid sequence and includes modified sequences such as glycoproteins. The term polypeptide, unless noted otherwise, is specifically intended to cover naturally occurring proteins, as well as those that are recombinantly or synthetically produced.

[0094] The term "nucleic acid" refers to a DNA or RNA nucleic acid and sequences of nucleic acids in either single or doublestranded form, and unless otherwise limited, encompasses known analogs of natural nucleotides that hybridize to nucleic acids in manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence includes the complementary sequence thereof.

[0095] Disclosed are materials, compositions, and components that can be used for, can be used in conjunction with, can be used in preparation for, or are products of the disclosed methods and compositions. It is understood that, when combinations, subsets, interactions, groups, etc., of these materials are disclosed, each of various individual and collective combinations is specifically contemplated, even though specific reference to each and every single combination and permutation of these compounds may not be explicitly disclosed. This concept applies to all aspects of this disclosure including, but not limited to, steps in the described methods. Specific elements of any foregoing embodiments can be combined or substituted for elements in other embodiments. For example, if there are a variety of additional steps that can be performed, each of these additional steps can be performed with any specific method step or combination of method steps of the disclosed methods, and each such combination or subset of combinations is specifically contemplated and disclosed. Additionally, it is understood that the embodiments described herein can be implemented using any suitable material such as those described elsewhere herein or as known in the art.

[0096] While illustrative embodiments have been illustrated and described, it will be appreciated that various changes can be made therein without departing from the scope of the invention.

EXPERIMENTAL EMBODIMENTS

[0097] Various experimental embodiments were directed to designing different nucleic acid vectors, sometimes herein referred to as vectors for ease of reference and which can include the previously described expression constructs or a portion thereof, such as a DNA or mRNA construct. The vectors encode a first intein fused to at least a portion of a first TALE, the first intein fused to at least a portion of a second TALE, and a second intein fused to at least a portion of a rare-cutting nuclease. Specific experiments were designed to show genetic editing by the above-described vectors. A number of experiments conducted are described herein.

[0098] In various experimental embodiments, several vectors were designed and constructed. The vectors included a TAL effector fused to an Ssp DnaE int-N peptide sequence, and an endonuclease fused to an Ssp DnaE int-C peptide sequence. Together, the Ssp DnaE int-N and int-C peptide sequences make up a trans intein, referred to here

as intein. Specific experiments were conducted to show genetic editing from joint activity of these intein vectors.

[0099] The different vectors are shown below in Table 1. The nucleic acid vectors in Table 1 include DNA constructs.

[0100] Table 1

Name	Composition	Description
pSW155	NosPro-TALE-BnFAD2(T03-L)-Ssp DnaE int-N-NosTerm	Left TAL effector with BnFAD2 DNA binding domain fused to an Ssp DnaE int-N sequence
pSW156	NosPro-TALE-BnFAD2(T03-R)-Ssp DnaE int-N-NosTerm	Right TAL effector with BnFAD2 DNA binding domain fused to an Ssp DnaE int-N sequence
pSW157	NosPro-TALE-Ssp DnaE int-N-NosTerm	Left TAL effector fused to an Ssp DnaE int-N sequence, includes BsaI sites for GG TALE cloning
pSW158	NosPro-TALE-Ssp DnaE int-N-NosTerm	Right TAL effector fused to an Ssp DnaE int-N sequence, includes BsaI sites for GG TALE cloning
pSW159	NosPro- Ssp DnaE int-C-FokI-NosTerm	FokI endonuclease fused to an Ssp DnaE int-C peptide sequence
pCLS29621	NosPro-TALE-BnFAD2(T03-L)-N-NosTerm	Control LHT with BnFAD2 DNA binding domain
pCLS29622	NosPro-TALE-BnFAD2(T03-R)-N-NosTerm	Control RHT with BnFAD2 DNA binding domain

[0101] The constructs in Table 1 were generated in the experimental embodiments are described in detail below. The vectors pSW155 and pSW156 include a TAL effector

that targets the gene BnFAD2 fused to an Ssp DnaE int-N peptide sequence. Vector pSW155 includes a promoter NosPro, a TAL effector targeting the gene BnFAD2, an Ssp DnaE int-N peptide sequence, a terminator NosTerm, and a left half TALEN (HT) backbone. The vector pSW156 includes a promoter NosPro, a TAL effector targeting the gene BnFAD2, an Ssp DnaE int-N peptide sequence, a terminator NosTerm, and a right HT backbone. Vector pSW159 includes an Ssp DnaE int-C peptide sequence fused to the FokI endonuclease. Vector pSW159 includes a promoter NosPro, an Ssp DnaE int-C peptide sequence, a FokI endonuclease, and a terminator NosTerm. In the experimental embodiments, vectors pSW155, pSW156, and pSW159 were jointly used to demonstrate TALEN gene editing activity.

[0102] A sequence of vector pSW155 is set forth in SEQ ID NO: 1, which includes a Nos promoter (SEQ ID NO: 2), left TALE N-terminal (SEQ ID NO: 3), BnFAD2(T03-L) binding domain (SEQ ID NO: 4), left TALE C-terminal (SEQ ID NO: 5), a linker (SEQ ID NO: 6), Ssp DnaE int-N (SEQ ID NO: 7), and a Nos terminator (SEQ ID NO: 8). A sequence of vector pSW156 is set forth in SEQ ID NO: 9, which includes a Nos promoter (SEQ ID NO: 2), right TALE N-terminal (SEQ ID NO: 10), BnFAD2(T03-R) binding domain (SEQ ID NO: 11), right TALE C-terminal (SEQ ID NO: 12), a linker (SEQ ID NO: 6), Ssp DnaE int-N (SEQ ID NO: 7), and a Nos terminator (SEQ ID NO: 8). A sequence of vector pSW159 is set forth in SEQ ID NO: 17, which includes a Nos promoter (SEQ ID NO: 2), Ssp DnaE int-C (SEQ ID NO: 18), a linker (SEQ ID NO: 19), FokI (SEQ ID NO: 20), and a Nos terminator (SEQ ID NO: 8).

[0103] Gene editing activity of the vectors pSW155, pSW156 and pSW159 were compared against control vectors pCLS29621 and pCLS29622. pCLS29621 consists of a promoter NosPro, a TALEN targeting BnFAD2(T03-L), and a terminator NosTerm in a left HT backbone. pCLS29622 consists of a promoter NosPro, a TALEN targeting BnFAD2(T03-R), and a terminator NosTerm in a right HT backbone.

[0104] The promoter NosPro and the terminator NosTerm are based on sequences from *Agrobacterium tumefaciens*, the TAL effector is based off a *Xanthomonas* sequence, the BnFAD2(T03) targets a sequence found in *Brassica napus*, the FokI endonuclease is based off a *Flavobacterium okeanokoites* sequence, and the Ssp DnaE int-C and int-N peptide sequences are from *Synechocystis* sp. PCC6803.

[0105] The remaining example constructs of Table 1 are described below. pSW157 and pSW158 are entry vectors that include a TAL effector fused to an Ssp DnaE int-N peptide sequence. Vector pSW157 includes a promoter NosPro, a TAL effector, a lacZ cassette flanked by BsaI sites for Golden Gate cloning, an Ssp DnaE int-N sequence, and terminator NosTerm in a left HT backbone. Vector pSW158 includes a promoter NosPro, a TAL effector, a lacZ cassette flanked by BsaI sites for Golden Gate cloning, an Ssp DnaE int-N sequence, and terminator NosTerm in a right HT backbone.

[0106] A sequence of vector pSW157 is set forth in SEQ ID NO: 13, which includes a Nos promoter (SEQ ID NO: 2), left TALE N-terminal (SEQ ID NO: 3), LacZ cassette (SEQ ID NO: 14), left TALE C-terminal (SEQ ID NO: 5), a linker (SEQ ID NO: 6), Ssp DnaE int-N (SEQ ID NO: 7), and a Nos terminator (SEQ ID NO: 8). A sequence of vector pSW158 is set forth in SEQ ID NO: 15, which includes a Nos promoter (SEQ ID NO: 2), right TALE N-terminal (SEQ ID NO: 10), LacZ cassette (SEQ ID NO: 16), right TALE c-terminus (SEQ ID NO: 12), a linker (SEQ ID NO: 6), Ssp DnaE int-N (SEQ ID NO: 7), and a Nos terminator (SEQ ID NO: 8).

[0107] Additional experiments were conducted to illustrate transformation of cells with the vectors of Table 1, as shown by Table 2. More specifically, canola protoplasts were transformed using the vectors illustrated in Table 1. These include the previously described vectors pSW155, pSW156, pSW159, pCLS29621, and pCLS29622. All samples were used to transform 200,000 protoplasts each, and all samples were tested as biological replicates. Samples A and D included 30 ug each of a process control plasmid, which functioned as a negative control during data analysis. Samples B and E included 30 ug each of control vectors for the genetic editing of target BnFAD2(T03), pCLS29621 and pCLS29622. Samples C and F included 30 ug of the three intein plasmids pSW155, pSW156, and pSW159. The intein samples C and F were compared to the positive control samples B and E to assess gene editing efficiency. All samples were prepared using the same Illumina sequence for analysis. The vectors were used to transform canola protoplasts to compare the gene editing efficiency of the intein TALEN vectors to the TALEN vectors without the intein peptide sequences. Table 2 illustrates the experiments conducted.

[0108] Table 2

Sample	A	B	C	D	E	F
Plasmid 1	pCLS31089	pCLS29621	pSW155	pCLS31089	pCLS29621	pSW155
Description	pVaUbi3_Y FP_NosTerm (negative control)	NosPro- TALE- BnFAD2(T 03-L)-N- NosTerm	NosPro- TALE- BnFAD2(T 03-L)- Ssp DnaE int-N- NosTerm	pVaUbi3_Y FP_NosTerm (negative control)	NosPro- TALE- BnFAD2(T 03-L)-N- NosTerm	NosPro- TALE- BnFAD2(T 03-L)- Ssp DnaE int-N- NosTerm
Type	DNA	DNA	DNA	DNA	DNA	DNA
Per 100k cells (ug)	15	15	15	15	15	15
Total Quantity (ug)	30	30	30	30	30	30
Conc. (ug/ul)	2.604	1.691	5.2	0.300	0.300	0.300
Vol. (ul)	11.52	17.74	5.77	100.00	100.00	100.00
Protoplast #	200K	200K	200K	200K	200K	200K
Plasmid 2		pCLS29622	pSW156		pCLS29622	pSW156
Description		NosPro- TALE- BnFAD2(T 03-R)-N- NosTerm	NosPro- TALE- BnFAD2(T 03-R)- Ssp DnaE int-N- NosTerm		NosPro- TALE- BnFAD2(T 03-R)-N- NosTerm	NosPro- TALE- BnFAD2(T 03-R)- Ssp DnaE int-N- NosTerm
Type		DNA	DNA		DNA	DNA
Per 100k cells (ug)		15	15		15	15
Total Quantity (ug)		30	30		30	30
Conc. (ug/ul)		1.838	6.3		0.300	0.300
Vol. (ul)		16.32	4.76		100.00	100.00
Plasmid 3			pSW159			pSW159
Description						
Type			DNA			DNA
Per 100k cells (ug)			15			15
Total Quantity (ug)			30			30
Conc. (ug/ul)			3.1			0.300

Vol. (ul)			9.68			100.00
Note	Negative control	TALEN positive control	Intein TALEN sample	Negative control	TALEN positive control	Intein TALEN sample

[0109] To perform the canola transformation, 30-60 days before the experiment, 10 canola seeds were washed with 1.5 mL 70% ethanol, and then 1.8 mL of sterile water. To sterilize the seeds, 1.5 mL of a 1% sodium hypochlorite solution was used to wash the seeds, and then the seeds were washed an additional five times with 1.8 mL sterile water. After imbibing, six of the newly sterilized seeds were planted on 8P-MS-G media in a PlantCon. The seeds were incubated at 25C in a 16/8hr light/dark ratio.

[0110] After 30-60 days of incubation, the germinated canola plantlets were digested. Sterile scissors were used to excise 4-6 young canola leaves and the leaves were placed into a petri plate containing 50-100 ul CPDS. The leaves were chopped into 0.5-1 mm pieces using a sterile, straight edge razor. Another 4 mL of CPDS was added to the plate, and the plate was placed inside a larger 100 mm petri plate to ensure sterility. The plates were moved to a vacuum chamber and vacuum at 30 inHg for 10 minutes. After 10 minutes, the plate was incubated at 25 rpm for 16 hours in the dark at 25C.

[0111] After 16 hours, the protoplast digestion was washed with 4 mL of W-5 plus Carb100 solution, and the protoplast solution was gently pipetted through a Falcon 40 um cell strainer into a 50 mL tube. This step was repeated once more. The tube was then centrifuged for 5 minutes at 100xg. After centrifugation, the supernatant was discarded and the pellet was resuspended with 4 mL of W-5 wash buffer, centrifuged one more time, and remaining supernatant was discarded. The washed pellet was then resuspended with 2 mL W-5 wash buffer, and 20 ul of the suspension was loaded onto a hemocytometer for cell counting. Cell counts among four grids of the hemocytometer were used to get an average number of protoplasts per grid. The total number of protoplasts in the sample was calculated as follows: $((\bar{x}\{a,b,c,d\})/0.2) \times 1000 \times 2(\text{mL}) = \text{total \# of protoplasts}$, where $\bar{x}$ is the average of the four grids {a,b,c,d} that were counted previously. To perform the transformation, the 50 mL tube containing the 2 mL of W-5 buffer and protoplast suspension was centrifuged for 5 minutes at 100xg. For samples A-C, the supernatant was removed and 1 mL of room temperature 1X MMG per 1×10^6 protoplasts was added, and a volume corresponding to 200,000 protoplasts was added to a

1.5 mL microcentrifuge tube containing the specified amount of vectors for each transformation. For samples D-F, 500 μ l of room temperature 2X MMG per 1×10^6 protoplasts was added along with the specified amount of DNA vectors to the 200,000 protoplasts in a 1.5 mL microcentrifuge tube. For all samples, the protoplast suspension was mixed with the vectors by slowly pipetting the liquid up and down. The protoplasts and plasmid vectors were incubated at room temperature for 5 minutes. After 5 minutes, a 1X volume of room temperature PEG was added to each microcentrifuge tube, and pipetted up and down until thoroughly mixed. The tubes were then incubated at room temperature for 20 minutes. After 20 minutes, 1.5 mL of W-5 wash buffer was added to resuspend the protoplasts. The tubes were centrifuged at 200xg for 5 minutes, the supernatant was removed, and an additional 800 μ l of W-5 wash buffer was added to resuspend the cells. 0.5-2 mL of the washed protoplasts were transferred to a 6-24 well plate and incubated for 24-48 hours.

[0112] To assess gene editing activity, the protoplasts were harvested by transferring the protoplasts to a 1.5 mL microcentrifuge tube. The suspension was centrifuged at 200xg for 5 minutes, the supernatant was discarded, and the tubes were placed in liquid nitrogen for two minutes. The tubes were then stored at -80C until the protoplast DNA was extracted and analyzed by Illumina.

[0113] Table 3 illustrates detected deletions from the protoplasts transformed with the vectors described in Table 1. The gene editing efficiencies, shown here as percent events, were compared across the samples A-F. Samples A and D included canola protoplasts transformed with vectors expressing YFP and served as a negative control, where no editing was expected. Samples B and E included canola protoplasts transformed with TALENs targeting the gene BnFAD2. Samples C and F included canola protoplasts transformed with TAL effectors and a FokI endonuclease, each fused to an Ssp DnaE int-N or int-C sequence, also targeting BnFAD2.

[0114] Table 3 shows the results of an NHEJ mutation assay that detects the number of deletions, or events, in the population of protoplast cells that were transformed with the above vectors. The assay amplifies three genomic regions containing the target BnFAD2, represented in Table 3 as Illumina 1, 2, and 3. As shown, the TALEN intein samples produced a significantly higher number of deletions than the YFP negative control,

although they did not produce deletions at the same frequency of the TALEN control vectors.

[0115] Table 3

Sample	Biological Replicate	Sample description	Percent events (avg. across Illumina copies 1-3)
A	1	pVaUbi3_YFP_NosTerm negative control	0.006074698
A	2	pVaUbi3_YFP_NosTerm negative control	0.007731365
B	1	TALEN positive control	2.460016807
B	2	TALEN positive control	2.114974022
C	1	Intein TALEN sample	0.208803999
C	2	Intein TALEN sample	0.42680571
D	1	pVaUbi3_YFP_NosTerm negative control	0.004661722
D	2	pVaUbi3_YFP_NosTerm negative control	0.011750975
E	1	TALEN positive control	4.474206527
E	2	TALEN positive control	5.84799401
F	1	Intein TALEN sample	0.130803747
F	2	Intein TALEN sample	0.223925849

[0116] The experiments tested the use of trans-splicing inteins as a method to reduce plasmid cargo size of the TALEN vectors, and also add flexibility when delivering gene editing materials to the cell. As previously described, the experiments were performed in Canola (Bn-Westar) protoplast. The target with TALEN T03.01 in the FAD2 gene, which has three known copies. The protoplasts were genotyped using amplicon sequencing to detect edits in the three FAD2 gene copies. Table 4 below provides a summary of the resulting sequence coverage.

[0117] For the experiments, there was good coverage across all three gene copies for each intein sample (e.g., average of greater than 28,000 reads, with a range of between

24,000 and 41,000 reads). Experiments 1 and 2 were performed to test different concentrations of transformation inputs. Specifically, experiment 2 (the “Mod” method) used a higher volume of DNA at a lower concentration.

[0118] Table 4 and Table 5 below provide the percent editing for experiments 1 and 2. In both experiments, the intein samples consistently showed some level of editing higher than the negative controls, but lower than the positive controls. Experiment 2 (Mod method, lower amount of DNA) have higher percent editing in the positive control, but lower in the intein samples. As may be appreciated, percent editing = (number of reads with edits / total number of reads analyzed) * 100.

[0119] Table 4. Experiment 1 Percent Editing

	Sample A: YFP (-) control	Sample A: YFP (-) control	Sample B: (+) BnFAD2 (T03)	Sample B: (+) BnFAD2 (T03)	Sample C: Inteins	Sample C: Inteins
BnaA.F AD2.a	0.007	0.003	2.568	2.110	0.265	0.504
BnaC.F AD2.a	0.009	0.006	2.687	2.228	0.237	0.548
BnaC.F AD2.b	0.003	0.014	2.125	2.007	0.125	0.228
Avg.	0.006	0.008	2.460	2.115	0.209	0.427

[0120] Table 5 Experiment 2 Percent Editing

	Sample D: YFP (-) control	Sample D: YFP (-) control	Sample E: (+) BnFAD2 (T03)	Sample E: (+) BnFAD2 (T03)	Sample F: Inteins	Sample F: Inteins
BnaA. FAD2. a	0.004	0.010	4.775	6.835	0.171	0.252

BnaC. FAD2. a	0.007	0.009	5.048	6.122	0.177	0.316
BnaC. FAD2. b	0.003	0.016	3.599	4.587	0.044	0.104
Avg.	0.005	0.012	4.474	5.848	0.131	0.224

[0121] FIG. 5 are images showing resulting genotyping from editing a cell using nucleotide sequences, consistent with the present disclosure. More particularly, FIG. 5 illustrates aligned images of genomic regions of the negative control 590, the positive control 591, and the intein 592.

[0122] FIGs. 6A-6C illustrate full images of the genomic regions from an YFP negative control, a TALEN positive control, and inteins, consistent with the present disclosure. For example, FIG. 6A illustrates an image of genomic regions of Sample A YFP negative control. FIG. 6B illustrates an image of genomic regions of Sample B (+) BnFAD2 (T03) TALEN positive control. FIG. 6C illustrates an image of genomic regions of Sample C (+) Inteins.

SEQUENCE LISTING

[0123] SEQ ID NO: 1 Vector pSW155

CGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGC
GCAACGCAATTAATACGCGTACCGCTAGCCAGGAAGAGTTTGTAGAAACGCAAAAA
GGCCATCCGTCAGGATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCG
TCCTGCCCCGCCACCCTCCGGGCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGA
TTTGTCTACTCAGGAGAGCGTTACCGACAAACAACAGATAAAACGAAAGGCCAG
TCTTCCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCAGTTCCCTACTCTCGCGTTCCG
AATACATCTAGATCCAAGTACATGGTACCATCGGAATCGAGATTGCCTCGGTACAAG
GATTCTAGGTACCTCGCGAATGCATCTAGATCCAATGATCATGAGCGGAGAATTAAG
GGAGTCACGTTATGACCCCCGCCGATGACGCGGGACAAGCCGTTTTACGTTTGGAAC
TGACAGAACCGCAACGTTGAAGGAGCCACTCAGCCGCGGGTTTCTGGAGTTTAATGA
GCTAAGCACATACGTCAGAAACCATTATTGCGCGTTCAAAAGTCGCCTAAGGTCCT

ATCAGCTAGCAAATATTTCTTGTCAAAAATGCTCCACTGACGTTCCATAAATTCCCT
CGGTATCCAATTAGAGTCTCATATTCACTCTCAATCCAAATAATCTGCACCGGATCTC
GCCCTTACCTGCTAGTCATGGGCGATCCTAAAAAGAAACGTAAGGTCATCGATTACC
CATACGATGTTCCAGATTACGCTATCGATATCGCCGATCTACGCACGCTCGGCTACAG
CCAGCAGCAACAGGAGAAGATCAAACCGAAGGTTTCGTTTCGACAGTGGCGCAGCACC
ACGAGGCACTGGTCGGCCACGGGTTTACACACGCGCACATCGTTGCGTTAAGCCAAC
ACCCGGCAGCGTTAGGGACCGTCGCTGTCAAGTATCAGGACATGATCGCAGCGTTGC
CAGAGGCGACACACGAAGCGATCGTTGGCGTCGGCAAACAGTGGTCCGGCGCACGC
GCTCTGGAGGCCTTGCTCACGGTGGCGGGAGAGTTGAGAGGTCCACCGTTACAGTTG
GACACAGGCCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGACCGCAGTGGAGGC
AGTGTCATGCATGGCGCAATGCACTGACGGGTGCCCGCTCAACTTGACCCCCCAGCA
GGTGGTGGCCATCGCCAGCAATAATGGTGGCAAGCAGGCGCTGGAGACGGTCCAGC
GGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTGACCCCGGAGCAGGTGGTGGCCA
TCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCG
GTGCTGTGCCAGGCCACGGCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAAT
AATGGTGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAG
GCCACGGCTTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAG
CAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTG
ACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATAATGGTGGCAAGCAGGCGCTGGA
GACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTGACCCCGGAGCA
GGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGG
CGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTGACCCCGGAGCAGGTGGTGGCCA
TCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCG
GTGCTGTGCCAGGCCACGGCTTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAAT
ATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCAG
GCCACGGCTTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAA
GCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTT
GACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGG
AGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTGACCCCCCAGC
AGGTGGTGGCCATCGCCAGCAATAATGGTGGCAAGCAGGCGCTGGAGACGGTCCAG
CGGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTGACCCCGGAGCAGGTGGTGGCC
ATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCC
GGTGTGTGCCAGGCCACGGCTTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCA
CGATGGCGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCA
GGGCCACGGCTTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCA

AGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCACGGCT
TGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTG
GAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTGACCCCTCAG
CAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAGGCCGGCGCTGGAGAGCATTGTT
GCCCAGTTATCTCGCCCTGATCCGGCGTTGGCCGCGTTGACCAACGACCACCTCGTCG
CCTTGGCCTGCCTCGGCGGGCGTCTGCGCTGGATGCAGTGAAAAAGGGATTGGGGG
ATCCTATCTGCCTTTTCATTTCGGAACCGAAATTTTGACCGTGGAGTATGGACCTCTTCC
AATTGGAAAGATTGTGTCAGAAGAGATTAAGTGTCTGTTTACTCAGTGGATCCAGA
AGGAAGAGTGTACACCCAAGCTATTGCACAGTGGCATGATAGAGGAGAGCAAGAAG
TGCTTGAGTACGAACTTGAGGATGGTTCTGTGATTAGGGCTACTTCTGATCACAGGTT
CTTGACCACTGATTACCAGTTGCTTGCAATTGAGGAAATTTTCGCTAGGCAATTGGAT
CTTTTGACTCTTGAGAACATTAAGCAAAGTCTTGAGGAAGCTCTTGATAACCATAGGCTTC
CATTTCCCTTTGCTTGATGCAGGAATAAAGTCTTAAGTGAATCCTGTTGCCGGTCTTGCGATG
TTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATG
ATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCA
TGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATA
CGCGATAGAAAACAAAATATAGCGCGCAAAGTCTAGGATAAATTATCGCGCGCGGTGT
CATCTATGTTACTAGATCGGGAATTCGTAATCATGGTCATAGCATTGGATCGGATCCC
GGGCCCCGTCGACTGCAGAGGCCTGCATGCAAGAGCGTCGGAATCGAGATTGCCTCGG
TACAAGGATATCGATTGTCTTCATCGGATCCCATCCCCTATAGTGAGTCGTATTACAT
GGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAATCTCTGATGTTACATT
GCACAAGATAAAAATATATCATCATGCCTCCTCTAGACCAGCCAGGACAGAAATGCC
TCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACACCGTGGAACGGATGAAG
GCACGAACCCAGTGGACATAAGCCTGTTTCGGTTCGTAAGCTGTAATGCAAGTAGCGT
ATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCA
GTGGCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTTGGGGTACAGTCTATGCCTCGGG
CATCCAAGCAGCAAGCGCGTTACGCCGTGGGTGATGTTTGATGTTATGGAGCAGCA
ACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAAACATCATGAGGGAA
GCGGTGATCGCCGAAGTATCGACTCAACTATCAGAGGTAGTTGGCGTCATCGAGCGC
CATCTCGAACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGGC
CTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTAAGGCTTGATGAA
ACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAAGTTCGGCTTCCCCTGGAGAG
AGCGAGATTCTCCGCGCTGTAGAAGTCACCATTGTTGTGCACGACGACATCATTCCGT
GGCGTTATCCAGCTAAGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTC
TTGCAGGTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTATCTTGCTGACAAA

AGCAAGAGAACATAGCGTTGCCTTGGTAGGTCCAGCGGCGGAGGAACTCTTTGATCC
GGTTCCTGAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAACGCTATGGAATC
GCCGCCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTACGTTGTCCCGCATTG
GTACAGCGCAGTAACCCGGCAAAATCGCGCCGAAGGATGTCGCTGCCGACTGGGCAA
TGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAGACAGGCTTATC
TTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCAGTTGGAAGAATTTGTCC
ACTACGTGAAAGGCGAGATCACCAAGGTAGTCGGCAAATAACCCTCGAGCCACCCAT
GACCAAAATCCCTTAACGTGAGTTACGCGTCGTTCCACTGAGCGTCAGACCCCGTAG
AAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCA
AACAAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCAAC
TCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTA
GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTC
GCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCG
GGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTTCGGGCTGAACGGGG
GGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTA
CAGCGTGAGCATTGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTA
TCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAA
ACGCCTGGTATCTTTATAGTCCTGTGGGTTTCCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTT
TTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCTGCGTTATCCC
CTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAG
CCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATAC
GCAAACCGCCTCTCCCCGCGCGTTGGC

[0124] SEQ ID NO: 2 Nos promoter

GATCATGAGCGGAGAAATTAAGGGAGTCACGTTATGACCCCCGCGATGACGCGGGA
CAAGCCGTTTTACGTTTGGAACTGACAGAACCGCAACGTTGAAGGAGCCACTCAGCC
GCGGGTTTCTGGAGTTTAATGAGCTAAGCACATACGTCAGAAACCATTATTGCGCGT
TCAAAAGTCGCCTAAGGTCCTATCAGCTAGCAAATATTTCTTGTCAAAAATGCTCCA
CTGACGTTCCATAAATTCCCCTCGGTATCCAATTAGAGTCTCATATTCACTCTCAATC
CAAATAATCTGCACCGGATCT

[0125] SEQ ID NO: 3 Left TALE N-terminal

ATGGGCGATCCTAAAAAGAAACGTAAGGTCATCGATTACCCATACGATGTTCCAGAT
TACGCTATCGATATCGCCGATCTACGCACGCTCGGCTACAGCCAGCAGCAACAGGAG
AAGATCAAACCGAAGGTTTCGTTTCGACAGTGGCGCAGCACACGAGGCACTGGTCGG
CCACGGGTTTACACACGCGCACATCGTTGCGTTAAGCCAACACCCGGCAGCGTTAGG

GACCGTCGCTGTCAAGTATCAGGACATGATCGCAGCGTTGCCAGAGGGCGACACACGA
AGCGATCGTTGGCGTCGGCAAACAGTGGTCCGGCGCACGCGCTCTGGAGGGCCTTGCT
CACGGTGGCGGGAGAGTTGAGAGGTCCACCGTTACAGTTGGACACAGGCCAACTTCT
CAAGATTGCAAAACGTGGCGGGCGTGACCGCAGTGGAGGCAGTGCATGCATGGCGCA
ATGCACTGACGGGTGCCCCGCTCAAC

[0126] SEQ ID NO: 4 BnFAD2 (T03-L) Binding Domain

TTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATAATGGTGGCAAGCAGGCGCTG
GAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTTGACCCCCGAG
CAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTCCA
GCGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTTGACCCCCCAGCAGGTGGTGGC
CATCGCCAGCAATAATGGTGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCC
GGTGCTGTGCCAGGCCCACGGCTTGACCCCCGAGCAGGTGGTGGCCATCGCCAGCAA
TATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCA
GGCCCACGGCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATAATGGTGGCAA
GCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTT
GACCCCCGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGG
AGACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTTGACCCCCGAGC
AGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTCCAG
CGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTTGACCCCCGAGCAGGTGGTGGCC
ATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCG
GTGCTGTGCCAGGCCCACGGCTTGACCCCCGAGCAGGTGGTGGCCATCGCCAGCCAC
GATGGCGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAG
GCCACGGCTTGACCCCCGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAA
GCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTT
GACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATAATGGTGGCAAGCAGGCGCTGG
AGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTTGACCCCCGAGC
AGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTCCAG
CGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTTGACCCCCGAGCAGGTGGTGGCC
ATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCC
GGTGCTGTGCCAGGCCCACGGCTTGACCCCCGAGCAGGTGGTGGCCATCGCCAGCCA
CGATGGCGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCA
GGCCCACGGCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAA
GCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGC

[0127] SEQ ID NO: 5 Left TALE C-terminal

TTGACCCCTCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAGGCCGGCGCTG
GAGAGCATTGTTGCCAGTTATCTCGCCCTGATCCGGCGTTGGCCGCGTTGACCAACG
ACCACCTCGTCGCCTTGGCCTGCCTCGGCGGGCGTCCTGCGCTGGATGCAGTGAAAA
AGGGA

[0128] SEQ ID NO: 6 Linker

TTGGGGGATCCTATC

[0129] SEQ ID NO: 7 Ssp DnaE int-N

TGCCTTTCATTCGGAACCGAAATTTTGACCGTGGAGTATGGACCTCTTCCAATTGGAA
AGATTGTGTCAGAAGAGATTAACTGTTCTGTTTACTCAGTGGATCCAGAAGGAAGAG
TGTACACCCAAGCTATTGCACAGTGGCATGATAGAGGAGAGCAAGAAGTGCTTGAGT
ACGAACTTGAGGATGGTTCTGTGATTAGGGCTACTTCTGATCACAGGTTCTTGACCAC
TGATTACCAGTTGCTTGCAATTGAGGAAATTTTCGCTAGGCAATTGGATCTTTTGACT
CTTGAGAACATTAAGCAAACCTGAGGAAGCTCTTGATAACCATAGGCTTCCATTTCCCTT
TGCTTGATGCAGGAACTATTAAG

[0130] SEQ ID NO: 8 Nos terminator

CGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTT
GCGATGATTATCATATAATTTCTGTTGAATTACGTAAAGCATGTAATAATTAACATGT
AATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACAT
TTAATACGCGATAGAAAACAAAATATAGCGCGCAAACCTAGGATAAATTATCGCGCGC
GGTGTCATCTATGTTACTAGATCGGGAATTTCGTAATCATGGTCATAGC

[0131] SEQ ID NO: 9 Vector pSW156

CGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGC
GCAACGCAATTAATACGCGTACCGCTAGCCAGGAAGAGTTTGTAGAAACGCAAAAA
GGCCATCCGTCAGGATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCG
TCCTGCCCCGCCACCCTCCGGGCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGA
TTTGTCTACTCAGGAGAGCGTTACCGACAAACAACAGATAAAACGAAAGGCCAG
TCTTCCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCAGTTCCCTACTCTCGCGTTCCG
AATACATCTAGATCCAAGTACATGGTGAATCGGAATCGAGATTGCCTCGGTACAAG
GAGCGTAGGTACCTCGCGAATGCATCTAGATCCAATGATCATGAGCGGAGAATTAAG
GGAGTCACGTTATGACCCCCGCCGATGACGCGGGACAAGCCGTTTTACGTTTGGAAC
TGACAGAACCGCAACGTTGAAGGAGCCACTCAGCCGCGGGTTTCTGGAGTTTAATGA
GCTAAGCACATACGTCAGAAACCATTATTGCGCGTTCAAAAGTCGCCTAAGGTCACT
ATCAGCTAGCAAATATTTCTTGTCAAAAATGCTCCACTGACGTTCCATAAATCCCCCT

CGGTATCCAATTAGAGTCTCATATTTCACTCTCAATCCAAATAATCTGCACCGGATCTC
GCCCTTACCTGCTAGTCATGGGCGATCCTAAAAAGAAACGTAAGGTCATCGATAAGG
AGACTGCCGCTGCCAAGTTCGAGAGACAGCACATGGACAGCATCGATATCGCCGATC
TACGCACGCTCGGCTACAGCCAGCAGCAACAGGAGAAGATCAAACCGAAGGTTTCGT
TCGACAGTGGCGCAGCACACGAGGCACTGGTCGGCCACGGGTTTACACACGCGCAC
ATCGTTGCGTTAAGCCAACACCCGGCAGCGTTAGGGACCGTCGCTGTCAAGTATCAG
GACATGATCGCAGCGTTGCCAGAGGCGACACAGCAAGCGATCGTTGGCGTCGGCAA
ACAGTGGTCCGGCGCACGCGCTCTGGAGGCCTTGCTCACGGTGGCGGGAGAGTTGAG
AGGTCCACCGTTACAGTTGGACACAGGCCAACTTCTCAAGATTGCAAAACGTGGCGG
CGTGACCGCAGTGGAGGCAGTGCATGCATGGCGCAATGCACTGACGGGTGCCCCGCT
CAACTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATAATGGTGGCAAGCAGGC
GCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTGACCCC
CCAGCAGGTGGTGGCCATCGCCAGCAATAATGGTGGCAAGCAGGCGCTGGAGACGG
TCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTGACCCCCCAGCAGGTGG
TGGCCATCGCCAGCAATAATGGTGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGT
TGCCGGTGCTGTGCCAGGCCACGGCTTGACCCCCGGAGCAGGTGGTGGCCATCGCCA
GCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGCTGT
GCCAGGCCACGGCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTG
GCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCACG
GCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGC
TGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTGACCCCC
AGCAGGTGGTGGCCATCGCCAGCAATAATGGTGGCAAGCAGGCGCTGGAGACGGTC
CAGCGGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTGACCCCCGGAGCAGGTGGTG
GCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTT
GCCGGTGCTGTGCCAGGCCACGGCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAG
CAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTG
CCAGGCCACGGCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGG
CAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCACGG
CTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCT
GGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTGACCCCCGA
GCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTCC
AGCGGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTGACCCCCCAGCAGGTGGTGG
CCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGC
CGGTGCTGTGCCAGGCCACGGCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCA
ATGGCGGTGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCC

AGGCCCACGGCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATAATGGTGGCA
AGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGTCTGTGCCAGGCCCACGGCT
TGACCCCTCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGGCGGCAGGCCGGCGCTGG
AGAGCATTGTTGCCAGTTATCTCGCCCTGATCCGGCGTTGGCCGCGTTGACCAACGA
CCACCTCGTCGCCTTGGCCTGCCTCGGCGGGCGTCTGCGCTGGATGCAGTGAAAAA
GGGATTGGGGGATCCTATCTGCCTTTCATTTCGGAACCGAAATTTTGACCGTGGAGTAT
GGACCTCTTCCAATTGGAAAGATTGTGTGTCAGAAGAGATTAACTGTTCTGTTTACTCAG
TGGATCCAGAAGGAAGAGTGTACACCCAAGCTATTGCACAGTGGCATGATAGAGGA
GAGCAAGAAGTGCTTGAGTACGAACTTGAGGATGGTTCTGTGATTAGGGCTACTTCT
GATCACAGGTTCTTGACCACTGATTACCAGTTGCTTGCAATTGAGGAAATTTTCGCTA
GGCAATTGGATCTTTTGACTCTTGAGAACATTAAGCAAACCTGAGGAAGCTCTTGATA
ACCATAGGCTTCCATTTCCCTTTGCTTGATGCAGGAACTATTAAGTGATAACTCGAGAA
GGGCGCGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCG
GTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAA
CATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTA
TACATTTAATACGCGATAGAAAAACAAAATATAGCGCGCAAACCTAGGATAAATTATCG
CGCGCGGTGTCATCTATGTTACTAGATCGGGAATTCGTAATCATGGTCATAGCATTGG
ATCGGATCCCGGGCCCGTTCGACTGCAGAGGCCTGCATGCAAACCAAGCTCGGAATCG
AGATTGCCTCGGTACAAGTCAGTTCGATTGTCTTCATCGGATCCCATCCCCTATAGTG
AGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAATCT
CTGATGTTACATTGCACAAGATAAAAAATATATCATCATGCCTCCTCTAGACCAGCCA
GGACAGAAATGCCTCGACTTCGCTGCTGCCCCAAGGTTGCCGGGTGACGCACACCGTG
GAAACGGATGAAGGCACGAACCCAGTGGACATAAGCCTGTTTCGGTTCGTAAGCTGTA
ATGCAAGTAGCGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGG
TGGTAAACGGCGCAGTGGCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTGGGGTACAG
TCTATGCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTGATGTTTGATG
TTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAAAC
ATCATGAGGGAAGCGGTGATCGCCGAAGTATCGACTCAACTATCAGAGGTAGTTGGC
GTCATCGAGCGCCATCTCGAACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCA
GTGGATGGCGGCCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTA
AGGCTTGATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAACCTTCGGCT
TCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTCACCATTGTTGTGCACGAC
GACATCATTCGGTGGCGTTATCCAGCTAAGCGCGAACTGCAATTTGGAGAATGGCAG
CGCAATGACATTCTTGCAAGTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTA
TCTTGCTGACAAAAGCAAGAGAACATAGCGTTGCCTTGGTAGGTCCAGCGGCGGAGG

AACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAAC
GCTATGGAACTCGCCGCCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTACGTT
GTCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGC
CGACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAG
ACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCAGTTGGA
AGAATTTGTCCACTACGTGAAAGGCGAGATCACCAAGGTAGTCGGCAAATAACCCCTC
GAGCCACCCATGACCAAATCCCTTAACGTGAGTTACGCGTCGTTCCACTGAGCGTC
AGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC
TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAA
GAGCTACCAACTCTTTTTCCGAAGGTAACGTGGCTTCAGCAGAGCGCAGATACCAAAT
ACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGC
CTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTC
GTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCCGG
CTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACT
GAGATACCTACAGCGTGAGCATTGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGG
CGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTT
CCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTG
AGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCA
ACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT
GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCG
CTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAG
CGCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGC

[0132] SEQ ID NO: 10 Right TALE N-terminal

ATGGGCGATCCTAAAAAGAAACGTAAGGTCATCGATAAGGAGACTGCCGCTGCCAA
GTTTCGAGAGACAGCACATGGACAGCATCGATATCGCCGATCTACGCACGCTCGGCTA
CAGCCAGCAGCAACAGGAGAAGATCAAACCGAAGGTTTCGTTTCGACAGTGGCGCAGC
ACCACGAGGCACTGGTCGGCCACGGGTTTACACACGCGCACATCGTTGCGTTAAGCC
AACACCCGGCAGCGTTAGGGACCGTCGCTGTCAAGTATCAGGACATGATCGCAGCGT
TGCCAGAGGGCGACACACGAAGCGATCGTTGGCGTCGGCAAACAGTGGTCCGGCGCA
CGCGCTCTGGAGGCCTTGCTCACGGTGGCGGGAGAGTTGAGAGGTCCACCGTTACAG
TTGGACACAGGCCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGACCGCAGTGGAG
GCAGTGCATGCATGGCGCAATGCACTGACGGGTGCCCCGCTCAAC

[0133] SEQ ID NO: 11 BnFAD2 (T03-R) Binding Domain

TTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATAATGGTGGCAAGCAGGCGCTG
GAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCACGGCTTGACCCCCCAG

CAGGTGGTGGCCATCGCCAGCAATAATGGTGGCAAGCAGGCGCTGGAGACGGTCCA
 GCGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTTGACCCCCCAGCAGGTGGTGGC
 CATCGCCAGCAATAATGGTGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCC
 GGTGCTGTGCCAGGCCCACGGCTTGACCCCCGGAGCAGGTGGTGGCCATCGCCAGCAA
 TATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCA
 GGCCACGGCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAA
 GCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTT
 GACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGG
 AGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTTGACCCCCCAGC
 AGGTGGTGGCCATCGCCAGCAATAATGGTGGCAAGCAGGCGCTGGAGACGGTCCAG
 CGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTTGACCCCCGGAGCAGGTGGTGGCC
 ATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCC
 GGTGCTGTGCCAGGCCCACGGCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAA
 TGGCGGTGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCA
 GGCCACGGCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAA
 GCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTT
 GACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGG
 AGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTTGACCCCCGGAGC
 AGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTCCAG
 CGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGCTTGACCCCCCAGCAGGTGGTGGCC
 ATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCC
 GGTGCTGTGCCAGGCCCACGGCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAA
 TGGCGGTGGCAAGCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCA
 GGCCACGGCTTGACCCCCCAGCAGGTGGTGGCCATCGCCAGCAATAATGGTGGCAA
 GCAGGCGCTGGAGACGGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCACGGC

[0134] SEQ ID NO: 12 Right TALE C-terminal

TTGACCCCTCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAGGCCGGCGCTG
 GAGAGCATTGTTGCCAGTTATCTCGCCCTGATCCGGCGTTGGCCGCGTTGACCAACG
 ACCACCTCGTCGCCTTGGCCTGCCTCGGCGGGCGTCCTGCGCTGGATGCAGTGAAAA
 AGGGA

[0135] SEQ ID NO: 13 Vector pSW157

CGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGC
 GCAACGCAATTAATACGCGTACCGCTAGCCAGGAAGAGTTTGTAGAAACGCAAAAA
 GGCCATCCGTCAGGATGGCCTTCTGCTTAGTTTGTATGCCTGGCAGTTTATGGCGGGCG
 TCCTGCCCCGCCACCCTCCGGGCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGA

TTTGTCTACTCAGGAGAGCGTTACCGACAAACAACAGATAAAACGAAAGGCCAG
TCTTCCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCAGTTCCTACTCTCGCGTTCC
AATACATCTAGATCCAAGTACATGGTACCATCGGAATCGAGATTGCCTCGGTACAAG
GATTCTAGGTACCTCGCGAATGCATCTAGATCCAATGATCATGAGCGGAGAATTAAG
GGAGTCACGTTATGACCCCCGCCGATGACGCGGGACAAGCCGTTTTACGTTTGGAAC
TGACAGAACCGCAACGTTGAAGGAGCCACTCAGCCGCGGGTTTCTGGAGTTTAATGA
GCTAAGCACATACGTCAGAAACCATTATTGCGCGTTCAAAAGTCGCCTAAGGTCCT
ATCAGCTAGCAAATATTTCTTGTCAAAAATGCTCCACTGACGTTCCATAAATTCCCCT
CGGTATCCAATTAGAGTCTCATATTCCTCTCAATCCAAATAATCTGCACCGGATCTC
GCCCTTACCTGCTAGTCATGGGCGATCCTAAAAAGAAACGTAAGGTCATCGATTACC
CATACGATGTTCCAGATTACGCTATCGATATCGCCGATCTACGCACGCTCGGCTACAG
CCAGCAGCAACAGGAGAAGATCAAACCGAAGGTTTCGTTTCGACAGTGGCGCAGCACC
ACGAGGCACTGGTCGGCCACGGGTTTACACACGCGCACATCGTTGCGTTAAGCCAAC
ACCCGGCAGCGTTAGGGACCGTCGCTGTCAAGTATCAGGACATGATCGCAGCGTTGC
CAGAGGCGACACACGAAGCGATCGTTGGCGTCGGCAAACAGTGGTCCGGCGCACGC
GCTCTGGAGGCCTTGCTCACGGTGGCGGGAGAGTTGAGAGGTCCACCGTTACAGTTG
GACACAGGCCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGACCGCAGTGGAGGC
AGTGCATGCATGGCGCAATGCACTGACGGGTGCCCCGCTCAACTTGACCAGAGACCG
GCGCCGCTACAGGGCGCGTCCCATTTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGG
CGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCA
AGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACG
GCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCC
CTCGAGGTCCTCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTGCGCGCTTGCGGTAA
TCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACA
TACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCA
CATTAAATTGCGTTGCGCTCACTGCCCCGCTTTCACCGGTGGTCTCACCTCAGCAGGTG
GTGGCCATCGCCAGCAATGGCGGCGGCAGGCCGCGCTGGAGAGCATTGTTGCCAG
TTATCTCGCCCTGATCCGGCGTTGGCCGCGTTGACCAACGACCACCTCGTCGCCTTGG
CCTGCCTCGGCGGGCGTCCTGCGCTGGATGCAGTGAAAAAGGGATTGGGGGATCCTA
TCTGCCTTTCATTTCGAACCGAAATTTTGACCGTGGAGTATGGACCTCTTCCAATTGG
AAAGATTGTGTCAGAAGAGATTAAGTGTCTGTTTACTCAGTGGATCCAGAAGGAAG
AGTGTACACCCAAGCTATTGCACAGTGGCATGATAGAGGAGAGCAAGAAGTGCTTG
AGTACGAACTTGAGGATGGTTCTGTGATTAGGGCTACTTCTGATCACAGGTTCTTGAC
CACTGATTACCAAGTTGCTTGCAATTGAGGAAATTTTCGCTAGGCAATTGGATCTTTTG
ACTCTTGAGAACATTAAGCAAACCTGAGGAAGCTCTTGATAACCATAGGCTTCCATTT

CCTTTGCTTGATGCAGGAACCTATTAAGTGATAACTCGAGAAGGGCGCGATCGTTCAA
ACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTAT
CATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACG
TTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGA
TAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTA
TGTTACTAGATCGGGAATTCGTAATCATGGTCATAGCATTGGATCGGATCCCGGGCC
CGTCGACTGCAGAGGCCTGCATGCAAGAGCGTCGGAATCGAGATTGCCTCGGTACAA
GGATATCGATTGTCTTCATCGGATCCCATCCCCTATAGTGAGTCGTATTACATGGTCA
TAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTACATTGCACA
AGATAAAAATATATCATCATGCCTCCTCTAGACCAGCCAGGACAGAAATGCCTCGAC
TTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACACCGTGGAACGGATGAAGGCACG
AACCAGTGGACATAAGCCTGTTCCGTTTCGTAAGCTGTAATGCAAGTAGCGTATGCG
CTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTGGC
GGTTTTCATGGCTTGTTATGACTGTTTTTTTTGGGGTACAGTCTATGCCTCGGGCATCCA
AGCAGCAAGCGCGTTACGCCGTGGGTTCGATGTTTGATGTTATGGAGCAGCAACGATG
TTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAAACATCATGAGGGAAGCGGTG
ATCGCCGAAGTATCGACTCAACTATCAGAGGTAGTTGGCGTCATCGAGCGCCATCTC
GAACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGGCCTGAAG
CCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTAAGGCTTGATGAAACAACG
CGGCGAGCTTTGATCAACGACCTTTTGGAACCTTCGGCTTCCCCTGGAGAGAGCGAG
ATTCTCCGCGCTGTAGAAGTCACCATTGTTGTGCACGACGACATCATTCCGTGGCGTT
ATCCAGCTAAGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGCGAG
GTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAA
GAGAACATAGCGTTGCCTTGGTAGGTCCAGCGGCGGAGGAACTCTTTGATCCGGTTC
CTGAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAACGCTATGGAACCTCGCCGC
CCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTACGTTGTCCCGCATTTGGTACA
GCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGCCGACTGGGCAATGGAG
CGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAGACAGGCTTATCTTGGA
CAAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCAGTTGGAAGAATTTGTCCACTAC
GTGAAAGGCGAGATCACCAAGGTAGTCGGCAAATAACCCTCGAGCCACCCATGACC
AAAATCCCTTAACGTGAGTTACGCGTCGTTCCACTGAGCGTCAGACCCCGTAGAAAA
GATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACA
AAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCAACTCTT
TTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCTTCTAGTGT
AGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT

GCTAATCCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTG
GACTCAAGACGATAGTTACCGGATAAAGGCGCAGCGGTCCGGGCTGAACGGGGGGTTC
GTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCG
TGAGCATTGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGG
TAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCC
TGGTATCTTTATAGTCCTGTCCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTG
ATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCCTTTTTACG
GTTCTTGGCCTTTTTGCTGGCCTTTTTGCTCACATGTTCTTTCTGCGTTATCCCCTGATT
CTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAA
CGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAA
CCGCCTCTCCCCGCGCGTTGGC

[0136] SEQ ID NO: 14 LacZ cassette

CCATTCGCCATTACAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTC
GCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAAC
GCCAGGGTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATA
CGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCCCTCGAGGTCCTCCAGCTTTTG
TTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCT
GTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAG
TGTAAGCCTGGGGTGCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGC

[0137] SEQ ID NO: 15 Vector pSW158

CGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGC
GCAACGCAATTAATACGCGTACCGCTAGCCAGGAAGAGTTTGTAGAAACGCAAAAA
GGCCATCCGTCAGGATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCG
TCCTGCCCCGCCACCCTCCGGGCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGA
TTTGTCTACTCAGGAGAGCGTTACCGACAAACAACAGATAAAACGAAAGGCCCAG
TCTTCCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCAGTTCCCTACTCTCGCGTTTCG
AATACATCTAGATCCAAGTACATGGTGGGAATCGGAATCGAGATTGCCTCGGTACAAG
GAGCGTAGGTACCTCGCGAATGCATCTAGATCCAATGATCATGAGCGGAGAATTAAG
GGAGTCACGTTATGACCCCCGCCGATGACGCGGGACAAGCCGTTTTACGTTTGGAAC
TGACAGAACCGCAACGTTGAAGGAGCCACTCAGCCGCGGGTTTCTGGAGTTTAATGA
GCTAAGCACATACGTCAGAAACCATTATTGCGCGTTCAAAAGTCGCCTAAGGTCACT
ATCAGCTAGCAAATATTTCTTGTCAAAAATGCTCCACTGACGTTCCATAAATCCCCT
CGGTATCCAATTAGAGTCTCATATTCCTCTCAATCCAAATAATCTGCACCGGATCTC
GCCCTTACCTGCTAGTCATGGGCGATCCTAAAAAGAAACGTAAGGTCATCGATAAGG
AGACTGCCGCTGCCAAGTTCGAGAGACAGCACATGGACAGCATCGATATCGCCGATC

TACGCACGCTCGGCTACAGCCAGCAGCAACAGGAGAAGATCAAACCGAAGGTTTCGT
TCGACAGTGGCGCAGCACCACGAGGCACTGGTCGGCCACGGGTTTACACACGCGCAC
ATCGTTGCGTTAAGCCAACACCCGGCAGCGTTAGGGACCGTCGCTGTCAAGTATCAG
GACATGATCGCAGCGTTGCCAGAGGCGACACACGAAGCGATCGTTGGCGTCGGCAA
ACAGTGGTCCGGCGCACGCGCTCTGGAGGCCTTGCTCACGGTGGCGGGAGAGTTGAG
AGGTCCACCGTTACAGTTGGACACAGGCCAACTTCTCAAGATTGCAAAACGTGGCGG
CGTGACCGCAGTGGAGGCAGTGCATGCATGGCGCAATGCACTGACGGGTGCCCCGT
CAACTTGACCAGAGACCGGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTG
CGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCG
AAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCA
CGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGA
ATTGGGTACCGGGCCCCCCTCGAGGTCCTCCAGCTTTTGTTCCTTTAGTGAGGGTT
AATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCG
CTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCC
TAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCCTTTCCACCGGT
GGTCTCACCTCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGGCGGCAGGCCGGCGCT
GGAGAGCATTGTTGCCAGTTATCTCGCCCTGATCCGGCGTTGGCCGCGTTGACCAAC
GACCACCTCGTCGCCTTGGCCTGCCTCGGCGGGCGTCCTGCGCTGGATGCAGTGAAA
AAGGGATTGGGGGATCCTATCTGCCTTTCATTCGGAACCGAAATTTTGACCGTGGAG
TATGGACCTCTTCCAATTGGAAAGATTGTGTGAGAAGAGATTAACTGTTCTGTTTACT
CAGTGGATCCAGAAGGAAGAGTGTACACCCAAGCTATTGCACAGTGGCATGATAGA
GGAGAGCAAGAAGTGCTTGAGTACGAACTTGAGGATGGTTCTGTGATTAGGGCTACT
TCTGATCACAGGTTCTTGACCACTGATTACCAGTTGCTTGCAATTGAGGAAATTTTCG
CTAGGCAATTGGATCTTTTGACTCTTGAGAACATTAAGCAAAGTCTGAGGAAGCTCTTG
ATAACCATAGGCTTCCATTTTCCTTTGCTTGATGCAGGAACTATTAAGTGATAACTCGA
GAAGGGCGCGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTT
GCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAA
TTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCA
ATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAAGTAGGATAAATT
ATCGCGCGCGGTGTCATCTATGTTACTAGATCGGGAATTCGTAATCATGGTCATAGCA
TTGGATCGGATCCCCGGGCCCCGTCGACTGCAGAGGCCTGCATGCAAACCAAGCTCGGA
ATCGAGATTGCCTCGGTACAAGTCAGTTCGATTGTCTTCATCGGATCCCATCCCCTAT
AGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCCGTGTCTCAA
ATCTCTGATGTTACATTGCACAAGATAAAAAATATATCATCATGCCTCCTCTAGACCAG
CCAGGACAGAAATGCCTCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACACC

GTGGAAACGGATGAAGGCACGAACCCAGTGGACATAAGCCTGTTCCGGTTCGTAAGCT
GTAATGCAAGTAGCGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAG
CGGTGGTAACGGGCGCAGTGGCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTGGGGTA
CAGTCTATGCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTG
ATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTA
AACATCATGAGGGAAGCGGTGATCGCCGAAGTATCGACTCAACTATCAGAGGTAGTT
GGCGTCATCGAGCGCCATCTCGAACCGACGTTGCTGGCCGTACATTTGTACGGCTCC
GCAGTGGATGGCGGCCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACC
GTAAGGCTTGATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAACCTTCG
GCTTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTCACCATTGTTGTGCAC
GACGACATCATTCCGTGGCGTTATCCAGCTAAGCGCGAACTGCAATTTGGAGAATGG
CAGCGCAATGACATTCTTGACAGGTATCTTCGAGCCAGCCACGATCGACATTGATCTG
GCTATCTTGCTGACAAAAGCAAGAGAACATAGCGTTGCCTTGGTAGGTCCAGCGGCG
GAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGCTAAATGAAACC
TTAACGCTATGGAACCTCGCCGCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTT
ACGTTGTCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTC
GCTGCCGACTGGGCAATGGAGCGCCTGCCGGCCAGTATCAGCCCGTCATACTTGAA
GCTAGACAGGCTTATCTTGACAAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCAG
TTGGAAGAATTTGTCCACTACGTGAAAGGCGAGATCACCAAGGTAGTCGGCAAATAA
CCCTCGAGCCACCCATGACCAAAAATCCCTTAACGTGAGTTACGCGTCGTTCCACTGA
GCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCG
TAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGG
ATCAAGAGCTACCAACTCTTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATAC
CAAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGC
ACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT
AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGG
TCGGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACC
GAACTGAGATACCTACAGCGTGAGCATTGAGAAAGCGCCACGCTTCCCGAAGGGAG
AAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGG
GAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCCGGTTTCGCCACCTCT
GACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACG
CCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTT
TTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTG
ATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCG
GAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGC

[0138] SEQ ID NO: 16 LacZ cassette

CCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTC
GCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAAC
GCCAGGGTTTTCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATA
CGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCTCGAGGTCCTCCAGCTTTTG
TTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCT
GTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAG
TGTAAGCCTGGGGTGCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGC

[0139] SEQ ID NO: 17 Vector pSW159

CGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGC
GCAACGCAATTAATACGCGTACCGCTAGCCAGGAAGAGTTTGTAGAAACGCAAAAA
GGCCATCCGTCAGGATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCG
TCCTGCCCCGCCACCCTCCGGGCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGA
TTTGTCTACTCAGGAGAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCCCAG
TCTTCCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCAGTTCCTACTCTCGCGTCGA
ATACATCTAGATCCAAGTACATGGATTGATCGGAATCGAGATTGCCTCGGTACAAGC
AAGCGCTTAGGTACCTCGCGAATGCATCTAGATCCAATGATCATGAGCGGAGAATTA
AGGGAGTCACGTTATGACCCCCGCCGATGACGCGGGACAAGCCGTTTTACGTTTGGA
ACTGACAGAACCGCAACGTTGAAGGAGCCACTCAGCCGCGGGTTTCTGGAGTTTAAT
GAGCTAAGCACATACGTCAGAAACCATTATTGCGCGTTCAAAAGTCGCCTAAGGTCA
CTATCAGCTAGCAAATATTTCTTGTCAAAAATGCTCCACTGACGTTCCATAAATTCCC
CTCGGTATCCAATTAGAGTCTCATATTTCACTCTCAATCCAAATAATCTGCACCGGATC
TCGCCCTTACCTGCTAGTCATGGATCCTAAAAAGAAACGTAAGGTCATGGTGAAGGT
TATTGGAAGGAGATCTCTTGGTGTGCAAAGGATTTTCGATATTGGACTTCCACAAGAT
CACAACCTTCCTTTTGGCTAACGGAGCAATTGCTGCTAACAGCCGTTCCCAGCTGGTGA
AGTCCGAGCTGGAGGAGAAGAAATCCGAGTTGAGGCACAAGCTGAAGTACGTGCCC
CACGAGTACATCGAGCTGATCGAGATCGCCCCGGAACAGCACCCAGGACCGTATCCTG
GAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGGCAAGCACCTG
GGCGGCTCCAGGAAGCCCCGACGGCGCCATCTACACCGTGGGCTCCCCCATCGACTAC
GGCGTGATCGTGGACACCAAGGCCTACTCCGGCGGCTACAACCTGCCCATCGGCCAG
GCCGACGAAATGCAGAGGTACGTGGAGGAGAACCAGACCAGGAACAAGCACATCAA
CCCCAACGAGTGGTGAAGGTGTACCCCTCCAGCGTGACCGAGTTCAAGTTCCTGTT
CGTGTCCGGCCACTTCAAGGGCAACTACAAGGCCAGCTGACCAGGCTGAACACAT
CACCAACTGCAACGGCGCCGTGCTGTCCGTGGAGGAGCTCCTGATCGGCGGCGAGAT
GATCAAGGCCCGGCACCCTGACCCTGGAGGAGGTGAGGAGGAAGTTCAACAACGGCG

AGATCAACTTCGCGGCCGACTGATAACTCGAGAAGGGCGCGATCGTTCAAACATTTG
GCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAA
TTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTA
TGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAA
ACAAAATATAGCGCGCAAACCTAGGATAAATTATCGCGCGCGGTGTCTATCTATGTTAC
TAGATCGGGAATTCGTAATCATGGTCATAGCATTGGATCGGATCCCGGGCCCCGTCGA
CTGCAGAGGCCTGCATGCAACCTAGACCAAGACTTCAGCTCGCGTCGTCGGAATCGA
GATTGCCTCGGTACAAGAACGTCGATTGTCTTCATCGGATCCCATCCCCTATAGTGAG
TCGTATTACATGGTCATAGCTGTTTCTTGGCAGCTCTGGCCCGTGTCTCAAAATCTCT
GATGTTACATTGCACAAGATAAAAAATATATCATCATGCCTCCTCTAGACCAGCCAGG
ACAGAAATGCCTCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACACCGTGGA
AACGGATGAAGGCACGAACCCAGTGGACATAAGCCTGTTTCGGTTCGTAAGCTGTAAT
GCAAGTAGCGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTG
GTAACGGCGCAGTGGCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTGGGGTACAGTC
TATGCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTGCGATGTTTGATGTT
ATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAAACA
TCATGAGGGAAGCGGTGATCGCCGAAGTATCGACTCAACTATCAGAGGTAGTTGGCG
TCATCGAGCGCCATCTCGAACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGT
GGATGGCGGCCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTAAG
GCTTGATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAACTTCGGCTTC
CCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTCACCATTGTTGTGCACGACGA
CATCATTCCGTGGCGTTATCCAGCTAAGCGCGAACTGCAATTTGGAGAATGGCAGCG
CAATGACATTCTTGCAGGTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTATC
TTGCTGACAAAAGCAAGAGAACATAGCGTTGCCTTGGTAGGTCCAGCGGCGGAGGA
ACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAACG
CTATGGAACTCGCCGCCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTACGTTG
TCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGCC
GACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAGA
CAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCAGTTGGAA
GAATTTGTCCACTACGTGAAAGGCGAGATACCAAGGTAGTCGGCAAATAACCCTCG
AGCCACCCATGACCAAAATCCCTTAACGTGAGTTACGCGTCGTTCCACTGAGCGTCA
GACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCT
GCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAG
AGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATA
CTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCC

TACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCG
TGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCCGACGCGGTTCGGGC
TGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTG
AGATACCTACAGCGTGAGCATTGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGC
GGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTC
CAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGA
GCGTCGATTTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAA
CGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCTG
CGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGC
TCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGC
GCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGC

[0140] SEQ ID NO: 18 Ssp DnaE int-C

ATGGTGAAGGTTATTGGAAGGAGATCTCTTGGTGTGCAAAGGATTTTCGATATTGGA
CTTCCACAAGATCACAACCTTCCTTTTGGCTAACGGAGCAATTGCTGCTAAC

[0141] SEQ ID NO: 19 Linker

AGCCGTTCC

[0142] SEQ ID NO: 20 FokI

CAGCTGGTGAAGTCCGAGCTGGAGGAGAAGAAATCCGAGTTGAGGCACAAGCTGAA
GTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCCGGAACAGCACCCAGGA
CCGTATCCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGG
CAAGCACCTGGGCGGCTCCAGGAAGCCCGACGGCGCCATCTACACCGTGGGCTCCCC
CATCGACTACGGCGTGATCGTGGACACCAAGGCCTACTCCGGCGGCTACAACCTGCC
CATCGGCCAGGCCGACGAAATGCAGAGGTACGTGGAGGAGAACCAGACCAGGAACA
AGCACATCAACCCCAACGAGTGGTGGAAAGGTGTACCCCTCCAGCGTGACCGAGTTCA
AGTTCCTGTTTCGTGTCCGGCCACTTCAAGGGCAACTACAAGGCCCAGCTGACCAGGC
TGAACCACATCACCAACTGCAACGGCGCCGTGCTGTCCGTGGAGGAGCTCCTGATCG
GCGGCGAGATGATCAAGGCCGGCACCCCTGACCCTGGAGGAGGTGAGGAGGAAGTTC
AACAACGGCGAGATCAACTTCGCGGCCGACTGATAA

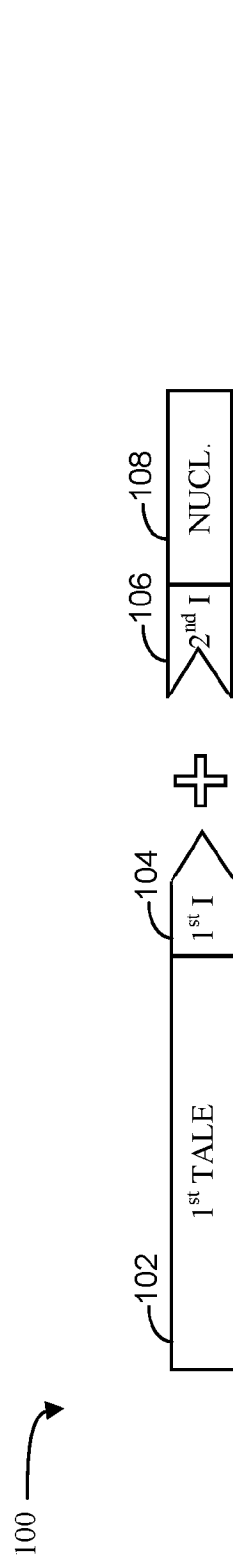


FIG. 1A

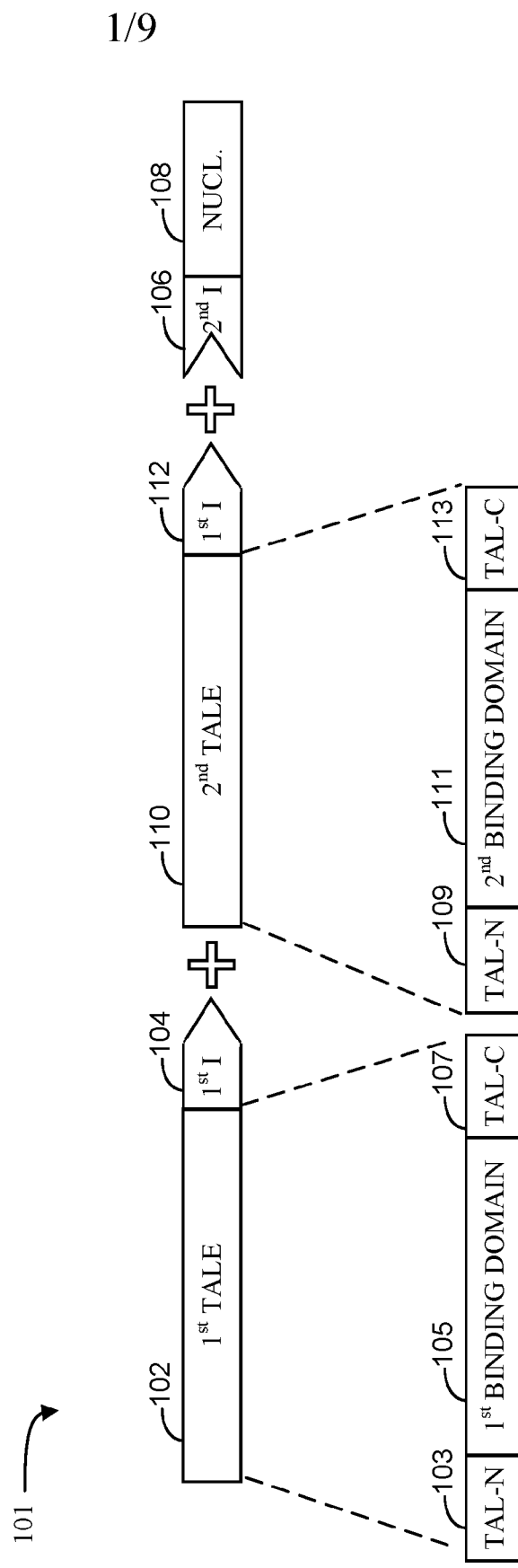


FIG. 1B

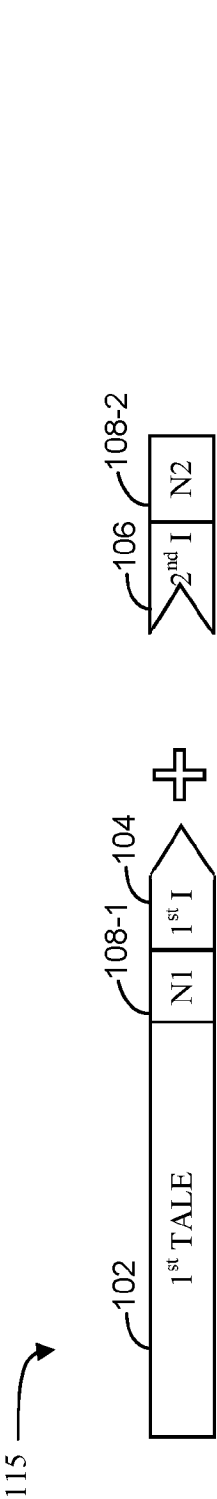


FIG. 1C

2/9

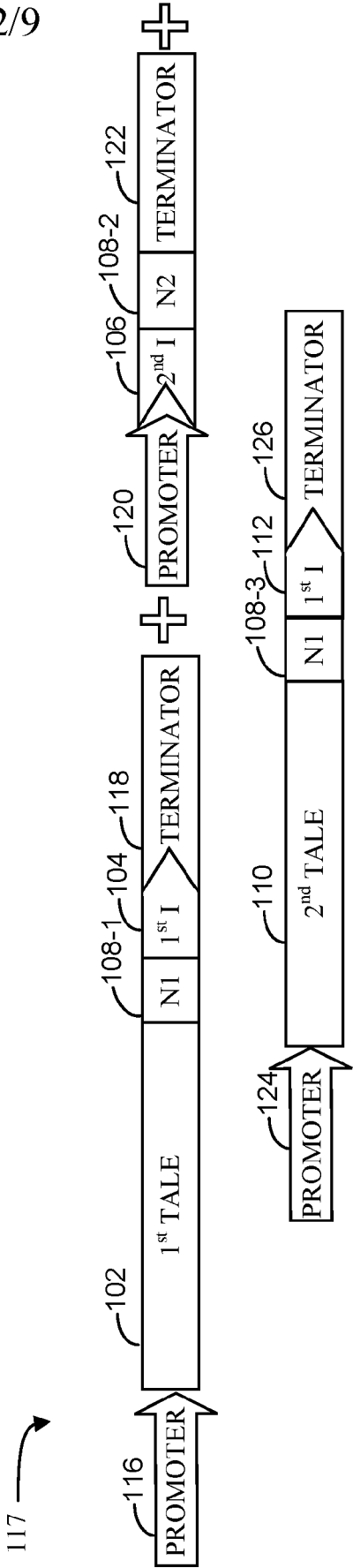


FIG. 1D

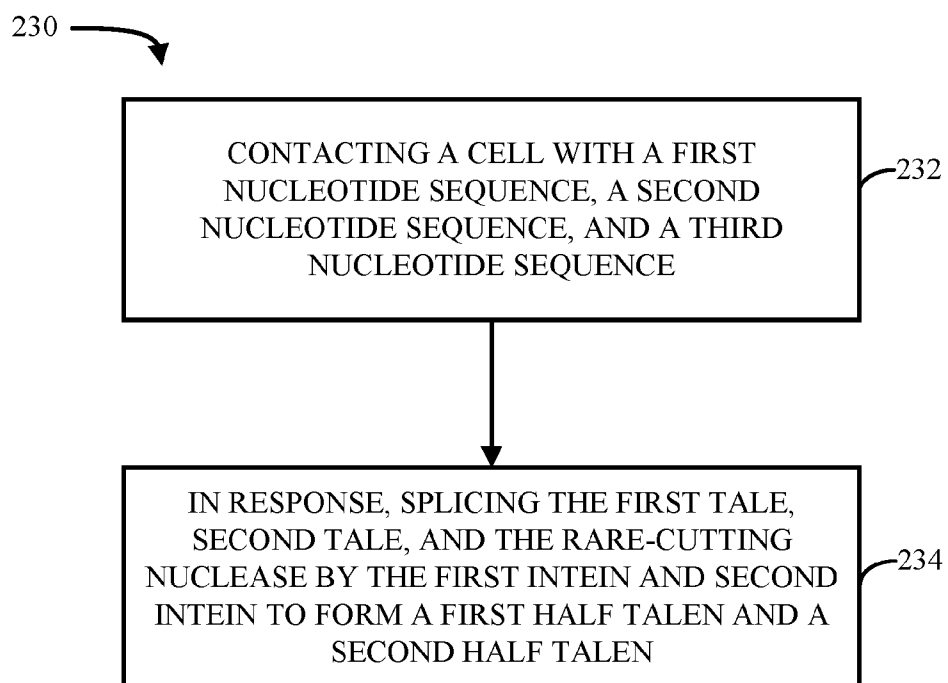


FIG. 2

340 →

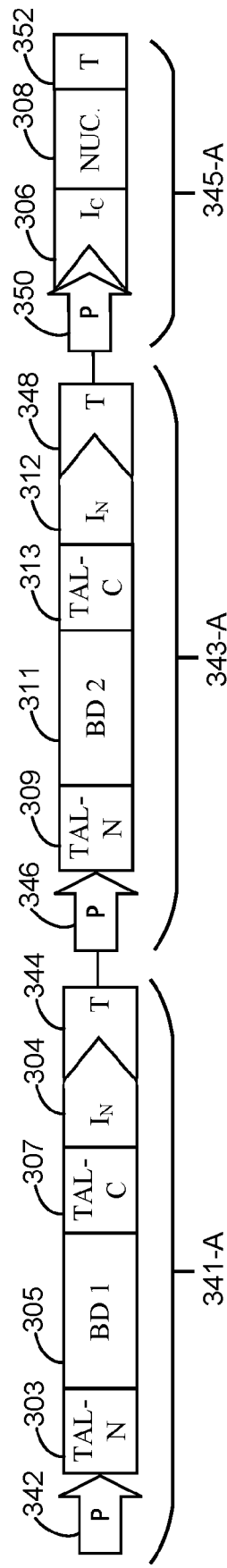


FIG. 3A

360 →

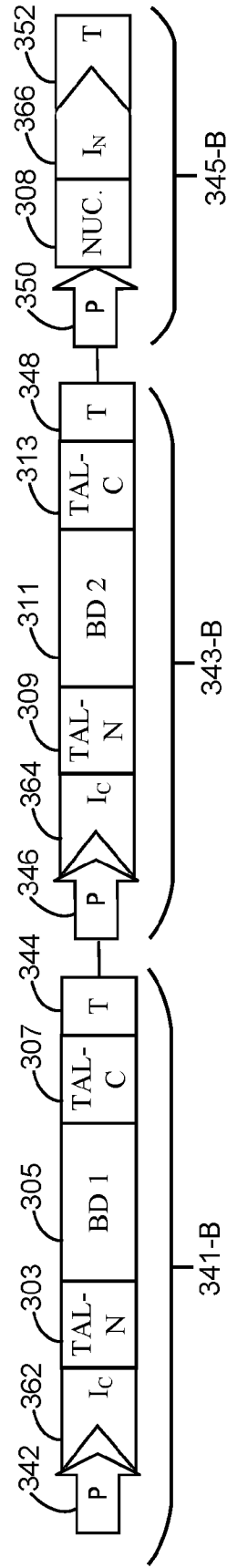


FIG. 3B

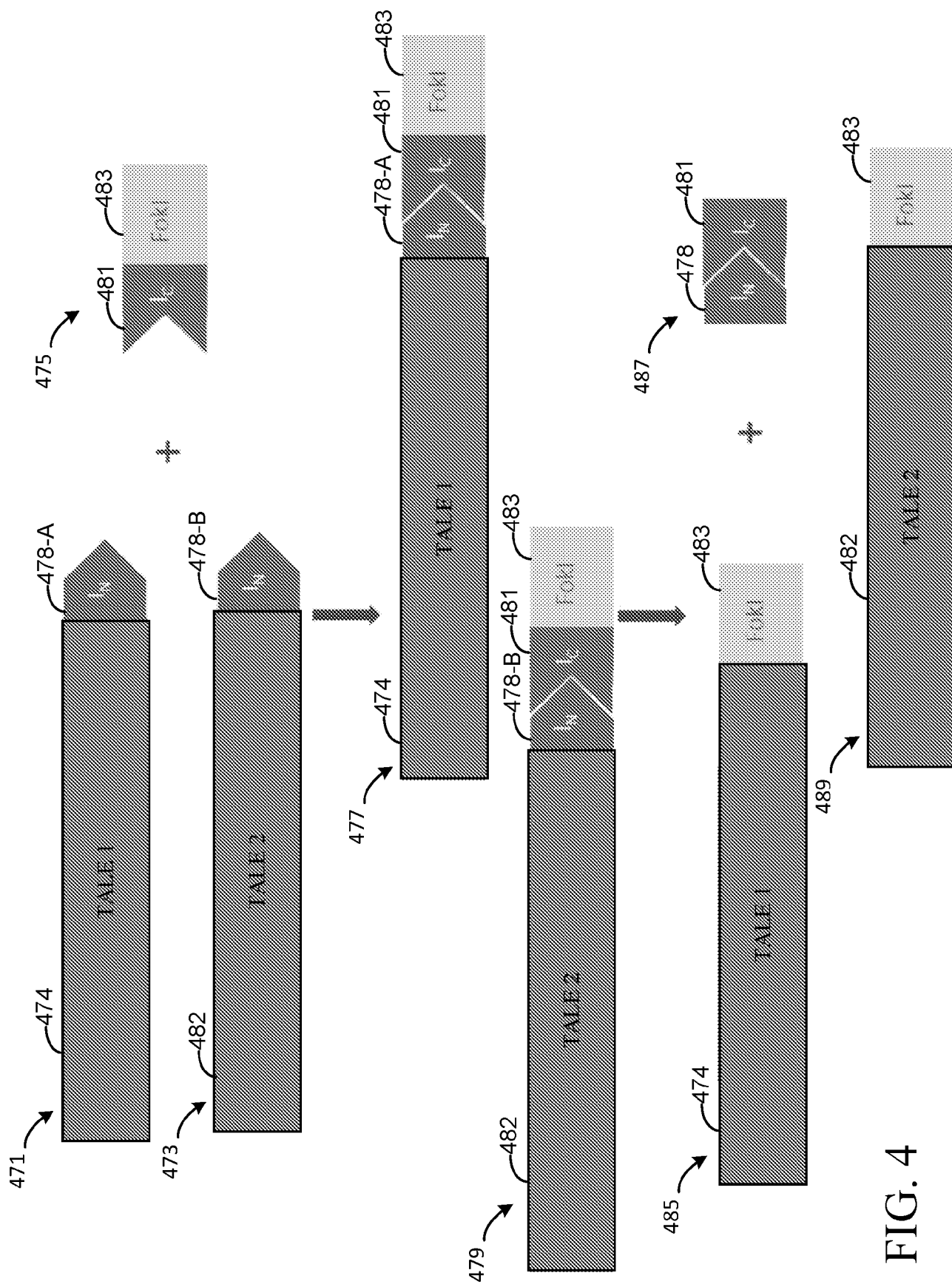


FIG. 4

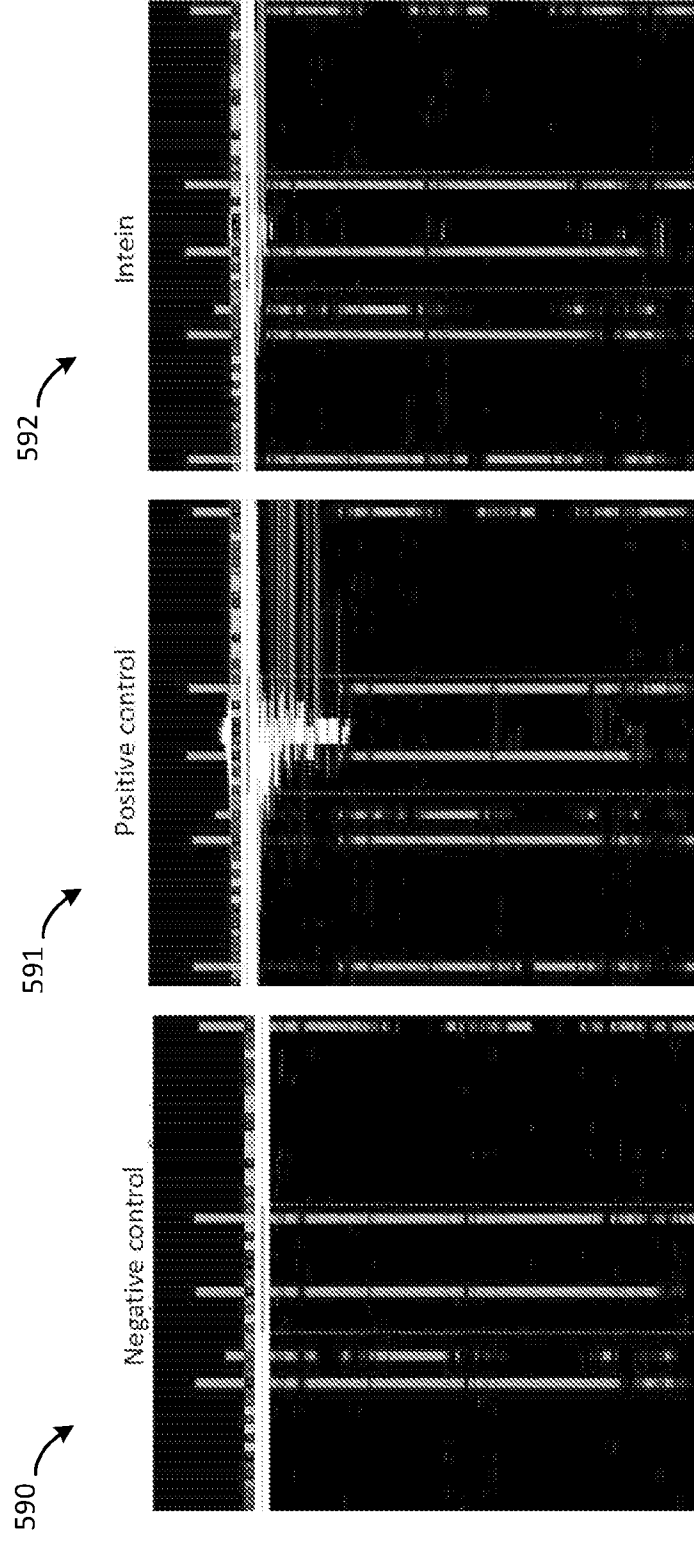


FIG. 5

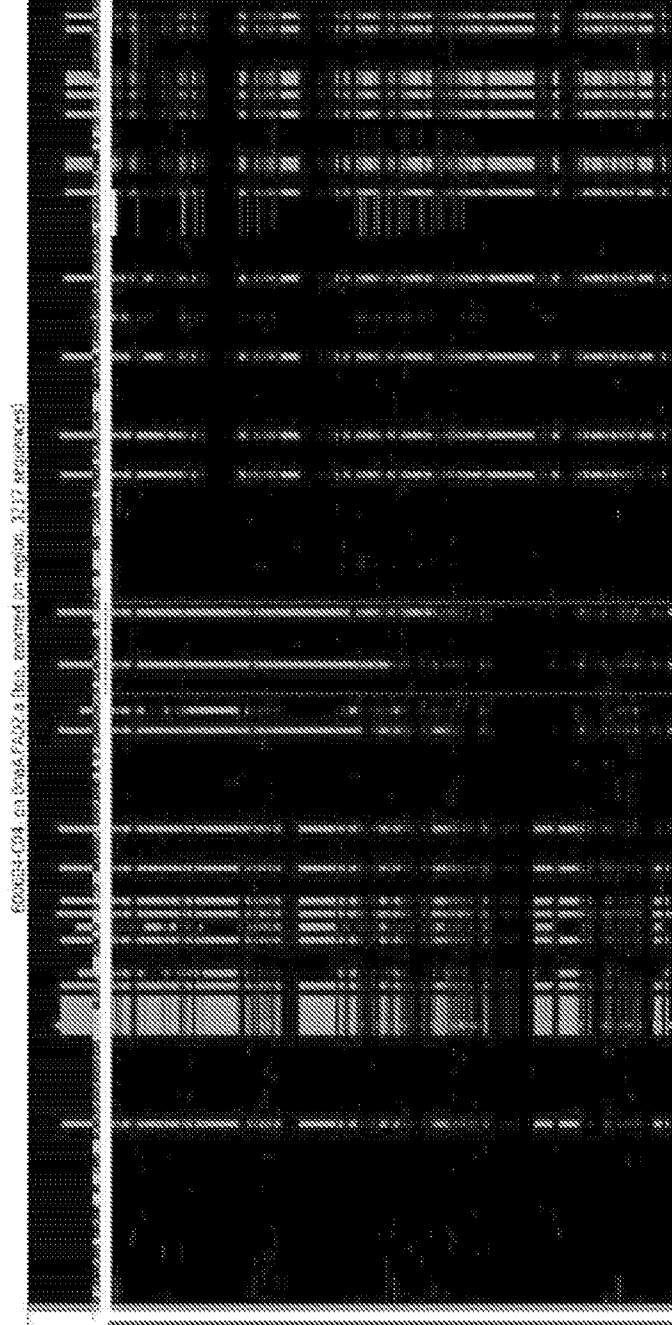


FIG. 6A

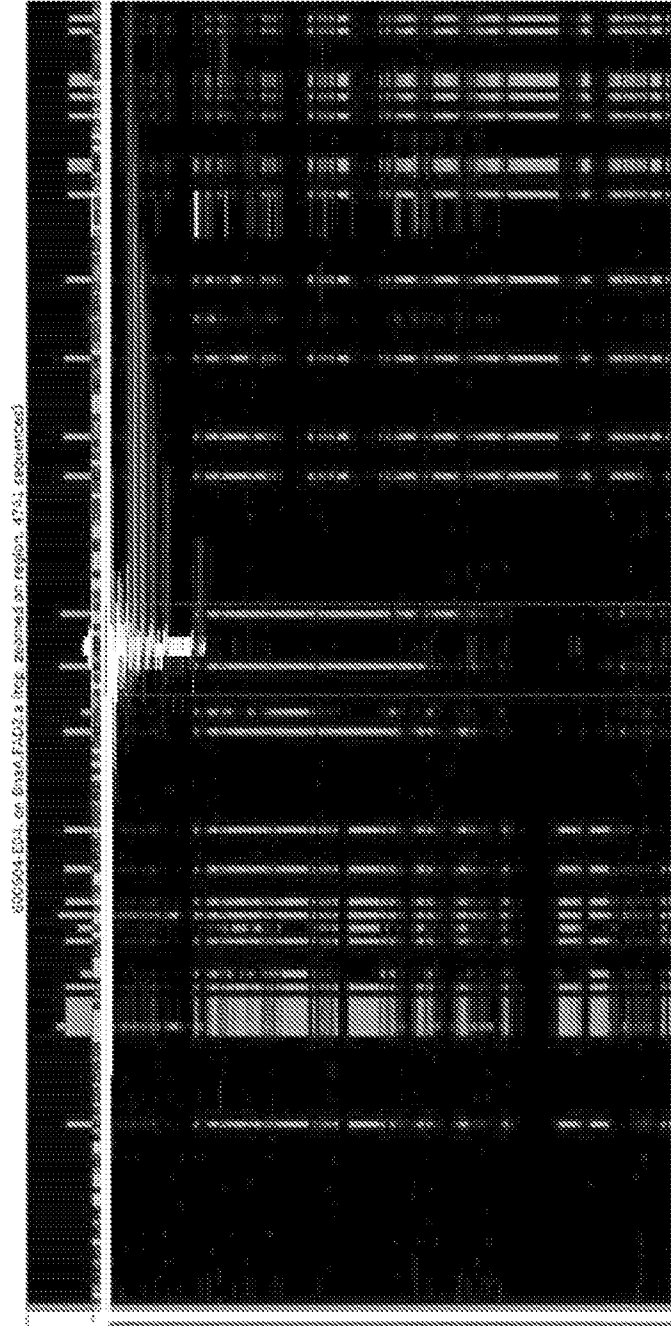


FIG. 6B

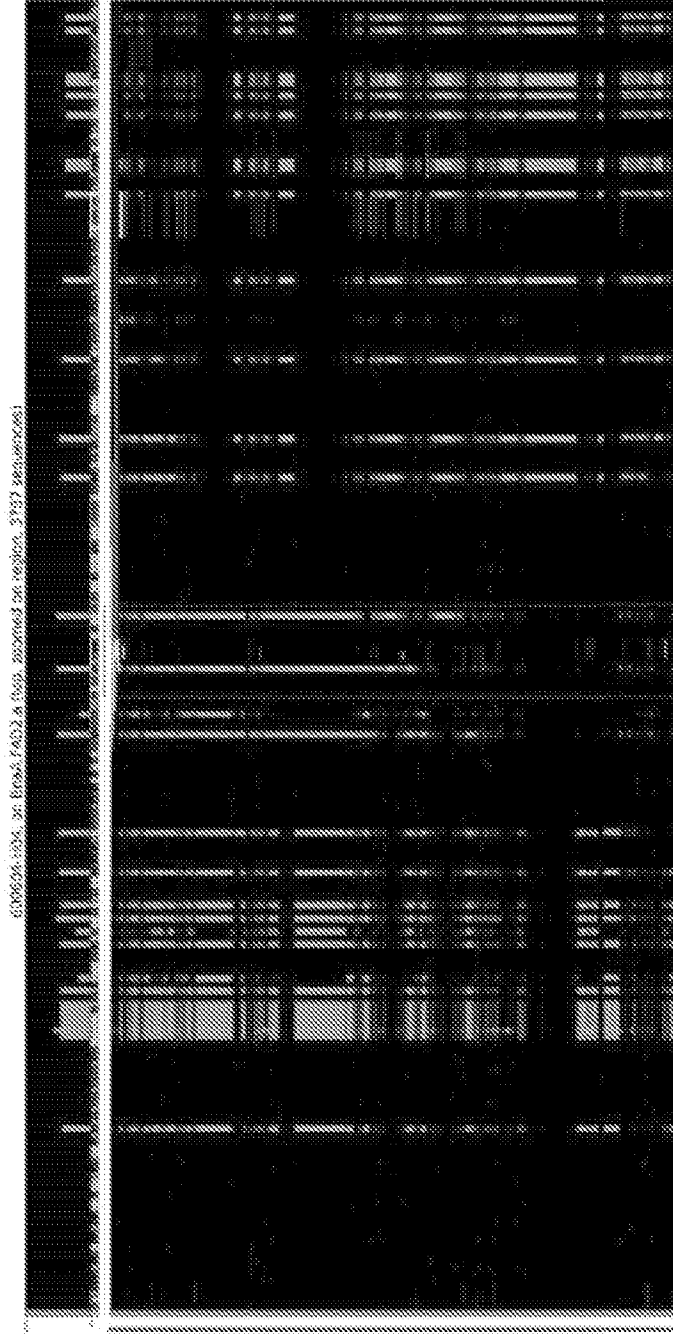


FIG. 6C

Doc Code: **TR.PROV**

Document Description: Provisional Cover Sheet (SB16)

PTO/SB/16 (02-18)

Approved for use through 11/30/2020. OMB 0651-0032

U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number

Provisional Application for Patent Cover Sheet					
This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c)					
Inventor(s)					
Inventor 1					Remove
Given Name	Middle Name	Family Name	City	State	Country ;
Sara		Wixon	Minneapolis	MN	US
Inventor 2					Remove
Given Name	Middle Name	Family Name	City	State	Country ;
Robert	W.	Williams	Minneapolis	MN	US
All Inventors Must Be Listed – Additional Inventor Information blocks may be generated within this form by selecting the Add button.					Add
Title of Invention		TRANSCRIPTION ACTIVATOR-LIKE EFFECTORS FUSED TO INTEINS			
Attorney Docket Number (if applicable)		C1633.112.101			
Correspondence Address					
Direct all correspondence to (select one):					
☒ The address corresponding to Customer Number			☐ Firm or Individual Name		
Customer Number			025281		

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.	
☒ No.	
☐ Yes, the invention was made by an agency of the United States Government. The U.S. Government agency name is:	
☐ Yes, the invention was under a contract with an agency of the United States Government. The name of the U.S. Government agency and Government contract number are:	

Doc Code: **TR.PROV**

Document Description: Provisional Cover Sheet (SB16)

PTO/SB/16 (02-18)

Approved for use through 11/30/2020. OMB 0651-0032

U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number

Entity Status

Applicant asserts small entity status under 37 CFR 1.27 or applicant certifies micro entity status under 37 CFR 1.29

- ☒ Applicant asserts small entity status under 37 CFR 1.27
- ☐ Applicant certifies micro entity status under 37 CFR 1.29. Applicant must attach form PTO/SB/15A or B or equivalent.
- ☐ No

Warning

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

Signature

Please see 37 CFR 1.4(d) for the form of the signature.

Signature	/Abigail A. Lorfing/			Date (YYYY-MM-DD)	2021-07-07
First Name	Abigail	Last Name	Lorfing	Registration Number (If appropriate)	72266

This collection of information is required by 37 CFR 1.51. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **This form can only be used when in conjunction with EFS-Web. If this form is mailed to the USPTO, it may cause delays in handling the provisional application.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that : (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	C1633.112.101
		Application Number	
Title of Invention	TRANSCRIPTION ACTIVATOR-LIKE EFFECTORS FUSED TO INTEINS		
The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76. This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.			

Secrecy Order 37 CFR 5.2:

☐	Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)
--------------------------	---

Inventor Information:

Inventor 1						
Legal Name						
Prefix	Given Name	Middle Name	Family Name	Suffix		
	Sara		Wixon			
Residence Information (Select One) ☒ US Residency ☐ Non US Residency ☐ Active US Military Service						
City	Minneapolis	State/Province	MN	Country of Residence	US	
Mailing Address of Inventor:						
Address 1		2800 Mount Ridge Road				
Address 2		c/o Calyxt. Inc.				
City	Roseville	State/Province	MN			
Postal Code	55113	Country	US			
Inventor 2						
Legal Name						
Prefix	Given Name	Middle Name	Family Name	Suffix		
	Robert	W.	Williams			
Residence Information (Select One) ☒ US Residency ☐ Non US Residency ☐ Active US Military Service						
City	Minneapolis	State/Province	MN	Country of Residence	US	
Mailing Address of Inventor:						
Address 1		2800 Mount Ridge Road				
Address 2		c/o Calyxt. Inc.				
City	Roseville	State/Province	MN			
Postal Code	55113	Country	US			
All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the Add button.						
						

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below. For further information see 37 CFR 1.33(a).
--

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	C1633.112.101
		Application Number	
Title of Invention	TRANSCRIPTION ACTIVATOR-LIKE EFFECTORS FUSED TO INTEINS		

☐ An Address is being provided for the correspondence information of this application.

Customer Number	025281		
Email Address	USPTO.PATENTS@DBCLAW.COM	Add Email	Remove Email

Application Information:

Title of the Invention	TRANSCRIPTION ACTIVATOR-LIKE EFFECTORS FUSED TO INTEINS		
Attorney Docket Number	C1633.112.101	Small Entity Status Claimed	☒
Application Type	Provisional		
Subject Matter	Utility		
Total Number of Drawing Sheets (if any)	9	Suggested Figure for Publication (if any)	

Filing By Reference:

Only complete this section when filing an application by reference under 35 U.S.C. 111(c) and 37 CFR 1.57(a). Do not complete this section if application papers including a specification and any drawings are being filed. Any domestic benefit or foreign priority information must be provided in the appropriate section(s) below (i.e., "Domestic Benefit/National Stage Information" and "Foreign Priority Information").

For the purposes of a filing date under 37 CFR 1.53(b), the description and any drawings of the present application are replaced by this reference to the previously filed application, subject to conditions and requirements of 37 CFR 1.57(a).

Application number of the previously filed application	Filing date (YYYY-MM-DD)	Intellectual Property Authority or Country

Publication Information:

☐ Request Early Publication (Fee required at time of Request 37 CFR 1.219)

☐ **Request Not to Publish.** I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application **has not and will not** be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32). Either enter Customer Number or complete the Representative Name section below. If both sections are completed the customer Number will be used for the Representative Information during processing.

Please Select One:	☒ Customer Number	☐ US Patent Practitioner	☐ Limited Recognition (37 CFR 11.9)
Customer Number	025281		

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	C1633.112.101
		Application Number	
Title of Invention	TRANSCRIPTION ACTIVATOR-LIKE EFFECTORS FUSED TO INTEINS		

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, 365(c), or 386(c) or indicate National Stage entry from a PCT application. Providing benefit claim information in the Application Data Sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78.

When referring to the current application, please leave the "Application Number" field blank.

Prior Application Status			Remove
Application Number	Continuity Type	Prior Application Number	Filing or 371(c) Date (YYYY-MM-DD)
Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the Add button.			

Foreign Priority Information:

This section allows for the applicant to claim priority to a foreign application. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55. When priority is claimed to a foreign application that is eligible for retrieval under the priority document exchange program (PDX)ⁱ the information will be used by the Office to automatically attempt retrieval pursuant to 37 CFR 1.55(i)(1) and (2). Under the PDX program, applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period specified in 37 CFR 1.55(g)(1).

Application Number	Country ⁱ	Filing Date (YYYY-MM-DD)	Remove
			Access Code ⁱ (if applicable)
Additional Foreign Priority Data may be generated within this form by selecting the Add button.			

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications

This application (1) claims priority to or the benefit of an application filed before March 16, 2013 and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013. ☐ NOTE: By providing this statement under 37 CFR 1.55 or 1.78, this application, with a filing date on or after March 16, 2013, will be examined under the first inventor to file provisions of the AIA.
--

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	C1633.112.101
		Application Number	
Title of Invention	TRANSCRIPTION ACTIVATOR-LIKE EFFECTORS FUSED TO INTEINS		

Authorization or Opt-Out of Authorization to Permit Access:

When this Application Data Sheet is properly signed and filed with the application, applicant has provided written authority to permit a participating foreign intellectual property (IP) office access to the instant application-as-filed (see paragraph A in subsection 1 below) and the European Patent Office (EPO) access to any search results from the instant application (see paragraph B in subsection 1 below).

Should applicant choose not to provide an authorization identified in subsection 1 below, applicant **must opt-out** of the authorization by checking the corresponding box A or B or both in subsection 2 below.

NOTE: This section of the Application Data Sheet is **ONLY** reviewed and processed with the **INITIAL** filing of an application. After the initial filing of an application, an Application Data Sheet cannot be used to provide or rescind authorization for access by a foreign IP office(s). Instead, Form PTO/SB/39 or PTO/SB/69 must be used as appropriate.

1. Authorization to Permit Access by a Foreign Intellectual Property Office(s)

A. Priority Document Exchange (PDX) - Unless box A in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the State Intellectual Property Office of the People's Republic of China (SIPO), the World Intellectual Property Organization (WIPO), and any other foreign intellectual property office participating with the USPTO in a bilateral or multilateral priority document exchange agreement in which a foreign application claiming priority to the instant patent application is filed, access to: (1) the instant patent application-as-filed and its related bibliographic data, (2) any foreign or domestic application to which priority or benefit is claimed by the instant application and its related bibliographic data, and (3) the date of filing of this Authorization. See 37 CFR 1.14(h)(1).

B. Search Results from U.S. Application to EPO - Unless box B in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the EPO access to the bibliographic data and search results from the instant patent application when a European patent application claiming priority to the instant patent application is filed. See 37 CFR 1.14(h)(2).

The applicant is reminded that the EPO's Rule 141(1) EPC (European Patent Convention) requires applicants to submit a copy of search results from the instant application without delay in a European patent application that claims priority to the instant application.

2. Opt-Out of Authorizations to Permit Access by a Foreign Intellectual Property Office(s)

☐ A. Applicant **DOES NOT** authorize the USPTO to permit a participating foreign IP office access to the instant application-as-filed. If this box is checked, the USPTO will not be providing a participating foreign IP office with any documents and information identified in subsection 1A above.

☐ B. Applicant **DOES NOT** authorize the USPTO to transmit to the EPO any search results from the instant patent application. If this box is checked, the USPTO will not be providing the EPO with search results from the instant application.

NOTE: Once the application has published or is otherwise publicly available, the USPTO may provide access to the application in accordance with 37 CFR 1.14.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	C1633.112.101
		Application Number	
Title of Invention	TRANSCRIPTION ACTIVATOR-LIKE EFFECTORS FUSED TO INTEINS		

Applicant Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.			
Applicant 1			
If the applicant is the inventor (or the remaining joint inventor or inventors under 37 CFR 1.45), this section should not be completed. The information to be provided in this section is the name and address of the legal representative who is the applicant under 37 CFR 1.43; or the name and address of the assignee, person to whom the inventor is under an obligation to assign the invention, or person who otherwise shows sufficient proprietary interest in the matter who is the applicant under 37 CFR 1.46. If the applicant is an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest) together with one or more joint inventors, then the joint inventor or inventors who are also the applicant should be identified in this section.			
Clear			
☒ Assignee	☐ Legal Representative under 35 U.S.C. 117	☐ Joint Inventor	
☐ Person to whom the inventor is obligated to assign.		☐ Person who shows sufficient proprietary interest	
If applicant is the legal representative, indicate the authority to file the patent application, the inventor is:			
Name of the Deceased or Legally Incapacitated Inventor:			
If the Applicant is an Organization check here. ☒			
Organization Name	Calyx, Inc.		
Mailing Address Information For Applicant:			
Address 1	2800 Mount Ridge Road		
Address 2			
City	Roseville	State/Province	MN
Country ¹	US	Postal Code	55113
Phone Number		Fax Number	
Email Address			
Additional Applicant Data may be generated within this form by selecting the Add button.			

Assignee Information including Non-Applicant Assignee Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	C1633.112.101
		Application Number	
Title of Invention	TRANSCRIPTION ACTIVATOR-LIKE EFFECTORS FUSED TO INTEINS		

Assignee 1

Complete this section if assignee information, including non-applicant assignee information, is desired to be included on the patent application publication. An assignee-applicant identified in the "Applicant Information" section will appear on the patent application publication as an applicant. For an assignee-applicant, complete this section only if identification as an assignee is also desired on the patent application publication.

If the Assignee or Non-Applicant Assignee is an Organization check here. ☐

Prefix	Given Name	Middle Name	Family Name	Suffix

Mailing Address Information For Assignee including Non-Applicant Assignee:

Address 1			
Address 2			
City		State/Province	
Country ¹		Postal Code	
Phone Number		Fax Number	
Email Address			

Additional Assignee or Non-Applicant Assignee Data may be generated within this form by selecting the Add button.

Signature:

NOTE: This Application Data Sheet must be signed in accordance with 37 CFR 1.33(b). **However, if this Application Data Sheet is submitted with the INITIAL filing of the application and either box A or B is not checked in subsection 2 of the "Authorization or Opt-Out of Authorization to Permit Access" section, then this form must also be signed in accordance with 37 CFR 1.14(c).**

This Application Data Sheet **must** be signed by a patent practitioner if one or more of the applicants is a **juristic entity** (e.g., corporation or association). If the applicant is two or more joint inventors, this form must be signed by a patent practitioner, **all** joint inventors who are the applicant, or one or more joint inventor-applicants who have been given power of attorney (e.g., see USPTO Form PTO/AIA/81) on behalf of **all** joint inventor-applicants.

See 37 CFR 1.4(d) for the manner of making signatures and certifications.

Signature	/Abigail A. Lorfing/		Date (YYYY-MM-DD)	2021-07-07
First Name	Abigail	Last Name	Lorfing	Registration Number
72266				

Additional Signature may be generated within this form by selecting the Add button.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	C1633.112.101
		Application Number	
Title of Invention	TRANSCRIPTION ACTIVATOR-LIKE EFFECTORS FUSED TO INTEINS		

This collection of information is required by 37 CFR 1.76. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 23 minutes to complete, including gathering, preparing, and submitting the completed application data sheet form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Electronic Acknowledgement Receipt

EFS ID:	43190740
Application Number:	63219291
International Application Number:	
Confirmation Number:	4201
Title of Invention:	TRANSCRIPTION ACTIVATOR-LIKE EFFECTORS FUSED TO INTEINS
First Named Inventor/Applicant Name:	Sara Wixon
Customer Number:	25281
Filer:	Abigail Ann Lorfing/Cindy Koski
Filer Authorized By:	Abigail Ann Lorfing
Attorney Docket Number:	C1633.112.101
Receipt Date:	07-JUL-2021
Filing Date:	
Time Stamp:	19:31:04
Application Type:	Provisional

Payment information:

Submitted with Payment	yes
Payment Type	DA
Payment was successfully received in RAM	\$ 150
RAM confirmation Number	E202177J32151542
Deposit Account	082025
Authorized User	Cindy Koski
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: 37 CFR 1.17 (Patent application and reexamination processing fees)	

File Listing:					
Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		2021-07-07ProvisionalApplication.pdf	206079	yes	51
			5f2ad7ea5a4ddfe5dcb212aa2b1c0cb3492f2083		
	Multipart Description/PDF files in .zip description				
	Document Description		Start	End	
	Abstract		51	51	
	Claims		46	50	
	Specification		1	45	
Warnings:					
Information:					
2	Drawings-only black and white line drawings	2021-07-07Drawings.pdf	788280	no	9
			7ad1c1cca54c2a1139bc626ccea4ee09174a59e		
Warnings:					
Information:					
3	Provisional Cover Sheet (SB16)	2021-07-07ProvisionalSheet.pdf	77559	no	3
			e0323d4398e304cd36654e8cc915b1abb7841e3c		
Warnings:					
This is not a USPTO supplied Provisional Cover Sheet SB16 form.					
Information:					
4	Application Data Sheet	2021-07-07ADS.pdf	120969	no	8
			6f11a715e07bfd7c692ae73eaa239fd38cc8cf9		
Warnings:					
Information:					
This is not an USPTO supplied ADS fillable form					

5	Fee Worksheet (SB06)	fee-info.pdf	29729 4c803a0c77b38d377f602532725e9731e5c4b0a7	no	2
Warnings:					
Information:					
Total Files Size (in bytes):			1222616		
This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503. <u>New Applications Under 35 U.S.C. 111</u> If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application. <u>National Stage of an International Application under 35 U.S.C. 371</u> If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course. <u>New International Application Filed with the USPTO as a Receiving Office</u> If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.					