

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 979 391**

51 Int. Cl.:

C12N 15/66 (2006.01)

C12N 15/10 (2006.01)

C12N 15/64 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.05.2018** **PCT/GB2018/051174**

87 Fecha y número de publicación internacional: **08.11.2018** **WO18203056**

96 Fecha de presentación y número de la solicitud europea: **02.05.2018** **E 18728221 (5)**

97 Fecha y número de publicación de la concesión europea: **03.01.2024** **EP 3619311**

54 Título: **Ensamblaje del ADN**

30 Prioridad:

03.05.2017 GB 201707049

23.08.2017 GB 201713546

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
25.09.2024

73 Titular/es:

TRIPLE HELIX BIOTECHNOLOGY LTD (100.0%)
12 New Fetter Lane
London EC4A1JP, GB

72 Inventor/es:

LIN, DA y
O'CALLAGHAN, CHRIS

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 979 391 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Ensamblaje del ADN

Esta invención se refiere a métodos de ensamblaje de ADN y a material de ácido nucleico relacionado.

Los avances en biología y biotecnología han llevado al desarrollo de una variedad de métodos para ensamblar grandes construcciones de ADN. Los métodos existentes pueden dividirse en dos grandes grupos: métodos basados en la homología, que utilizan secuencias largas y solapadas entre fragmentos para especificar el orden de ensamblaje, como el ensamblaje Gibson y los enfoques de recombinación homóloga *in vivo* basados en levaduras o *B. subtilis*, y los métodos basados en enzimas de restricción/recombinasas que utilizan una secuencia específica de enzimas definida para especificar el orden de ensamblaje, como Moclo (revisado en Nat Rev Mol Cell Biol. septiembre de 2015;16(9):568-76. doi: 10.1038/nrm4014 (PMID 2608161)2 y Crit Rev Biotecnología. mayo de 2017; 37(3):277-286. doi: 10.3109/07388551.2016.1141394 (PMID 26863154). Una lucha constante es lograr un equilibrio entre la reducción de las limitaciones del diseño de secuencias, como la secuencia prohibida en el ADN ensamblado y la secuencia de cicatrices no deseadas introducidas durante el proceso de ensamblaje, y la mejora de la modularidad del ensamblaje y la reutilización de las piezas.

El sistema inicial de ensamblaje de ADN basado en enzimas de restricción tipo IIS se denominó ensamblaje Golden Gate. En el sistema de ensamblaje Golden Gate, los fragmentos de ADN insertados que se ensamblarán dentro de los plásmidos insertados están flanqueados por sitios de reconocimiento para una enzima de restricción de tipo IIS que corta fuera de la secuencia de reconocimiento y genera extremos adhesivos arbitrarios con secuencias independientes de la secuencia de reconocimiento de la enzima. Los plásmidos de inserción para el ensamblaje de Golden Gate contienen inserciones flanqueadas por un sitio de restricción de tipo IIS dentro del esqueleto del vector de inserción frente a los insertos, de modo que una vez que los insertos son liberados del plásmido por la enzima de restricción de tipo IIS, no contienen el sitio de restricción. El vector de ensamblaje para el ensamblaje Golden Gate contiene un marcador de selección negativo flanqueado por sitios de tipo IIS ubicados en el marcador de selección frente al vector, de modo que una vez que el marcador de selección se libera del vector, la columna vertebral del vector no contiene el sitio de restricción de tipo IIS. Estos plásmidos especialmente diseñados permiten el ensamblaje simple en un solo recipiente de múltiples fragmentos de ADN mezclando plásmidos insertados y vectores junto con una enzima de restricción de tipo IIS y ADN ligasa. En la reacción de ensamblaje Golden Gate, la digestión de restricción de los insertos y el vector y la ligación entre los insertos y el vector se producen en el mismo tubo. Se generará una mezcla de productos de ligadura durante la reacción en un solo recipiente cuando los insertos y la columna vertebral del vector de ensamblaje tengan extremos adhesivos compatibles. Esto incluye productos de ligamiento favorables entre los insertos y el esqueleto del vector de ensamblaje, así como productos de ligamiento desfavorables tales como entre el marcador de selección negativo y el esqueleto del vector de ensamblaje, entre el marcador de selección negativo y el esqueleto del vector de inserto, y entre el inserto y el esqueleto del vector de inserto. Debido a que los sitios de restricción de tipo IIS están ubicados dentro del marcador de selección negativa y la cadena principal del vector de inserción, todos los productos de ligación desfavorables son susceptibles a la digestión por la enzima de restricción de tipo IIS, mientras que los productos de ligación favorables son refractarios a la digestión. Como resultado, la reacción favorece la generación de productos de ligación ensamblados correctamente entre los insertos y la columna vertebral del vector de ensamblaje a largo plazo. Seguido de la selección usando marcadores de selección de antibióticos positivos dentro de la estructura principal del vector de ensamblaje y un marcador de selección negativo dentro del inserto del vector de ensamblaje, se pudieron seleccionar con alta eficiencia plásmidos que contenían insertos correctamente ensamblados.

Los sistemas basados en Moclo se desarrollaron a partir del sistema de ensamblaje Golden Gate agregando un segundo par de sitios de restricción tipo IIS diferentes de la enzima utilizada para el ensamblaje al vector de ensamblaje para liberar el ADN ensamblado para el ensamblaje de la siguiente etapa. En el ensamblaje Golden Gate, el ADN ensamblado no pudo ser liberado por la misma enzima de restricción utilizada para el ensamblaje, porque los insertos y el esqueleto del vector plasmídico de ensamblaje carecen de la secuencia de reconocimiento de la enzima de restricción tipo IIS utilizada. Al agregar un par diferente de sitios de restricción de tipo IIS para liberar el fragmento ensamblado, el proceso de ensamblaje de Golden Gate podría repetirse usando esta segunda enzima de restricción y un vector diferente que porta un marcador de selección de antibiótico diferente. Los usos secuenciales alternativos de dos enzimas de restricción de tipo IIS diferentes y marcadores de selección de antibióticos permiten el ensamblaje jerárquico de grandes fragmentos de ADN mediante una reacción de restricción/ligadura en un solo recipiente.

Un inconveniente importante del ensamblaje de ADN basado en enzimas de tipo IIS es la necesidad de eliminar los sitios de restricción internos de tipo IIS dentro de las partes de ADN que se utilizarán para el ensamblaje. Se requieren un mínimo de dos enzimas para el ensamblaje jerárquico básico y tres enzimas para otros esquemas, como la adición lineal de múltiples pasos de partes de ADN. Debido a que la mayoría de las enzimas de restricción de tipo IIS conocidas reconocen una secuencia ≤ 6 pb, la mayoría de los sistemas basados en Moclo utilizan cortadores de 6 pb. Esto requiere que la parte de ADN esté libre de dos secuencias asimétricas de 6 pb, que ocurren con una frecuencia de ~ 1 en 1 kb para una secuencia de ADN aleatoria con un contenido de GC del 50 %. Una mejora potencial del sistema Moclo actual sería utilizar enzimas de restricción de tipo IIS con 7 pb o más de especificidad. Sin embargo, esta opción está limitada por la disponibilidad de enzimas de restricción comerciales de tipo IIS. Actualmente sólo están disponibles comercialmente dos tipos de enzimas de restricción de tipo IIS de 7 pb: LglI/SapI que reconoce GCTCTTC y deja un

extremo pegajoso de 3 pb, y AarI que reconoce CACCTGC y deja un extremo pegajoso de 4 pb. AarI también requiere la adición de oligonucleótidos para una digestión completa, lo cual no es deseable para el ensamblaje del ADN.

Los sistemas de ensamblaje de ADN basados en enzimas de restricción de tipo IIS existentes también están diseñados para el ensamblaje modular de ADN que deja secuencias cicatriciales no deseadas en el ADN ensamblado final y no se pueden usar para ensamblar secuencias de ADN arbitrarias sin crear vectores de ensamblaje personalizados.

El documento US20160002644 describe un sistema de ensamblaje de ADN. Utiliza dos enzimas de restricción LglI y EarI, y una cepa con expresión inducible de M.TaqI para bloquear sitios de restricción solapados en el vector de ensamblaje. El sitio M.TaqI se solapa con la secuencia del extremo adhesivo de 3 pb generada por LglI/EarI, lo que deja graves limitaciones en el diseño de la secuencia del adaptador.

Lo que se requiere es un método de ensamblaje de ADN mejorado que proporcione a uno o más una mayor libertad para elegir y diseñar secuencias adaptadoras, ensamblaje con cicatrices mínimas, menos pasos de reacción o complejidad y la capacidad de ensamblar secuencias de ADN más grandes. Por lo tanto, un objetivo de la presente invención es proporcionar un método de ensamblaje de ADN mejorado y material asociado.

Según un primer aspecto de la invención, se proporciona un ácido nucleico para su uso en el ensamblaje de ADN según la reivindicación 1 del presente documento.

En el presente documento se describe un ácido nucleico para su uso en el ensamblaje de ADN, en donde el ácido nucleico comprende al menos un elemento de restricción protegible por metilación, comprendiendo el elemento de restricción protegible por metilación:

- una secuencia de reconocimiento de enzima de restricción que es reconocida por una enzima de restricción que escinde fuera de la secuencia de reconocimiento; y
- una secuencia de reconocimiento de metilasa del ADN,

en donde la secuencia de reconocimiento de la enzima de restricción y la secuencia de reconocimiento de la metilasa del ADN se solapan de manera que la base modificada por la metilasa del ADN se encuentra dentro de la secuencia de reconocimiento de la enzima de restricción,

en donde la secuencia de reconocimiento de la metilasa del ADN no es idéntica a, o está incluida por, la secuencia de reconocimiento de la enzima de restricción, y

en donde la secuencia de reconocimiento de la metilasa del ADN no se solapa con la secuencia que formaría la secuencia del extremo sobresaliente generada por la enzima de restricción.

La presente invención puede permitir ventajosamente el uso de una única enzima de restricción en diferentes etapas del ensamblaje del ADN. Esto reduce el coste y la complejidad de preparar una parte inicial del fragmento de ADN para que esté libre de la enzima de restricción utilizada. Como la secuencia de reconocimiento de la metilasa no se solapa con la secuencia que formaría la secuencia del extremo saliente generada por la enzima de restricción, la invención proporciona además más libertad para diseñar adaptadores (es decir, salientes) para el ensamblaje de fragmentos de ADN. La invención también proporciona el primer sistema de ensamblaje de ADN jerárquico sin cicatrices que utiliza un método de ensamblaje basado en enzimas de restricción. A diferencia de los métodos de ensamblaje de ADN modulares existentes, que dejan secuencias cicatriciales no deseadas en el ADN ensamblado final, la presente invención se puede utilizar para ensamblar secuencias de ADN grandes y arbitrarias sin dejar cicatrices. Los sistemas de ensamblaje sin cicatrices existentes, todos los cuales se basan en métodos de ensamblaje basados en homología, tales como el ensamblaje de Gibson y la recombinación homóloga de levadura, no pueden ensamblar construcciones de ADN grandes con secuencias de ADN repetitivas, en contraste con la presente invención. El método de la invención también permite un diseño simple para el ensamblaje jerárquico y sin cicatrices de ADN utilizando un conjunto universal de 6 plásmidos.

La metilación del elemento de control de corte puede bloquear/deteriorar el sitio de reconocimiento de la enzima de restricción solapada.

En una descripción, la enzima de restricción que escinde fuera de su secuencia de reconocimiento es una enzima de restricción de tipo IIS.

En una descripción, el ácido nucleico comprende al menos dos elementos de restricción protegibles por metilación. En otra descripción, el ácido nucleico presenta dos elementos de restricción protegibles por metilación. En una descripción que comprende dos elementos de restricción protegibles por metilación, la secuencia de ácido nucleico entre los sitios de corte de los dos elementos de restricción protegibles por metilación puede ser una secuencia de descarte (es decir, un fragmento no deseado que se eliminará del ácido nucleico durante una reacción de restricción). En una descripción alternativa, la secuencia de descarte puede eliminarse previamente. Por ejemplo, el ácido nucleico puede ser un vector linealizado previamente cortado que tiene un elemento de restricción protegible por metilación en cada extremo.

La secuencia de descarte puede comprender un marcador seleccionable, un gen informador o una etiqueta. El experto estará familiarizado con los marcadores y genes indicadores típicos utilizados en el ensamblaje y la clonación de ADN. Por ejemplo, un marcador seleccionable puede comprender un gen que codifica una resistencia a antibióticos, una actividad enzimática, una luciferasa o similares. El marcador podría ser una etiqueta, tal como una radiomarca. En una descripción, la secuencia de descarte codifica *lacZ* como reportero. El experto reconocerá que la secuencia descartada, que puede ser reemplazada por la secuencia ensamblada en un vector, puede no contener ningún elemento funcional en absoluto, ya que el proceso favorece la generación de ADN ensamblado.

Proporcionar un marcador seleccionable, gen indicador o etiqueta en la secuencia de descarte permite ventajosamente seleccionar clones a los que se les ha eliminado o reemplazado con éxito la secuencia de descarte por un fragmento/secuencia de interés.

En una descripción, se proporciona una secuencia de reconocimiento de enzima de restricción no protegible opuesta en el lado opuesto del sitio de corte del elemento de restricción protegible por metilación. Por ejemplo, en una descripción que comprende una secuencia de descarte, la secuencia de reconocimiento de la enzima de restricción no protegible opuesta se ubicaría en la secuencia de descarte. La secuencia de reconocimiento de la enzima de restricción no protegible opuesta también puede ser reconocida por una enzima de restricción que escinde fuera de su secuencia de reconocimiento, tal como una enzima de restricción de tipo IIS. La secuencia de reconocimiento de la enzima de restricción no protegible opuesta puede no estar protegida por (o dispuesta para ser protegida por) metilación, o al menos puede no estar protegida por metilación de una metilasa que reconoce la secuencia de reconocimiento de metilasa del elemento de restricción protegible por metilación. Tal secuencia de reconocimiento de enzima de restricción no protegible opuesta en la secuencia de descarte puede denominarse de otro modo una "secuencia de reconocimiento de enzima de restricción interna", mientras que la secuencia de reconocimiento de enzima de restricción del elemento de restricción protegible por metilación puede denominarse "secuencia externa de reconocimiento de enzima de restricción". secuencia de reconocimiento de enzimas de restricción".

La secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación puede ser reconocida por la misma especie de enzima de restricción que la secuencia de reconocimiento de la enzima de restricción no protegible opuesta. La secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación puede comprender la misma secuencia que la secuencia de reconocimiento de la enzima de restricción no protegible opuesta.

En una realización, el ácido nucleico puede comprender un "elemento compuesto de mantenimiento", en donde la secuencia de reconocimiento de la enzima de restricción no protegible opuesta puede estar dispuesta para dirigir la enzima de restricción para cortar el ácido nucleico en el mismo sitio que la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación, de modo que se produciría el mismo saliente/extremo pegajoso en la misma posición. Esto se puede lograr proporcionando la secuencia de reconocimiento de la enzima de restricción y la secuencia de reconocimiento de la enzima de restricción opuesta no protegible a una distancia específica, dependiendo del sitio de corte proporcionado por la(s) enzima(s) de restricción seleccionada. En esta realización, la secuencia del saliente producido por la enzima de restricción que corta el ácido nucleico según la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación se mantiene después de cortar el ácido nucleico con la enzima de restricción que reconoce el elemento opuesto no protegible. secuencia de reconocimiento de la enzima de restricción de la secuencia de descarte y una ligadura posterior con un inserto de fragmento de ADN. Como tal, esta realización puede denominarse "elemento compuesto de mantenimiento", es decir, el elemento compuesto de mantenimiento puede comprender dos secuencias de restricción dispuestas en dirección cabeza a cabeza (una de las cuales es parte del elemento protegible contra la metilación, y la otra la secuencia de reconocimiento de la enzima de restricción no protegible opuesta). El saliente producido a partir de un elemento compuesto de mantenimiento puede estar predeterminado por la secuencia entre las secuencias de reconocimiento de restricción opuestas.

En una realización alternativa al elemento compuesto de mantenimiento, se proporciona un "elemento compuesto truncado", en donde la distancia entre el elemento de restricción protegible por metilación y la secuencia de reconocimiento de la enzima de restricción no protegible opuesta se puede reducir de manera que la no protegible opuesta La secuencia de reconocimiento de la enzima de restricción puede disponerse para dirigir la enzima de restricción para que corte el ácido nucleico en un sitio que no se solape al sitio de corte de la enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación (es decir, de manera que la la posición de los salientes posteriores creados por los sitios de corte opuestos no se solapan y los salientes posteriores pueden ser diferentes en secuencia). El sitio de corte de la secuencia de reconocimiento de la enzima de restricción no protegible opuesta en un elemento compuesto truncado puede estar dentro de la secuencia de reconocimiento de la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación, y/o dentro de los nucleótidos entre el sitio de corte y el secuencia de reconocimiento de la enzima de restricción secuencia de reconocimiento del elemento de restricción protegible por metilación. En una realización, el punto de escisión entre pares de bases al final de un sitio de corte y el punto de escisión entre pares de bases al inicio del sitio de corte opuesto puede estar en pares de bases adyacentes/vecinos (es decir, los sitios de corte pueden estar inmediatamente adyacentes a entre sí, pero sin solaparse). Esto se puede lograr proporcionando la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación y la secuencia de reconocimiento de la enzima de restricción no protegible opuesta a una

distancia específica, dependiendo del sitio de corte proporcionado por la enzima(s) de restricción seleccionada. La distancia será lo suficientemente cercana como para causar el corte en la secuencia de reconocimiento de enzimas de restricción opuestas no protegibles y/o nucleótidos entre la secuencia de reconocimiento y el sitio de corte, o viceversa (es decir, más cerca que las secuencias de reconocimiento de enzimas de restricción opuestas en el "elemento compuesto de mantenimiento"). "para un equivalente/la misma enzima de restricción seleccionada). En esta realización de elemento compuesto truncado, la secuencia del saliente producido por la enzima de restricción que corta el ácido nucleico según la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación puede mantenerse o no después de cortar el ácido nucleico con la enzima de restricción. que reconoce la secuencia de reconocimiento de la enzima de restricción no protegible opuesta de la secuencia de descarte y una ligadura posterior con un inserto de fragmento de ADN, mediante el cual la secuencia del inserto de fragmento de ADN dictaría la secuencia del saliente (por ejemplo, el saliente producido a partir de un elemento compuesto truncado puede no estar predeterminado por la secuencia entre los sitios de reconocimiento de enzimas de restricción opuestas).

En una realización alternativa al elemento compuesto de mantenimiento y al elemento compuesto truncado, se proporciona un "elemento compuesto de inserción", en donde se proporciona un inserto de secuencia entre el elemento de restricción protegible por metilación y la secuencia de reconocimiento de enzima de restricción no protegible opuesta. El inserto de secuencia puede comprender una secuencia funcional de manera que proporcione una función. Por ejemplo, el inserto de secuencia puede comprender una secuencia promotora o una secuencia terminadora. En una realización que comprende un elemento compuesto de inserción, el elemento de restricción protegible por metilación y la secuencia de reconocimiento de la enzima de restricción no protegible opuesta pueden no solaparse y pueden estar separados por el inserto de secuencia. El inserto de secuencia puede tener cualquier longitud adecuada. Por ejemplo, la distancia entre el elemento de restricción protegible por metilación y la secuencia de reconocimiento de la enzima de restricción no protegible opuesta puede ser cualquier número de nucleótidos según sea necesario para el inserto de secuencia, tal como un inserto de secuencia funcional. La distancia entre la secuencia de reconocimiento de la enzima de restricción dentro del elemento de restricción protegible por metilación y la secuencia de reconocimiento de la enzima de restricción no protegible opuesta puede oscilar entre 6 pb y 300 kb.

Ventajosamente, después del ensamblaje del ADN, la secuencia funcional se agregará al extremo 5' o 3' del ADN ensamblado. Este diseño es conveniente para agregar elementos comunes a un plásmido ensamblado mediante el uso de vectores especializados (por ejemplo, agregar promotor y terminadores en una reacción de ensamblaje de genes de región codificante de múltiples partes).

En una descripción, para el "elemento compuesto de mantenimiento", la secuencia de reconocimiento de la enzima de restricción no protegible opuesta (tal como una secuencia Bsal) puede distanciarse de la secuencia de reconocimiento de la enzima de restricción (tal como una secuencia Bsal) de la secuencia protegible por metilación. elemento de restricción por 6 pares de bases (es decir, 6 pb en el espacio entre el final de una secuencia de reconocimiento y el inicio de la secuencia de reconocimiento opuesta). El experto reconocerá que dicha distancia depende de la propiedad de la(s) enzima(s) de restricción seleccionada(s). Por ejemplo, suponiendo que la enzima corta x bases de la secuencia de reconocimiento y genera y bases del extremo/protuberancia adhesiva (por ejemplo, para Bsal $x = 1$ e $y = 4$), entonces la distancia (d) para el número de bases entre enzimas de restricción opuestas Las secuencias de reconocimiento del elemento compuesto de mantenimiento pueden proporcionarse mediante la siguiente ecuación: $d = (2 * x) + y$ (por ejemplo, $d = 6$ pb para Bsal). La distancia (d) para el elemento compuesto de mantenimiento puede variar desde 5 pb (por ejemplo, cuando se usa Earl) hasta 24 pb (por ejemplo, cuando se usa BtgZI). En una descripción, la diferencia entre las secuencias de reconocimiento de enzimas de restricción opuestas del elemento compuesto de mantenimiento es d, en donde $d = (2 * x) + y$, y en donde x es el número de pares de bases que tiene el sitio de corte de una enzima de restricción determinada. de su secuencia de reconocimiento, y y es la longitud del saliente que produciría la enzima de restricción.

En la descripción alternativa del elemento compuesto truncado, por ejemplo cuando el saliente producido por la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación puede mantenerse o no, la secuencia de reconocimiento de la enzima de restricción no protegible opuesta (tal como para Bsal) puede estar distanciados de la secuencia de reconocimiento de la enzima de restricción (tal como para Bsal) del elemento de restricción protegible por metilación en 2 pares de bases. Como anteriormente, el experto reconocerá que dicha distancia depende de la propiedad de la(s) enzima(s) de restricción seleccionada(s). Por ejemplo, suponiendo que la enzima corta x bases de la secuencia de reconocimiento y genera y bases con el extremo adhesivo (para Bsal $x = 1$ e $y = 4$), la distancia (d) para el número de bases entre secuencias de reconocimiento de enzimas de restricción opuestas de la secuencia truncada El elemento compuesto puede ser proporcionado por la siguiente ecuación: $d = 2 * x$ (por ejemplo, $d = 2$ pb para Bsal). La distancia (d) para el elemento compuesto truncado puede oscilar entre 2 pb (por ejemplo, cuando se usa Bsal) y 20 pb (por ejemplo, cuando se usa BtgZI). En una descripción, la diferencia entre las secuencias de reconocimiento de enzimas de restricción opuestas del elemento compuesto truncado es d, en donde $d = 2 * x$, y en donde x es el número de pares de bases que el sitio de corte de una enzima de restricción dada es de su secuencia de reconocimiento.

Ventajosamente, la provisión de un elemento compuesto de mantenimiento facilita el ensamblaje de ADN en un solo recipiente, mediante el cual se pueden ensamblar múltiples fragmentos de ADN en una única reacción en un solo recipiente con una única especie de enzima de restricción. Ventajosamente, la provisión de un elemento compuesto

truncado facilita el ensamblaje de ADN sin cicatrices, por lo que los extremos sobresalientes de un inserto de fragmento de ADN particular pueden determinarse para adaptarse para la ligación en un vector de destino o proporcionarse según la secuencia nativa del fragmento de ADN. de interés para unirse con fragmentos vecinos en la secuencia.

En una descripción, el ácido nucleico puede comprender un ácido nucleico circular, tal como un vector, que comprende dos elementos compuestos de mantenimiento con una secuencia de descarte entre ellos, o una versión linealizada del mismo con la secuencia de descarte recortada. En otra descripción, el ácido nucleico puede comprender un ácido nucleico circular, tal como un vector, que comprende un elemento compuesto de mantenimiento y un elemento compuesto truncado con una secuencia de descarte entre ellos, o una versión linealizada del mismo con la secuencia de descarte recortada. En otra descripción, el ácido nucleico puede comprender un ácido nucleico circular, tal como un vector, que comprende dos elementos compuestos truncados con una secuencia de descarte entre ellos, o una versión linealizada del mismo con la secuencia de descarte recortada. En otra descripción, el ácido nucleico puede comprender un ácido nucleico circular, tal como un vector, que comprende un elemento compuesto de inserción y un elemento compuesto de inserción y un elemento compuesto de mantenimiento con una secuencia de descarte entre ellos, o una versión linealizada del mismo con la secuencia de descarte recortada. En otra descripción, el ácido nucleico puede comprender un ácido nucleico circular, tal como un vector, que comprende dos elementos compuestos de inserción con una secuencia de descarte entre ellos, o una versión linealizada del mismo con la secuencia de descarte recortada. En realizaciones de ácido nucleico linealizado, la secuencia de descarte puede haberse cortado mediante restricción con la enzima de restricción que reconoce la secuencia de enzima de restricción opuesta no protegible presente en la secuencia de descarte (es decir, la secuencia de reconocimiento de la enzima de restricción interna).

En descripciones que comprenden dos elementos compuestos de mantenimiento, las secuencias de reconocimiento de enzimas de restricción de los dos elementos compuestos de mantenimiento pueden ser las mismas y/o las secuencias de reconocimiento de metilasa de los dos elementos compuestos de mantenimiento pueden ser las mismas. En descripciones que comprenden dos elementos compuestos truncados, las secuencias de reconocimiento de enzimas de restricción de los dos elementos compuestos truncados pueden ser las mismas y/o las secuencias de reconocimiento de metilasa de los dos elementos compuestos truncados pueden ser las mismas. En descripciones que comprenden dos elementos compuestos de inserción, las secuencias de reconocimiento de enzimas de restricción de los dos elementos compuestos de inserción pueden ser las mismas y/o las secuencias de reconocimiento de metilasa de los dos elementos compuestos de inserción pueden ser las mismas. En las descripciones que comprenden un elemento compuesto de mantenimiento y un elemento compuesto de truncamiento, las secuencias de reconocimiento de enzimas de restricción de los elementos compuestos de mantenimiento y truncamiento pueden ser las mismas y/o las secuencias de reconocimiento de metilasa de los elementos compuestos de mantenimiento y truncamiento pueden ser las mismas. En las descripciones que comprenden un elemento compuesto de mantenimiento y un elemento compuesto de inserción, las secuencias de reconocimiento de enzimas de restricción de los elementos compuestos de mantenimiento y de inserción pueden ser las mismas y/o las secuencias de reconocimiento de metilasa de los elementos compuestos de mantenimiento y de inserción pueden ser las mismas. En las descripciones que comprenden un elemento compuesto truncado y un elemento compuesto de inserción, las secuencias de reconocimiento de enzimas de restricción de los elementos compuestos truncados e de inserción pueden ser las mismas y/o las secuencias de reconocimiento de metilasa de los elementos compuestos truncados e de inserción pueden ser las mismas. Por ejemplo, la digestión del ácido nucleico para eliminar una secuencia de descarte o un fragmento de ADN de interés puede llevarse a cabo mediante la misma enzima de restricción. Además, la metilación del ácido nucleico para proteger todos los sitios de reconocimiento de enzimas de restricción protegibles por metilasa en el ácido nucleico puede llevarse a cabo mediante la misma metilasa.

En una descripción, la metilasa comprende una metilasa del ADN de tipo II. En otra descripción, la metilasa comprende una metilasa del ADN de tipo I. En una descripción, la metilasa comprende una metilasa de dominio único sin actividad de enzima de restricción, o una metilasa modificada que tiene una actividad de enzima de restricción no funcional. Por ejemplo, la modificación puede comprender modificar la posición N6 de metiladenina. En una descripción, la secuencia de reconocimiento de la metilasa del ADN comprende al menos 4 pares de bases. La metilasa del ADN puede estar activa, tal como óptimamente activa, a aproximadamente 37°C.

En una descripción, la secuencia de reconocimiento de la enzima de restricción no protegible opuesta no es capaz de bloquearse mediante metilación con la metilasa que reconoce la secuencia de reconocimiento de la metilasa del ADN del elemento de restricción protegible por metilación. En una descripción, la metilasa que reconoce la secuencia de reconocimiento de metilasa del ADN del elemento de restricción protegible por metilación puede no reconocer una secuencia dentro o solapada con la secuencia de reconocimiento de la enzima de restricción no protegible opuesta. Por ejemplo, la secuencia de reconocimiento de la enzima de restricción opuesta no protegible puede no solaparse con o comprender una secuencia que sea reconocida por la metilasa que reconoce la secuencia de reconocimiento de la metilasa del ADN del elemento de restricción protegible por metilación. La secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación puede estar metilada, por ejemplo mediante la metilasa que reconoce la secuencia de reconocimiento de metilasa del ADN del elemento de restricción protegible por metilación. La metilasa que reconoce la secuencia de reconocimiento de la metilasa del ADN del elemento de restricción protegible por metilación puede disponerse para metilar un nucleótido dentro de la secuencia de

reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación. La metilación puede ser capaz de bloquear el corte del ADN mediante la enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación. En una descripción, la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación puede volverse no funcional como resultado de la metilación de al menos uno de los nucleótidos de la secuencia. En una descripción, el nucleótido metilado puede ser una adenina, o el nucleótido dispuesto para ser metilado puede ser una adenina. En otra descripción, el nucleótido metilado puede ser una citosina, o el nucleótido dispuesto para ser metilado puede ser una citosina. La metilación puede ser cualquiera de los tipos de metilación seleccionados entre N4-metilcitosina (m4C), C5-metilcitosina (m5C) o N6-metiladenina (m6A). El experto reconocerá que el tipo de metilación proporcionada debería bloquear/deteriorar la función del sitio de restricción solapado. Además, para metilación *in vivo*, la cepa utilizada para expresar la metilasa puede ser deficiente en enzimas de restricción dependientes de modificación que pueden reconocer las bases metiladas, como las enzimas de restricción de la familia Mcr/Mrr. En una descripción, el tipo de metilación puede comprender N6-metiladenina (m6A).

En una descripción, la metilasa puede metilar o no las bases más externas de la secuencia de reconocimiento de la enzima de restricción. Por ejemplo, en una secuencia de reconocimiento de enzimas de restricción de 6 pb, las bases 1 ó 6 pueden no estar metiladas. En una descripción, la metilasa puede metilar o no los 2^o base de la secuencia de reconocimiento de la enzima de restricción. En una descripción, la metilasa puede metilar la posición 3, 4 o 5 para sitios de reconocimiento de enzimas de restricción de 6 pb. En una descripción, la metilasa puede metilar la posición 3, 4, 5 o 6 para los sitios de reconocimiento de enzimas de restricción.

En una descripción que comprende dos elementos de restricción protegibles por metilación, las secuencias de reconocimiento de metilasa del ADN de cada elemento de restricción protegible por metilación pueden ser reconocidas por la misma metilasa. La secuencia de reconocimiento de metilasa del ADN de los elementos de restricción protegibles por metilación puede comprender o consistir en la misma secuencia.

En una descripción que comprende dos elementos de restricción protegibles por metilación, las secuencias de reconocimiento de enzimas de restricción de cada elemento de restricción protegible por metilación pueden ser reconocidas por la misma especie de enzima de restricción. Cada secuencia de reconocimiento de enzimas de restricción de los elementos de restricción protegibles por metilación puede comprender o consistir en la misma secuencia. Los dos elementos de restricción protegibles por metilación pueden comprender o consistir en la misma secuencia.

El uso de la misma secuencia de reconocimiento de enzimas de restricción proporciona ventajosamente que se puede usar una única especie de enzima de restricción en una reacción de ensamblaje de ADN. El uso de una única especie de enzima de restricción abre la posibilidad de un ensamblaje de ADN menos complejo, porque la presencia de más especies de enzimas de restricción en una reacción aumenta significativamente la probabilidad de que la secuencia contenga aleatoriamente/inadvertidamente un sitio de reconocimiento y corte de restricción, causando un fallo en el procedimiento de clonación.

En una descripción, el elemento de restricción protegible por metilación comprende o consiste en una secuencia según cualquiera de los sitios de metilación/restricción solapados identificados en la Tabla 3 del presente documento. En una descripción alternativa, el elemento de restricción protegible por metilación comprende o consiste en una secuencia según cualquiera de los sitios de metilación/restricción solapados identificados en la Tabla 3 en el presente documento, excluyendo aquellos que proporcionan metilación de la primera o última base en la secuencia de restricción. En una descripción, el elemento de restricción protegible por metilación comprende o consiste en la secuencia GACNNGGTCTCNNNNN (BsaI/M.Osp807II - SEQ ID NO: 1). En otra descripción, el elemento de restricción protegible por metilación comprende o consiste en la secuencia GAAGACGCNNNNNN (BpiI/M2.NmeMC58II - SEQ ID NO: 2). En otra descripción, el elemento de restricción protegible por metilación comprende o consiste en la secuencia GAAGCTCTTCNNNN (LglI/M.Xmnl - SEQ ID NO: 3).

En una descripción, la secuencia de reconocimiento de enzima de restricción protegible y/o no protegible por metilación comprende o consiste en una secuencia según cualquiera de las secuencias de reconocimiento de enzima de restricción identificadas en la Tabla 3 del presente documento. En una descripción alternativa, la secuencia de reconocimiento de la enzima de restricción protegible y/o no protegible por metilación comprende o consiste en una secuencia según cualquiera de las secuencias de reconocimiento de la enzima de restricción identificadas en la Tabla 3 del presente documento, excluyendo aquellas que proporcionan la metilación de la primera o la última base en la secuencia de restricción. En una descripción, la secuencia de reconocimiento de la enzima de restricción protegible y/o no protegible por metilación comprende o consiste en la secuencia GGTCTC (BsaI - SEQ ID NO: 4). En otra descripción, la secuencia de reconocimiento de la enzima de restricción protegible y/o no protegible por metilación comprende o consiste en una secuencia GAAGAC (BpiI - SEQ ID NO: 5). En otra descripción, la secuencia de reconocimiento de la enzima de restricción protegible y/o no protegible por metilación comprende o consiste en una secuencia GCTCTTC (LglI - SEQ ID NO: 6).

En una descripción, la secuencia de reconocimiento de metilasa del ADN comprende o consiste en una secuencia según cualquiera de las secuencias de reconocimiento de metilasa del ADN de tipo II identificadas en la Tabla 3 del presente documento. En una descripción alternativa, la secuencia de reconocimiento de metilasa del ADN comprende

o consiste en una secuencia según cualquiera de las secuencias de reconocimiento de metilasa del ADN de tipo II identificadas en la Tabla 3 en el presente documento, excluyendo aquellas que proporcionan metilación de la primera o última base en la secuencia de restricción. En una descripción, la secuencia de reconocimiento de metilasa del ADN comprende o consiste en la secuencia GACNNNGTC (M.Osp807II - SEQ ID NO: 7). En otra descripción, la secuencia de reconocimiento de metilasa del ADN comprende o consiste en la secuencia GACGC (M2.NmeMC58II - SEQ ID NO: 8). En otra descripción, la secuencia de reconocimiento de metilasa del ADN comprende o consiste en la secuencia GAANNNTTC (M.XmnI - SEQ ID NO: 9).

La enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación, y/o la secuencia de reconocimiento de la enzima de restricción opuesta, puede ser capaz de cortar el ácido nucleico para dejar al menos un saliente/extremo pegajoso de 2 pb. Alternativamente, la enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación, y/o la secuencia de reconocimiento de la enzima de restricción opuesta, puede ser capaz de cortar el ácido nucleico para dejar al menos un saliente/extremo pegajoso de 3 pb. En otra descripción, la enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación, y/o la secuencia de reconocimiento de la enzima de restricción opuesta, puede ser capaz de cortar el ácido nucleico para dejar un saliente/extremo pegajoso de 3 pb. En otra descripción, la enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación, y/o la secuencia de reconocimiento de la enzima de restricción opuesta, puede ser capaz de cortar el ácido nucleico para dejar un saliente/extremo pegajoso de 4 pb. En otra descripción, la enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación, y/o la secuencia de reconocimiento de la enzima de restricción opuesta, puede ser capaz de cortar el ácido nucleico para dejar un saliente/extremo pegajoso de 2 a 6 pb. En otra descripción, la enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación, y/o la secuencia de reconocimiento de la enzima de restricción opuesta, puede ser capaz de cortar el ácido nucleico para dejar un saliente/extremo pegajoso de 2 a 5 pb. En otra descripción, la enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación, y/o la secuencia de reconocimiento de la enzima de restricción opuesta, puede ser capaz de cortar el ácido nucleico para dejar un saliente/extremo pegajoso de 3 a 5 pb. En otra descripción, la enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible por metilación, y/o la secuencia de reconocimiento de la enzima de restricción opuesta, puede ser capaz de cortar el ácido nucleico para dejar un saliente/extremo pegajoso de 4 a 5 pb.

En una descripción, la enzima de restricción comprende o consiste en cualquier enzima de restricción de tipo IIS identificada en la Tabla 3 del presente documento. En otra descripción, la enzima de restricción de tipo IIS comprende o consiste en BsaI, BpiI o LglI.

En una descripción, la metilasa del ADN comprende o consiste en cualquier metilasa del ADN de tipo II identificada en la Tabla 3 del presente documento. En una descripción alternativa, la metilasa del ADN comprende o consiste en cualquier metilasa del ADN de tipo II identificada en la Tabla 3 del presente documento, excluyendo aquellas que proporcionan metilación de la primera o la última base en la secuencia de restricción. En otra descripción, la metilasa del ADN comprende o consiste en M.Osp807II, M2.NmeMC58II o M.XmnI.

En una descripción, la enzima de restricción y la metilasa del ADN comprenden o consisten en cualquier par de enzimas de restricción de tipo IIS y las metilasas del ADN de tipo II identificadas en la Tabla 3 del presente documento. En una descripción, la enzima de restricción y la metilasa del ADN comprenden o consisten en los pares BsaI/M.Osp807II, BpiI/M2.NmeMC58II o LglI/M.XmnI.

En una descripción, la secuencia de reconocimiento de la enzima de restricción y la secuencia de reconocimiento de la metilasa del ADN comprenden o consisten en cualquier par de secuencias de reconocimiento de la enzima de restricción de tipo IIS y las secuencias de reconocimiento de la metilasa del ADN de tipo II identificadas en la Tabla 3 del presente documento. En una descripción alternativa, la secuencia de reconocimiento de la enzima de restricción y la secuencia de reconocimiento de la metilasa del ADN comprenden o consisten en cualquier par de secuencias de reconocimiento de la enzima de restricción de tipo IIS y las secuencias de reconocimiento de la metilasa del ADN de tipo II identificadas en la Tabla 3 del presente documento, excluyendo aquellas que proporcionan metilación de la primera o última base en la secuencia de restricción.

En una descripción que comprende dos elementos de restricción protegibles por metilación (por ejemplo, con una secuencia de descarte entre ellos), el saliente/extremo pegajoso producido cortando con la enzima de restricción en el primer elemento de restricción protegible por metilación puede tener una secuencia diferente que el saliente/extremo pegajoso. extremo producido cortando con la enzima de restricción en el segundo elemento de restricción protegible por metilación. El experto estará familiarizado con proporcionar diferentes salientes para el ensamblaje/clonación direccional de ADN o controlar qué extremo del ácido nucleico está ligado a un inserto/fragmento de interés.

En una descripción que implica un elemento compuesto de mantenimiento, el ácido nucleico puede comprender o consistir en la secuencia de GACNNGGTCTCNNNNNNGAGACC (SEQ ID NO: 10). En otra descripción que implica un elemento compuesto de mantenimiento, el ácido nucleico puede comprender o consistir en la secuencia de

GAAGACGCNNNNNNGTCTTC (SEQ ID NO: 11). En otra descripción que implica un elemento compuesto de mantenimiento, el ácido nucleico puede comprender o consistir en la secuencia de GAAGCTCTTCNNNNNGAAGAGC (SEQ ID NO: 12). N puede significar A, C, T o G. Se pueden proporcionar otras secuencias según las secuencias de reconocimiento de metilasa y enzima de restricción combinadas/solapadas proporcionadas en la tabla 3, en donde después de la serie de N bases, la secuencia comprende además las secuencias palindrómicas/inversas. secuencia complementaria de la secuencia de reconocimiento de la enzima de restricción.

En una descripción que implica dos elementos compuestos de mantenimiento, el ácido nucleico puede comprender o consistir en la secuencia de GACNNGGTCTCNNNNNGAGACC(N^x)GGTCTCNNNNNGAGACCNNGTC (SEQ ID NO: 13). En otra descripción que implica dos elementos compuestos de mantenimiento, el ácido nucleico puede comprender o consistir en la secuencia de GAAGACGCNNNNNNGTCTTC(N^x)GAAGACNNNNNNGCGTCTTC (SEQ ID NO: 14). En otra descripción que implica dos elementos compuestos de mantenimiento, el ácido nucleico puede comprender o consistir en la secuencia de GAAGCTCTTCNNNNNGAAGAGC(N^x)GCTCTTCNNNNNGAAGAGCTTC (SEQ ID NO: 15). N puede significar A, C, T o G. (N^x) puede significar cualquier número de A, C, T o G, o entre 0 pb y 300 kbp de A, C, T o G.

El experto entenderá que para algunas secuencias de reconocimiento de enzimas de restricción, tales como para LguI, el número de nucleótidos (N^x) entre dos elementos compuestos de mantenimiento podría ser negativo con los dos sitios internos de reconocimiento de enzimas de restricción no protegibles solapados. Por lo tanto, en otra descripción que implica dos elementos compuestos de mantenimiento, el ácido nucleico puede comprender o consistir en la secuencia de GAAGCTCTTCNNNNNGAAGA.**GCT**CTTCNNNNNGAAGAGCTTC (SEQ ID NO: 16) (los nucleótidos solapados se muestran en negrita y subrayados).

En una descripción que implica un elemento compuesto truncado, el ácido nucleico puede comprender o consistir en la secuencia de GACNNGGTCTCNGAGACC (SEQ ID NO: 17). En una descripción que implica un elemento compuesto truncado, el ácido nucleico puede comprender o consistir en la secuencia de GAAGCTCTTCNNGAAGAGGC (SEQ ID NO: 18). N puede significar A, C, T o G. N puede significar A, C, T o G. Se pueden proporcionar otras secuencias según las secuencias de reconocimiento de enzima de restricción y metilasa combinadas/solapadas proporcionadas en la tabla 3, en donde siguiendo la serie de N bases, la secuencia comprende además la secuencia palindrómica/complementaria inversa de la secuencia de reconocimiento de la enzima de restricción, y en donde el número de N bases entre las secuencias de reconocimiento de restricción se reduce por el número de bases sobresalientes producidas por la enzima de restricción seleccionada.

En una descripción que implica un elemento compuesto de mantenimiento y truncamiento, el ácido nucleico puede comprender o consistir en la secuencia de GACNNGGTCTCNNNNNGAGACC(N^x)GGTCTCNGAGACCNNGTC (SEQ ID NO: 19). En una descripción que implica un elemento compuesto de mantenimiento y truncamiento, el ácido nucleico puede comprender o consistir en la secuencia de GAAGCTCTTCNNNNNGAAGAGC(N^x)GCTCTTCNNGAAGAGCTTC (SEQ ID NO: 20). N puede significar A, C, T o G. (N^x) puede significar cualquier número de A, C, T o G, o entre 0 pb y 300 kbp de A, C, T o G.

El experto entenderá que, para algunas secuencias de reconocimiento de enzimas de restricción, tales como para LguI, el número de nucleótidos (N^x) entre un elemento compuesto de mantenimiento y truncamiento podría ser negativo con los dos sitios internos de reconocimiento de enzimas de restricción no protegibles solapados. Por lo tanto, en otra descripción que implica un elemento compuesto de mantenimiento y truncamiento, el ácido nucleico puede comprender o consistir en la secuencia de GAAGCTCTTCNNNNNGAAGA.**GCT**CTTCNNGAAGAGCTTC (SEQ ID NO: 21) (los nucleótidos solapados se muestran en negrita y subrayados).

En una descripción que implica un elemento compuesto de mantenimiento y truncamiento, el ácido nucleico puede comprender o consistir en la secuencia de GACNNGGTCTCNGAGACC(N^x)GGTCTCNNNNNGAGACCNNGTC (SEQ ID NO: 22). En una descripción que implica un elemento compuesto de mantenimiento y truncamiento, el ácido nucleico puede comprender o consistir en la secuencia de GAAGCTCTTCNNGAAGAGC(N^x)GCTCTTCNNNNNGAAGAGCTTC (SEQ ID NO: 23). N puede significar A, C, T o G. (N^x) puede significar cualquier número de A, C, T o G, o entre 0 pb y 300 kbp de A, C, T o G. En otra descripción que implica un elemento compuesto que mantiene y trunca, el ácido nucleico puede comprender o consistir en la secuencia de GAAGCTCTTCNNGAAGAG**GCT**CTTCNNNNNGAAGAGCTTC (SEQ ID NO: 24) (los nucleótidos solapados se muestran en negrita y subrayados).

En una descripción que implica dos elementos compuestos truncados, el ácido nucleico puede comprender o consistir en la secuencia de GACNNGGTCTCNGAGACC(N^x)GGTCTCNGAGACCNNGTC (SEQ ID NO: 25). En una descripción que implica dos elementos compuestos truncados, el ácido nucleico puede comprender o consistir en la secuencia de GAAGCTCTTCNNGAAGAGC(N^x)GCTCTTCNNGAAGAGCTTC (SEQ ID NO: 26). N puede significar A, C, T o G. (N^x) puede significar cualquier número de A, C, T o G, o entre 0 pb y 300 kbp de A, C, T o G.

El experto entenderá que, para algunas secuencias de reconocimiento de enzimas de restricción, tales como para LguI, el número de nucleótidos (N^x) entre dos elementos compuestos truncados podría ser negativo con los dos sitios internos de reconocimiento de enzimas de restricción no protegibles solapados. Por lo tanto, en otra descripción que involucra dos elementos compuestos truncados, el ácido nucleico puede comprender o consistir en la secuencia de

GAAGCTCTTCNNGAAGA.GCTCTTCNNGAAGAGCTTC (SEQ ID NO: 27) (los nucleótidos solapados se muestran en negrita y subrayados).

El ácido nucleico puede ser un ácido nucleico aislado. En una descripción, el ácido nucleico puede ser sintético (es decir, no encontrarse en la naturaleza). En una descripción, el ácido nucleico puede aislarse o derivarse de una cepa bacteriana, tal como *E. coli*, que expresa una metilasa del ADN que reconoce (y metila) la secuencia de reconocimiento de metilasa del elemento de restricción protegible por metilación. Por ejemplo, una cepa bacteriana, como *E. coli*, que expresa cualquier metilasa del ADN de tipo II identificada en la Tabla 3 del presente documento. En una descripción, el ácido nucleico puede aislarse o derivarse de una cepa bacteriana, tal como *E. coli*, que expresa M.Osp807II, M2.NmeMC58II o M.Xmnl. La metilasa del ADN expresada puede ser funcional. La cepa bacteriana puede modificarse/diseñarse genéticamente para expresar dicha metilasa del ADN.

El ácido nucleico puede ser un vector, tal como un vector de clonación. El experto estará familiarizado con diversos vectores de clonación, que pueden incluir, por ejemplo, un origen de replicación, un marcador seleccionable, un gen indicador, elementos de expresión o combinaciones de los mismos. El vector puede comprender ácido nucleico que comprende un elemento de ADN que soporta la replicación del ADN que lo contiene (por ejemplo, un origen de replicación) y un marcador de selección (por ejemplo, un marcador de selección de antibiótico). El vector puede ser un vector con un número de copias alto o bajo. En una descripción, el vector es un vector con un número elevado de copias. El experto estará familiarizado con los métodos de clonación y los materiales/vectores, por ejemplo como se proporciona en Green, M.R., Sambrook, J. Clonación molecular: manual de laboratorio. 4to. Prensa de laboratorio de Cold Spring Harbor, Cold Spring Harbor, Nueva York; 2012.

En versiones linealizadas del ácido nucleico, el ácido nucleico no puede comprender más de dos copias/sitios de la secuencia de reconocimiento de la enzima de restricción.

También se describe un método de ensamblaje de ADN que comprende:

proporcionar un ácido nucleico metilado linealizado como se describe en el presente documento, en donde el elemento de restricción protegible por metilación está metilado con la metilasa del ADN;

proporcionar dos o más fragmentos de ADN de interés para el ensamblaje con el ácido nucleico metilado linealizado, en donde un primer fragmento de ADN comprende un saliente complementario para la ligación con un primer extremo del ácido nucleico metilado linealizado, y un segundo fragmento de ADN comprende salientes complementarios para la ligación con el otro/segundo extremo del ácido nucleico metilado linealizado; y

(i) el primer y segundo fragmentos de ADN comprenden además salientes complementarios para ligarse entre sí; o

(ii) el primer y segundo fragmentos de ADN comprenden además salientes complementarias para la ligación con uno o más fragmentos de ADN adicionales que tienen salientes complementarias, de modo que ligar los fragmentos de ADN y el ácido nucleico metilado linealizado con una ADN ligasa daría como resultado una única molécula de ADN ensamblada;

ligar los fragmentos de ADN y el ácido nucleico metilado linealizado con una ligasa para formar una única molécula de ADN ensamblada que comprende la secuencia de los fragmentos de ADN ensamblados flanqueados por las secuencias de reconocimiento de enzimas de restricción de los elementos de restricción protegibles por metilación.

La única molécula de ADN ensamblada puede ser circular, tal como un vector de ADN circular.

Ventajosamente, el corte adicional de la única molécula de ADN ensamblada mediante la enzima de restricción que reconoce las secuencias de reconocimiento de las enzimas de restricción de los elementos de restricción protegibles por metilación se bloquea mediante la metilación de las secuencias de reconocimiento de las enzimas de restricción.

El ácido nucleico linealizado puede proporcionarse proporcionando un ácido nucleico como se describe en el presente documento en forma de un vector de destino circular que comprende dos elementos de restricción metilados protegibles por metilación y una secuencia de descarte entre ellos, en donde el método comprende el paso de cortar el vector de destino circular con enzimas de restricción que reconocen las secuencias de reconocimiento de enzimas de restricción no protegibles opuestas en la secuencia de descarte, dejando así un ácido nucleico linealizado que tiene salientes definidos por las enzimas de restricción. Las enzimas de restricción pueden ser de la misma especie. Las secuencias de reconocimiento de enzimas de restricción de los elementos de restricción metilados protegibles por metilación pueden comprender la misma secuencia. En una descripción, las secuencias de reconocimiento de enzimas de restricción de los elementos de restricción metilados protegibles por metilación y las secuencias de reconocimiento de enzimas de restricción no protegibles opuestas pueden comprender la misma secuencia.

La enzima de restricción y la metilasa (o secuencias de reconocimiento de las mismas) utilizadas en el método se pueden seleccionar según la enzima de restricción y las metilasas (o secuencias de reconocimiento de las mismas) proporcionadas para la descripción del ácido nucleico en el presente documento.

El vector de destino circular puede purificarse/aislarse a partir de una cepa bacteriana que exprese la metilasa de ADN que reconoce la secuencia de reconocimiento de la metilasa de ADN del elemento de restricción protegible contra la metilación. En otra descripción, el vector de destino circular o linealizado puede ser metilado por la metilasa de ADN *in vitro*.

5 Los dos o más fragmentos de ADN de interés para ensamblar con el ácido nucleico pueden proporcionarse en vectores de donación circulares, en donde el método comprende el paso de cortar los vectores de donación circulares para liberar los fragmentos de ADN de interés. El corte de los vectores donantes puede realizarse con una enzima de restricción que deja salientes complementarios para la ligación con el ácido nucleico metilado linealizado. El corte de los vectores donantes puede realizarse con una enzima de restricción que deja salientes complementarios para la ligación con el ácido nucleico metilado linealizado y con otro fragmento de ADN de interés para la inserción. El corte de los vectores donantes puede realizarse con una enzima de restricción que deja salientes complementarios para la ligación con otros dos fragmentos de ADN de interés para la inserción.

10 La enzima de restricción para cortar los vectores de donación circulares puede comprender una enzima de restricción de tipo IIS. La enzima de restricción para cortar los vectores de donación circulares puede comprender la misma enzima de restricción, o una enzima de restricción que reconoce la misma secuencia, que la enzima de restricción utilizada para cortar el vector de destino circular. Los salientes complementarios de los fragmentos de ADN de interés para la ligación con el ácido nucleico linealizado pueden definirse mediante la enzima de restricción.

15 Ventajosamente, la provisión de la misma secuencia de reconocimiento de la enzima de restricción para los elementos de restricción protegibles por metilación, la secuencia de reconocimiento de la enzima de restricción opuesta en la secuencia de descarte y las secuencias de reconocimiento de la enzima de restricción en los vectores donantes proporciona que se pueda utilizar un único tipo de enzima de restricción. usado. Esto permite reducir el coste y la complejidad a la hora de eliminar el sitio de restricción interno de tipo IIS de las partes de ADN, porque hay menos sitios de restricción que eliminar. Esto hace que el ensamblaje de secuencias más largas sea más fácil y menos costoso. Además, la invención permite la denominada reacción "en un solo recipiente", mediante la cual la restricción y la ligación se pueden llevar a cabo en un recipiente en un solo paso.

20 Por lo tanto, el método puede combinar los pasos de restringir y ligar. En particular, el vector de destino circular, los vectores donantes, la enzima de restricción y la ligasa pueden proporcionarse en la misma composición.

Después del ensamblaje de la única molécula de ADN ensamblada, el método puede comprender además el paso de transformar la única molécula de ADN ensamblada en una cepa bacteriana que no expresa la metilasa del ADN que reconoce la secuencia de reconocimiento de la metilasa del ADN del elemento de restricción protegible por metilación.

25 Ventajosamente, transformar la única molécula de ADN ensamblada en una cepa bacteriana que no expresa la metilasa del ADN que reconoce la secuencia de reconocimiento de la metilasa del ADN del elemento de restricción protegible por metilación da como resultado la clonación y producción de una secuencia de ácido nucleico, tal como un vector. que comprende una secuencia de reconocimiento de enzima de restricción no metilada (es decir, no bloqueada) para permitir reacciones de restricción posteriores.

30 El método puede comprender además un proceso de ensamblaje de múltiples etapas para formar fragmentos de ADN ensamblados o vectores de mayor longitud de secuencia. Por ejemplo, el ADN ensamblado, que es el producto del método, puede usarse posteriormente para un ensamblaje de orden superior. Por ejemplo, el método puede comprender una segunda ronda de ensamblaje de ADN según la descripción en el presente documento en donde los fragmentos de ADN de interés para el ensamblaje con el ácido nucleico metilado linealizado pueden comprender ADN ensamblado de una primera/anterior ronda de ensamblaje de ADN según la descripción en el presente documento.

35 Proporcionar múltiples rondas de ensamblaje de ADN para formar fragmentos de ADN más grandes reduce ventajosamente el número de secuencias adaptadoras/protuberantes que son necesarias. La tasa de éxito del ensamblaje del ADN también aumenta al ensamblar menos fragmentos de ADN. Además, puede ser más fácil detectar el ADN ensamblado exitosamente con menos fragmentos ensamblados en una sola reacción.

Método de ensamblaje de ADN sin cicatrices

Según otro aspecto de la presente invención, se proporciona un método de ensamblaje de fragmentos de ADN sin cicatriz según la reivindicación 7 del presente documento.

50 Descrito en el presente documento, se proporciona un método de ensamblaje de fragmentos de ADN sin cicatriz comprende los pasos de:

(A) proporcionar un primer ácido nucleico metilado linealizado como se describe en el presente documento, que tiene un elemento compuesto de mantenimiento en un extremo y un elemento compuesto de truncamiento en el otro extremo, en donde los elementos compuestos de mantenimiento y truncamiento están metilados con la metilasa del ADN;

proporcionar un primer fragmento de ADN para ensamblaje que tiene extremos salientes que están adaptados para ligarse a los extremos salientes del primer ácido nucleico metilado linealizado;

ligar el primer fragmento de ADN para el ensamblaje y el primer ácido nucleico metilado linealizado con una ligasa para formar un primer vector intermedio metilado;

- 5 transformar el primer vector intermedio metilado en una cepa bacteriana que no expresa la(s) metilasa del ADN(s) que reconoce cualquiera de las secuencias de reconocimiento de metilasa del ADN de los elementos compuestos de mantenimiento y truncamiento, formando de este modo un primer vector intermedio que no está metilado;

aislar el primer vector intermedio;

- 10 (B) proporcionar un segundo ácido nucleico metilado linealizado como se describe en el presente documento, que tiene un elemento compuesto de mantenimiento en un extremo y un elemento compuesto truncado en el otro extremo, en donde los elementos compuestos de mantenimiento y truncamiento están metilados con la metilasa del ADN, y en donde el saliente proporcionado por el elemento compuesto de mantenimiento del primer ácido nucleico metilado linealizado es diferente en secuencia al saliente proporcionado por el elemento de restricción protegible por metilación del segundo ácido nucleico metilado linealizado;

proporcionar un segundo fragmento de ADN para ensamblaje que tiene extremos salientes que están adaptados para ligarse a los extremos salientes del segundo ácido nucleico metilado linealizado;

ligar el segundo fragmento de ADN para el ensamblaje y el segundo ácido nucleico metilado linealizado con una ligasa para formar un segundo vector intermedio metilado;

- 20 transformar el segundo vector intermedio metilado en una cepa bacteriana que no expresa la(s) metilasa del ADN(s) que reconoce(n) cualquiera de las secuencia(s) de reconocimiento de metilasa del ADN de los elementos compuestos de mantenimiento y truncamiento, formando de ese modo un segundo vector intermedio que no está metilado;

aislar el segundo vector intermedio;

- 25 (C) cortar el primer vector intermedio con una enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento compuesto de mantenimiento y una enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento compuesto truncado, formando de este modo un primer inserto de fragmento de ADN adaptado que comprende una secuencia saliente mantenida que está determinada por el elemento compuesto de mantenimiento y una secuencia saliente nativa opuesta que está determinada por la secuencia nativa del primer fragmento de ADN para ensamblaje;

- 30 (D) cortar el segundo vector intermedio con una enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento compuesto de mantenimiento y una enzima de restricción que reconoce la secuencia de reconocimiento de la enzima de restricción del elemento compuesto truncado, formando de este modo un segundo inserto de fragmento de ADN adaptado que comprende una secuencia saliente mantenida que está determinada por el elemento compuesto de mantenimiento y una secuencia saliente nativa opuesta que está determinada por la secuencia nativa del segundo fragmento de ADN para ensamblaje;

en donde

(i) el primer y segundo fragmentos de ADN adaptados son fragmentos terminales en donde sus secuencias salientes nativas son complementarias, de modo que están dispuestas para ligarse entre sí; o

- 40 (ii) se proporcionan uno o más fragmentos de ADN medios para ensamblaje en donde el primer y segundo fragmentos de ADN adaptados son fragmentos terminales respectivos en el ensamblaje, y el uno o más fragmentos de ADN medios están dispuestos para ligarse entre el primer y segundo fragmentos de ADN adaptados a través de secuencias nativas salientes complementarias;

- 45 que comprende además el paso de ligar juntos, con una ligasa, el primer y segundo fragmentos de ADN adaptados, o el primer y segundo fragmentos de ADN adaptados y uno o más fragmentos de ADN intermedios, para formar un fragmento de ADN ensamblado que tiene salientes mantenidos en cada extremo.

- 50 El uno o más fragmentos de ADN intermedios pueden comprender secuencias salientes nativas en ambos extremos que están determinadas por su secuencia nativa. En una descripción que comprende un único fragmento de ADN medio para ensamblaje, el fragmento de ADN medio puede comprender una primera secuencia saliente nativa que es complementaria a la saliente nativa del primer fragmento de ADN adaptado y una segunda secuencia saliente nativa que es complementaria a la saliente nativa del segundo fragmento de ADN adaptado.

En una descripción que comprende dos o más fragmentos de ADN medios para ensamblaje, los fragmentos de ADN medios pueden comprender cada uno de ellos secuencias salientes nativas que son complementarias al saliente nativo

de un fragmento de ADN medio vecino, de modo que se pueden ligar entre sí en una secuencia previa. -orden determinado. El primer y último fragmento de ADN intermedio en una secuencia se dispondrán para ligarse al primer fragmento de ADN adaptado y al segundo fragmento de ADN adaptado respectivos mediante secuencias sobresalientes nativas complementarias.

5 En una descripción, un fragmento de ADN medio para ensamblaje puede ser proporcionado por

proporcionar un ácido nucleico metilado linealizado adicional como se describe en el presente documento, en donde el ácido nucleico metilado linealizado comprende un elemento compuesto truncado en cada extremo, en donde los elementos compuestos truncados están metilados con la metilasa del ADN;

10 proporcionar un fragmento de ADN medio adaptado para ensamblaje que tiene extremos salientes que están adaptados para ligarse a los extremos salientes del ácido nucleico metilado linealizado;

ligar el fragmento de ADN medio adaptado para el ensamblaje y el ácido nucleico metilado linealizado con una ligasa para formar un vector intermedio metilado;

15 transformar el vector intermedio metilado en una cepa bacteriana que no expresa la(s) metilasa del ADN(s) que reconoce(n) cualquiera de las secuencias de reconocimiento de metilasa del ADN de los elementos compuestos truncados, formando de ese modo un vector intermedio que no está metilado;

aislar el vector intermedio;

20 cortar el vector intermedio con enzimas de restricción que reconocen la secuencia de reconocimiento de la enzima de restricción de los elementos compuestos truncados, formando así un inserto de fragmento de ADN medio que comprende secuencias salientes nativas en cada extremo que están determinadas por la secuencia nativa del fragmento de ADN medio para el ensamblaje.

25 Los fragmentos de ADN para el ensamblaje pueden comprender ADN bicatenario que comprende los salientes/extremos pegajosos en cada extremo. Se puede preparar un fragmento de ADN para ensamblaje mediante PCR a partir de una plantilla, mediante síntesis génica o mediante hibridación de dos oligonucleótidos. Tales procesos, p.ej. La PCR o la síntesis de genes pueden generar ADN bicatenario con extremos romos. Para el ensamblaje que comienza con el producto de la PCR, se puede incluir una secuencia adaptadora adecuada que contenga el sitio de restricción en el cebador de la PCR para recortar el ADN y generar los salientes para la clonación en los vectores intermedios. Para descripciones que comprenden el uso de oligonucleótidos hibridados hay más opciones. El diseño de la secuencia podría ser el mismo que para la PCR/síntesis de genes con sitios de restricción en ambos extremos para recortar con una enzima de restricción para generar los salientes. Alternativamente, dos oligos recocidos pueden formar los propios salientes.

30 En una descripción, los fragmentos de ADN para ensamblaje, tales como el primer fragmento de ADN para ensamblaje, el segundo fragmento de ADN para ensamblaje y uno o más fragmentos de ADN medios adaptados para ensamblaje, pueden proporcionarse mediante PCR a partir de una secuencia plantilla (es decir, los fragmentos de ADN para ensamblaje pueden ser productos de PCR). La secuencia plantilla puede proporcionar la secuencia completa que se desea clonar como fragmento de ADN ensamblado. Los fragmentos de ADN para el ensamblaje pueden replicarse, por ejemplo mediante PCR, a partir de diferentes secciones de la plantilla, y los productos de PCR vecinos generados pueden solaparse parcialmente en secuencia (es decir, la secuencia cerca del final de un producto de PCR puede comprender la misma secuencia cerca del inicio) del siguiente producto de PCR para su ensamblaje). El solapamiento puede ser parcial, pero de longitud suficiente para proporcionar un extremo adhesivo complementario. En una descripción, la secuencia de solapamiento dentro de un producto de PCR (fragmento de ADN para ensamblaje) y la solapamiento en el siguiente producto de PCR vecino (fragmento de ADN para ensamblaje) pueden tener al menos 2 pares de bases de longitud. En otra descripción, la secuencia de solapamiento dentro de un producto de PCR (fragmento de ADN para ensamblaje) y la solapamiento en el siguiente producto de PCR vecino (fragmento de ADN para ensamblaje) pueden tener al menos 3 pares de bases de longitud. En otra descripción, la secuencia de solapamiento dentro de un producto de PCR (fragmento de ADN para ensamblaje) y la solapamiento en el siguiente producto de PCR vecino (fragmento de ADN para ensamblaje) pueden tener una longitud de 3 a 6 pares de bases. En otra descripción, la secuencia de solapamiento dentro de un producto de PCR (fragmento de ADN para ensamblaje) y la solapamiento en el siguiente producto de PCR vecino (fragmento de ADN para ensamblaje) pueden tener una longitud de 3 a 5 pares de bases. En otra descripción, la secuencia de solapamiento dentro de un producto de PCR (fragmento de ADN para ensamblaje) y la solapamiento en el siguiente producto de PCR vecino (fragmento de ADN para ensamblaje) pueden tener una longitud de 4-5 pares de bases. En otra descripción, la secuencia de solapamiento dentro de un producto de PCR (fragmento de ADN para ensamblaje) y la solapamiento en el siguiente producto de PCR vecino (fragmento de ADN para ensamblaje) pueden tener una longitud de 3-4 pares de bases. En otra descripción, la secuencia de solapamiento dentro de un producto de PCR (fragmento de ADN para ensamblaje) y la solapamiento en el siguiente producto de PCR vecino (fragmento de ADN para ensamblaje) pueden tener 4 pares de bases de longitud.

Los cebadores de PCR pueden disponerse para amplificar las secciones parcialmente solapadas del ADN molde con el fin de formar los fragmentos de ADN para el ensamblaje que comprenderían salientes complementarias. La reacción

- de PCR puede agregar secuencias adaptadoras que contienen la secuencia de reconocimiento de restricción (es decir, la secuencia adaptadora universal, por ejemplo la secuencia de reconocimiento de la enzima de restricción puede comprender la misma secuencia que la secuencia de reconocimiento de la enzima de restricción del elemento compuesto de mantenimiento y/o del elemento compuesto truncado) más la secuencia del saliente que sería complementario al saliente del fragmento vecino/consecutivo para su ensamblaje. El saliente que comprende la secuencia que se solapa, por ejemplo en 4 pb, entre fragmentos consecutivos/vecinos puede formarse después de la digestión con la enzima de restricción (es decir, el solapamiento entre el producto de PCR se refiere a la secuencia, por ejemplo de 4 pb, que está dentro del adaptador completo). secuencia que comprende la secuencia de reconocimiento de la enzima de restricción).
- Los cebadores de PCR pueden proporcionar además las secuencias adaptadoras para los fragmentos de ADN, que forman el saliente complementario para la ligación con los ácidos nucleicos metilados linealizados (tales como el primer ácido nucleico metilado linealizado, el segundo ácido nucleico metilado linealizado o ácido nucleico metilado linealizado adicional) y formación de los vectores intermedios. Los cebadores de PCR pueden proporcionar además una secuencia de reconocimiento de enzima de restricción que está dispuesta para dirigir la enzima de restricción para que corte el producto de PCR para formar las secuencias adaptadoras. La secuencia de reconocimiento de la enzima de restricción proporcionada por los cebadores de PCR puede ser una secuencia de reconocimiento de la enzima de restricción. La secuencia de reconocimiento de la enzima de restricción puede comprender la misma secuencia que la secuencia de reconocimiento de la enzima de restricción del elemento compuesto de mantenimiento y/o del elemento compuesto truncado.
- En una descripción, el fragmento de ADN medio adaptado comprende una región de secuencia nativa solapada que es la misma secuencia que una secuencia saliente nativa proporcionada en un fragmento de ADN consecutivo/vecino destinado a la ligación. Por ejemplo, cuando fragmentos de ADN vecinos/consecutivos para ensamblarse entre sí están codificados a partir de una plantilla, cada uno de ellos puede comprender una región de secuencia solapada/misma de la plantilla, que formaría los salientes nativos complementarios entre las dos secuencias vecinas/consecutivas en un ligadura posterior. En una descripción, el fragmento de ADN medio adaptado comprende dos regiones de secuencia nativa solapada, que están dispuestas para formar los salientes nativos del fragmento de ADN medio una vez cortado por la enzima de restricción de reconocimiento.
- El método puede comprender además el paso de proporcionar un vector de destino linealizado para la inserción de los fragmentos de ADN ensamblados. El vector de destino linealizado puede comprender secuencias sobresalientes respectivas que son complementarias a las proyecciones mantenidas del fragmento de ADN ensamblado. Las secuencias salientes del vector de destino linealizado pueden comprender un primer saliente que es complementario a la secuencia saliente mantenida del primer fragmento de ADN adaptado y un segundo saliente que es complementario a la segunda secuencia saliente mantenida del segundo fragmento de ADN adaptado.
- El vector de destino linealizado se puede proporcionar cortando un vector de destino circular con la(s) enzima(s) de restricción que están dispuestas para proporcionar los mismos salientes complementarios que los fragmentos de ADN ensamblados. Por ejemplo, el vector de destino linealizado puede proporcionarse cortando un vector de destino circular con la(s) enzima(s) de restricción que reconocen las secuencias de reconocimiento de enzimas de restricción del elemento compuesto de mantenimiento y/o truncamiento. El vector de destino linealizado puede ser un vector de clonación.
- En una descripción, el elemento compuesto de mantenimiento comprende la misma secuencia en todos los vectores intermedios que comprenden el elemento compuesto de mantenimiento. En una descripción, el elemento compuesto truncado comprende la misma secuencia en todos los vectores intermedios. En una descripción, el elemento compuesto de mantenimiento comprende la misma secuencia de reconocimiento de enzimas de restricción que el elemento compuesto truncado. En una descripción, las secuencias de reconocimiento de enzimas de restricción de los elementos compuestos de mantenimiento y/o los elementos compuestos truncados de los vectores intermedios son las mismas. En particular, se pueden usar las mismas especies de enzimas de restricción para la restricción de los vectores intermedios y, opcionalmente, el vector de destino.
- Los vectores utilizados en el método de ensamblaje sin cicatrices como se describe en el presente documento pueden ser vectores con un número de copias alto. Por ejemplo, el primer y/o segundo ácido nucleico metilado linealizado pueden comprender un vector metilado linealizado de alto número de copias.
- En una descripción, el método de ensamblaje de ADN descrito en el presente documento se lleva a cabo en una única composición (es decir, el llamado proceso "de un solo recipiente"). Por ejemplo, el proceso de ensamblaje de ADN (incluyendo digestión con enzima de restricción y ligación) se puede llevar a cabo en una única composición que contiene la enzima de restricción y ADN ligasa usando plásmidos de inserción circular o fragmentos de inserción linealizados y vectores circulares intermedios y/o de destino, o linealizados. versiones de los mismos.
- En una descripción, la metilación del ácido nucleico se lleva a cabo en una cepa bacteriana modificada para expresar la metilasa que reconoce la secuencia de reconocimiento de la metilasa. En una descripción, la metilación del ácido nucleico se lleva a cabo *in vitro*.

También se describe una cepa bacteriana modificada que se modifica para expresar una metilasa del ADN descrita en el presente documento.

- 5 La cepa bacteriana modificada que se modifica para expresar una metilasa del ADN puede ser una cepa de *E. coli*. La modificación puede comprender la inserción del gen de la metilasa en el lugar *arsb* del genoma *E. coli*. La cepa bacteriana modificada que se modifica para expresar una metilasa del ADN puede formarse mediante ensamblaje de un vector con un casete de expresión de metilasa seguido de un casete de selección de antibióticos, tal como un casete de selección de antibióticos de zeocina, y flanqueado por una secuencia homóloga del gen *arsb* de *E. coli*, seguido de PCR usando esto como plantilla para generar ADN lineal que contiene el elemento anterior, y recombinación con este fragmento seguido de selección con el antibiótico, tal como zeocina, para generar cepas de bacterias con el casete de expresión de metilasa y el marcador de selección insertados de manera estable en el *arsb* lugar del genoma *E. coli*.

También se describe el uso de una cepa bacteriana modificada que se modifica para expresar una metilasa del ADN descrita en el presente documento, para fabricar una molécula de ácido nucleico como se describe en el presente documento.

- 15 La cepa bacteriana modificada puede ser *E. coli*.

También se describe una composición que comprende el ácido nucleico según la descripción en el presente documento. La composición puede comprender tampón.

También se describe un kit que comprende uno o más ácidos nucleicos según la descripción en el presente documento.

- 20 El kit puede comprender además una enzima de restricción y/o una ligasa, tal como ADN ligasa T4. Adicional o alternativamente, el kit puede comprender una cepa bacteriana modificada que se modifica para expresar una metilasa del ADN según la descripción en el presente documento y/o una metilasa del ADN.

Adicional o alternativamente, el kit puede comprender una o más soluciones tampón. Adicional o alternativamente, el kit puede comprender antibióticos para la selección de clones.

- 25 Uno o más, o todos, el ácido nucleico, el tampón, la cepa bacteriana, la enzima de restricción, la ligasa y la metilasa pueden proporcionarse en un recipiente, tal como un tubo Eppendorf.

También se describe una célula huésped que comprende ácido nucleico según la descripción en el presente documento.

- 30 También se describe el uso del ácido nucleico según la descripción en el presente documento para ensamblar fragmentos de ADN de interés, opcionalmente en donde el uso es *in vitro*.

- 35 La referencia al término "sin cicatrices" o "ensamblaje sin cicatrices" en el presente documento pretende referirse a una secuencia de ADN que se ha ensamblado a partir de múltiples secuencias, en donde la unión entre las secuencias no comprende artefactos, tales como pares de bases no deseados, de la restricción. y ligadura del proceso de ensamblaje del ADN. Por ejemplo, el ensamblaje de ADN sin cicatrices no introduce secuencias adicionales en el fragmento de ADN ensamblado. En algunas publicaciones también se le llama ensamblaje de ADN sin fisuras. En un ejemplo, una secuencia adaptadora diseñada para proporcionar extremos adhesivos complementarios para facilitar la unión de fragmentos de ADN puede considerarse una "cicatriz" entre las secuencias unidas.

La referencia al término "adaptador" en el presente documento pretende referirse a una secuencia diseñada para proporcionar extremos adhesivos complementarios para facilitar la unión de fragmentos de ADN.

- 40 La referencia al término "solapar" o "solaparse" en el contexto de secuencias de reconocimiento solapadas pretende referirse a que al menos parte de la secuencia de reconocimiento de una enzima (por ejemplo cualquier nucleótido distinto de N en una secuencia de reconocimiento) está en una posición de la secuencia de nucleótidos que especifica la secuencia de reconocimiento de otra enzima. (es decir, secuencias solapadas comparten al menos algunos nucleótidos comunes).

- 45 La referencia al término "nativo" en el contexto de una secuencia de extremo sobresaliente/pegajoso pretende referirse a la secuencia natural, no adaptada, del fragmento de ADN de interés que se insertará o unirá al ácido nucleico. Por ejemplo, la secuencia no se diseñará ni modificará con respecto a la secuencia del ADN molde de interés. Esto contrasta con los salientes determinados por la secuencia del vector para facilitar el ensamblaje del ADN, que podrían introducirse como una cicatriz en la secuencia ensamblada.

- 50 La referencia a una "molécula de ADN ensamblada única" pretende referirse a un único tipo de molécula, y no pretende referirse al número de copias de la misma molécula. En particular, en cualquier reacción o composición dada, puede haber una pluralidad de copias de una única molécula de ADN ensamblada.

La referencia a plásmidos/vectores de “alto número de copias” puede referirse a aquellos con un número de copias de 100 o más copias por célula, incluidos, entre otros, plásmidos con orígenes de replicación tales como 1. origen de replicación pMB1 mutado a partir de plásmidos pUC (Vieira y Messing 1982. Gen 19:259-268; Vieira y Messing 1987. Métodos Enzymol. 153:3-11; Messing 1983. Métodos Enzymol. 101:20-78; Lin-Chao et al. 1992. Mol. Microbiol. 6:3385-3393); y 2. Origen de replicación derivado de Co1E1 a partir de plásmidos pbluescript (<https://www.chem-agilent.com/pdf/strata/212205.pdf> - Vectores fagémidos pbluescript II, manual de instrucciones, catálogo n.º 212205, n.º 212206. #212207 y #212208, Revisión A.01). El origen de replicación que porta pbluescript es de pUC con una mutación de un solo nucleótido. Se pueden encontrar guías particulares sobre la regulación de plásmidos/vectores de copias bajas y altas en Molecular Cloning A Laboratory Manual, 3ra (2001), Joseph Sambrook y David W Russell, volumen 1 página 1.4 Tabla 1-1 y volumen 1 página 4.2 Tabla 4-1.

La referencia a plásmidos/vectores de “bajo número de copias” puede referirse a aquellos con un número de copias inferior a 100 copias por célula, incluidos, entre otros, plásmidos con origen de replicación tal como 1. origen de replicación p15A a partir de plásmidos pACYC (Chang y Cohen 1978. J. Bacteriol. 134:1141-1156); 2. Origen de replicación de pSC101 a partir de plásmidos pSC101 (Stoker et al. 1982. Gen 18:335-341); 3. Origen de la replicación F a partir de plásmidos F (Kim y cols. 1996. Genómica 34:213-218; Asakawa et al. 1997. Gen 191: 69-79); 4. Origen de replicación de pMB1 a partir de plásmidos pBR322 (Bolívar et al. 1977. Gen 2:95-113); y 5. Origen de replicación de Co1E1 a partir del plásmido Co1E1 (Kahn et al. 1979. Métodos Enzymol. 68:268-280)

El experto entenderá que las características opcionales de una descripción pueden ser aplicables, cuando sea apropiado, a otras descripciones.

Figura 1. Identificación de metilasas que bloquean la solapamiento de sitios de metilación/restricción.

A. Diagrama de sitios de metilación/restricción solapados para el cribado inicial. Los sitios de reconocimiento de enzimas de restricción están encuadrados en líneas continuas. El extremo adhesivo generado por la enzima de restricción se muestra con una línea continua. Los sitios de reconocimiento de las enzimas de metilación están encuadrados con líneas discontinuas y las bases metiladas están en negrita. Todas las metilasas enumeradas modifican la N6-adenina, excepto M.SacI y M.AspJHL3I, que modifican la C5-citosina y la N4-citosina respectivamente.

B. Diseños experimentales para probar el bloqueo de sitios de metilación/restricción solapados mediante metilación *in vivo*. Para cada combinación de metilasa/enzima de restricción analizada, el plásmido de prueba contiene un sitio de metilación/restricción solapado y un sitio de restricción no metilable. La digestión de restricción del plásmido de prueba preparado a partir de una cepa normal daría como resultado una restricción en ambos sitios y liberaría un fragmento de ~600 pb de la estructura principal del vector de ~4,3 kb. La digestión de restricción del plásmido de prueba preparado a partir de una cepa que expresa metilasa daría como resultado una única banda de ~4,9 kb, debido al bloqueo de los sitios de metilación/restricción solapados por metilación *in vivo*. Se muestran los sitios de restricción del plásmido de prueba para cada enzima de restricción, con el sitio de restricción encuadrado en línea continua, la secuencia de metilación encuadrada en línea discontinua y las bases metiladas en negrita. La disposición directa del sitio de metilación/restricción solapado permite analizar la actividad de la nickasa en el mismo ensayo. C. Análisis de electroforesis en gel de agarosa del plásmido de prueba para cada combinación solapada de metilasa/enzima de restricción preparada en una cepa normal o en una cepa que expresa metilasa del ADN y digerida con las enzimas de restricción de tipo IIS correspondientes. Las combinaciones ensayadas son BsaI con M.Osp807II (M.Osp) utilizando el plásmido de prueba pMOP_BsaINC, BpiI con M2.NmeMC58II (M2.Nme) utilizando el plásmido de prueba pMOP_BpiINC y LguI con M.XmnI utilizando el plásmido de prueba pMOP_LguINC. Las condiciones de la prueba son 60 fmol de plásmido de prueba digeridos usando 5 U de BsaI o BpiI, o 2,5 U de LguI en 10 µl de reacción a 37 °C durante 1 h. El resultado muestra que cada una de las metilasas probadas produce un bloqueo exitoso a través de la metilación *in vivo* de un sitio de restricción para la enzima de restricción de tipo IIS correspondiente cuando el sitio de reconocimiento de metilasa se solapa al sitio de la enzima de restricción. Los datos mostrados representan resultados de tres experimentos independientes.

Figura 2. Sistema MetcIo basado en BsaI-M.Osp807II

Los plásmidos donantes contienen fragmentos de ADN para ensamblar (Frag A-D) flanqueados por sitios BsaI que generan extremos adhesivos compatibles (etiquetados aaaa-eeee). Los sitios BsaI se solapan con la secuencia de reconocimiento de metilasa M.Osp807II. Los plásmidos donantes se prepararon a partir de una cepa normal que carece de actividad metilasa M.Osp807II. Como resultado, los sitios BsaI no están metilados y, por lo tanto, los fragmentos de ADN insertados pueden liberarse mediante digestión con BsaI. El vector de ensamblaje contiene un marcador de selección LacZ flanqueado por sitios BsaI enfrentados. El par exterior de sitios BsaI más cercanos al esqueleto del vector se solapa con la secuencia de metilación de M.Osp807II, mientras que el par interior no. La preparación del vector de ensamblaje en la cepa DH10B que expresa M.Osp807II da como resultado el bloqueo específico del par externo de sitios BsaI. BsaI aún podría liberar el fragmento LacZ a través de la acción en el par interior de sitios BsaI. Después de una reacción en un solo recipiente utilizando ADN ligasa BsaI y T4, la ligación entre extremos adhesivos compatibles da como resultado un ensamblaje ordenado de fragmentos de ADN en la columna vertebral del vector de ensamblaje. El fragmento

ensamblado en el plásmido ensamblado está flanqueado por sitios Bsal bloqueados refractarios a la digestión con Bsal. Después de la transformación en una cepa normal que carece de actividad M.Osp807II, la metilación de los sitios de restricción flanqueantes se pierde y Bsal podría liberar el fragmento ensamblado para el ensamblaje de la siguiente etapa.

Figura 3. Ensamblaje de ADN utilizando el vector Metc10 que contiene elementos de ADN funcionales entre el sitio de restricción protegible por metilación y el sitio de restricción no protegible.

A. Diagrama que muestra el ensamblaje de una unidad de transcripción para la expresión de una proteína con tres dominios diferentes (D1, D2, D3) utilizando el vector Metc10 con el elemento promotor (Prom) y la señal poliA (poliA) preincrustados en el vector. Los sitios Bsal protegidos por metilación están encuadrados en línea continua, los sitios Bsal funcionales en línea de puntos y las bases metiladas en negrita. XXXX y YYYY representan la secuencia adaptadora en los extremos 5' y 3' del fragmento de ADN ensamblado (unidad de transcripción) que podría liberarse para la siguiente ronda de ensamblaje de ADN. Pueden o no ser diferentes de las secuencias del adaptador interno (AATG, ACTA, ATCT y GCTT), porque no participan en la reacción de ensamblaje.

B. Diagrama que muestra el ensamblaje de la misma unidad de transcripción usando Metc10 con el elemento promotor y la señal poliA proporcionada en plásmidos de inserción separados. XXXX y YYYY representan la secuencia adaptadora en los extremos 5' y 3' del fragmento de ADN ensamblado (unidad de transcripción) que podría liberarse para la siguiente ronda de ensamblaje de ADN. Son diferentes de las secuencias del adaptador interno.

Figura 4. Principio de conmutación de extremos adhesivos utilizando vectores de clonación universales.

A. Diagrama de tres tipos de vectores de ensamblaje universal.

B. Detalles del proceso de cambio de extremo adhesivo de vectores de clonación universal a través de Metc10, utilizando el vector VL como ejemplo. Los sitios de restricción de tipo funcional IIS están encuadrados en línea continua, los sitios de restricción bloqueados en línea discontinua y las bases metiladas en negrita.

C. Sistema simbólico de representación del proceso de cambio de extremo del adhesivo. u y v representan los extremos adhesivos universales CTCC y CGAG respectivamente. x e y representan la secuencia de 4 pb dentro de los extremos adhesivos universales definidos por la secuencia. El ensamblaje utilizando diferentes tipos de vectores da como resultado el cambio de diferentes extremos adhesivos.

D. Proceso de cambio de extremo adhesivo aplicado al ensamblaje de múltiples fragmentos.

Figura 5. Ensamblaje de ADN sin cicatrices utilizando vectores de clonación universales

A. El fragmento de ADN que se va a ensamblar no lleva sitios de restricción Bsal internos. El objetivo del proceso de ensamblaje es generar un único fragmento ensamblado que lleva extremos adhesivos x e y definidos por los extremos del fragmento.

B. La secuencia primero se divide digitalmente en 9 fragmentos diferentes (fragmentos A-I) con solapamientos de 4 pb (a-h).

C. Luego se agregaron secuencias adaptadoras a los fragmentos digitalmente, lo que dio como resultado fragmentos de ADN que podrían recortarse con Bsal para generar fragmentos flanqueados por los extremos adhesivos universales u y v.

D. Los fragmentos de ADN bicatenario sintetizados se clonaron primero en vectores de ensamblaje universales específicos para cambiar los extremos adhesivos dependiendo de la posición del fragmento en la siguiente etapa de ensamblaje.

E. Los fragmentos clonados se ensamblaron 3 por grupo en vectores de ensamblaje universales específicos, dependiendo el tipo de vector de la posición del fragmento ensamblado en la siguiente etapa de ensamblaje.

F. La etapa final del ensamblaje genera el fragmento completamente ensamblado con extremos adhesivos x e y de secuencia especificada.

Figura 6. Ventaja del ensamblaje universal sin cicatrices basado en Metc10 en comparación con el ensamblaje universal basado en Moc10

Para el ensamblaje de múltiples etapas utilizando vectores de ensamblaje universales, el sistema basado en Moc10 utiliza dos enzimas de restricción diferentes (Bsal y Bpil) para etapas consecutivas, lo que requiere diferentes adaptadores universales para un bloque determinado (CTCC para Bsal y AGAA para Bpil). Para clonar los fragmentos de diferentes etapas en vectores de ensamblaje universal, el extremo 5' del primer bloque y el extremo 3' del último bloque en cada subconjunto deben cambiar entre diferentes secuencias adaptadoras

dependiendo de la enzima utilizada para el ensamblaje. La ilustración muestra el cambio del adaptador 5' para el primer bloque (bloque A) durante 3 etapas consecutivas de la reacción de ensamblaje, con los sitios BsaI encuadrados en línea continua y los sitios BpiI encuadrados en líneas discontinuas. Es necesario agregar un adaptador muy largo al extremo de 5' del bloque A para hacer frente al cambio de adaptadores durante las diferentes etapas del ensamblaje. El adaptador se "mastica" 4 pb después de cada etapa de ensamblaje, por lo que la longitud de la secuencia del adaptador depende del número de etapas. El diseño basado en MetcI es mucho más sencillo, porque utiliza una única enzima BsaI para las diferentes etapas de ensamblaje. Aquí los sitios BsaI no metilables están recuadrados en línea continua y los sitios BsaI bloqueados por metilación están recuadrados en línea discontinua. La secuencia de adaptador universal CTCC en el extremo 5' del primer bloque (bloque A) se puede utilizar para diferentes etapas de ensamblaje. No es necesario cambiar la secuencia del adaptador para los bloques exteriores. Como resultado, el diseño del adaptador para la secuencia inicial es mucho más sencillo.

Figura 7. Ensamblaje de ADN usando MetcI. La figura ilustra el ensamblaje de cuatro fragmentos de ADN (Fragmentos A, B, C y D) en un único fragmento de ADN (Fragmento ABCD) a través de dos ciclos de MetcI utilizando una única enzima de restricción de tipo IIS. La etapa 1 y la etapa 2 son efectivamente el mismo proceso con inserciones progresivamente más grandes. En la Etapa 1, los plásmidos donantes del inserto contienen fragmentos del inserto (Fragmentos A, B, C o D) flanqueados por sitios de restricción metilables de tipo IIS que generan extremos adhesivos compatibles. La preparación de plásmidos insertados en cepas normales de *E. coli*, que carecen de la metilasa específica del sitio adecuada, da como resultado fragmentos insertados liberables por la enzima de restricción tipo IIS. El vector de ensamblaje contiene el marcador de selección negativa LacZα flanqueado por una configuración de cabeza a cabeza diseñada expresamente de sitios de restricción de tipo IIS. El par interior de los sitios de restricción está diseñado para que no sean metilables, pero el par exterior sí lo es mediante una metilasa específica del sitio porque la secuencia de reconocimiento de la enzima de restricción se solapa con la secuencia de reconocimiento de la metilasa. La preparación del vector de ensamblaje en cepas de *E. coli* que expresan la metilasa correspondiente conduce a la metilación selectiva del par externo de sitios de restricción metilables, y la digestión con la enzima de restricción de tipo IIS genera una columna vertebral del vector que retiene el par externo de sitios de restricción de tipo IIS metilados. La ligación de los fragmentos del inserto con la estructura principal del vector genera un plásmido correctamente ensamblado refractario a la restricción mediante la enzima de restricción tipo IIS. Todos los demás fragmentos no deseados, incluida la cadena principal del vector del plásmido insertado y el fragmento LacZα del vector de ensamblaje, contienen sitios de restricción activos de tipo IIS, que son sensibles a la restricción por parte de la enzima. Este esquema permite una reacción de ensamblaje en un solo recipiente que combina el paso de restricción y ligación porque la reacción favorece la generación del plásmido ensamblado correctamente. La transformación de la reacción de ensamblaje en una cepa de *E. coli* normal, que carece de la metilasa específica del sitio apropiada, da como resultado efectivamente la desmetilación de los sitios de restricción metilados de tipo IIS que flanquean el fragmento ensamblado en el plásmido ensamblado. En la Etapa 2, los insertos ensamblados (Fragmentos AB y CD) se pueden usar para una ronda posterior de ensamblaje de MetcI en un fragmento más grande (Fragmento ABCD).

Figura 8. Salto de nivel: esquema de modificación de ADN utilizando componentes de diferentes niveles de ensamblaje. El diagrama muestra el esquema para el ensamblaje de dos fragmentos de ADN (Fragmentos A y B) mediante el esquema estándar de MetcI y la posterior modificación del fragmento ensamblado (Fragmento AB) utilizando un tercer fragmento (Fragmento C). Los fragmentos iniciales utilizados (Fragmentos A, B y C) son del mismo nivel en la jerarquía de ensamblaje, en el sentido de que llevan el mismo marcador de selección de antibióticos. El paso de modificación modifica un fragmento existente (Fragmento AB) utilizando un fragmento (Fragmento C) de un nivel diferente de ensamblaje, por lo tanto, "salto de nivel". El fragmento modificado (Fragmento ABC) lleva los mismos extremos adhesivos que el fragmento inicial (Fragmento AB) y un marcador de selección de antibiótico (M3) diferente de los bloques de nivel 1 y 2 (M1 y M2). Por lo tanto, el nuevo fragmento ABC puede usarse para el siguiente nivel de ensamblaje de ADN usando el marcador de selección M1 en lugar del fragmento AB. El mantenimiento de los extremos adhesivos durante el paso de modificación se logra mediante el uso de una disposición especial de sitios de restricción de tipo IIS cabeza a cabeza (a la derecha del fragmento LacZα durante el paso de modificación), en el que el interior de tipo IIS no metilable. El sitio de restricción genera un extremo adhesivo que es compatible con el fragmento de modificación a insertar (Fragmento C), y el sitio de restricción exterior metilable tipo IIS genera un extremo adhesivo idéntico al extremo del fragmento de ADN a modificar (el extremo derecho del Fragmento AB en el diagrama). El paso de modificación no es posible con el método Moclo estándar que utiliza dos enzimas de restricción de tipo IIS, porque los fragmentos de ADN de niveles consecutivos de ensamblaje tendrían que ser liberados por dos enzimas de restricción diferentes, cada una de las cuales tiene que ser diferente de la enzima utilizada para el siguiente nivel de ensamblaje de ADN después del paso de modificación. El esquema de modificación de salto de nivel es útil para la inserción de elementos genéticos funcionales, como unidades de transcripción, en un proceso de ensamblaje existente sin cambiar el diseño de ensamblaje de nivel superior, para maximizar la reutilización de las partes de ADN existentes con secuencia verificada.

Figura 9. Inserción de nivel: esquema para el ensamblaje del ADN a través de una etapa intermedia de ensamblaje. El diagrama muestra el esquema de ensamblaje de cuatro fragmentos de ADN (Fragmentos A, B, C y D) en un solo fragmento (Fragmento ABCD) a través de una etapa intermedia de ensamblaje de ADN (Nivel de

inserción). A diferencia del esquema de ensamblaje estándar de MetClo (Figura 3) que ensambla los cuatro fragmentos en una sola reacción, este esquema implica una etapa de ensamblaje intermedia mediante el subensamblaje de los fragmentos B y C en un vector de ensamblaje intermedio que lleva un marcador de selección (M3) diferente del de los insertos de ADN de nivel 1 (M1 transportado por los fragmentos A, B, C y D) y del fragmento de ADN ensamblado de nivel 2 (M2 transportado por los fragmentos ABCD). El esquema de subensamblaje de inserción de nivel ofrece más opciones en las rutas de ensamblaje del ADN y aumenta la reutilización de los fragmentos de ADN ensamblados intermedios.

Figura 10. Suma lineal: esquema para la construcción de vectores de ensamblaje usando MetClo basado en Bsal.

El diagrama muestra el esquema para la adición lineal de partes de ADN para la construcción de vectores de ensamblaje que transportan elementos genéticos funcionales existentes (Fragmentos A y D). Comenzando con un vector de ensamblaje que lleva un fragmento de ADN de relleno no seleccionable flanqueado por sitios de restricción de tipo IIS no metilables, primero se ensamblan dos fragmentos de ADN (Fragmentos A y D) con un marcador de selección LacZα en este vector. El marcador de selección LacZα está flanqueado por una disposición cabeza a cabeza especialmente diseñada de sitios de restricción tipo IIS, de modo que los sitios internos son metilables y los sitios externos no. La preparación del plásmido de inserción LacZα en una cepa de *E. coli* que expresa la metilasa correspondiente da como resultado el bloqueo selectivo de los pares internos de sitios de restricción metilables de tipo IIS. El ensamblaje en un solo recipiente del inserto LacZα con otros insertos preparados a partir de la cepa *E. coli* normal (Fragmentos A y D) da lugar a un fragmento de ADN ensamblado refractario a la restricción por la enzima de restricción de tipo IIS utilizada para el ensamblaje del ADN. La transformación del ADN ensamblado en una cepa normal de *E. coli* que carece de la correspondiente actividad metilasa da como resultado la desmetilación de los sitios de restricción metilables de tipo IIS que flanquean el fragmento LacZα. Como resultado, el ADN ensamblado se puede utilizar como vector de ensamblaje para la adición de nuevos fragmentos de ADN (Fragmentos B y C) en lugar del marcador de selección LacZα en la siguiente ronda de ensamblaje de ADN. El resultado neto es la adición lineal de fragmentos de ADN en el vector de ensamblaje (fragmentos A y D, luego fragmentos B y C). El esquema es útil para la construcción de vectores de ensamblaje de etapa final que portan elementos genéticos funcionales comunes, como marcadores de selección de mamíferos.

Ejemplo 1: ensamblaje modular asistido por metilación utilizando una única enzima de restricción (MetClo)

Introducción

Descripción general de MetClo

Aquí desarrollamos un método de ensamblaje jerárquico basado en enzimas de restricción tipo IIS mejorado llamado MetClo, que reduce la cantidad de enzimas de restricción utilizadas a una. LoG's esto agregando un par de sitios de restricción de tipo IIS específicos para la misma enzima utilizada para el ensamblaje en el vector de ensamblaje, que se bloquean selectivamente mediante la metilación del ADN. La transformación del ADN ensamblado en cepas normales que carecen de la metilasa específica desbloquea los sitios de restricción flanqueantes, lo que da como resultado plásmidos que transportan el fragmento ensamblado que puede liberarse usando la misma enzima de restricción tipo IIS y luego usarse para una ronda posterior de ensamblaje de ADN. Probamos varias metilasas para determinar su capacidad para bloquear sitios de metilación/restricción solapados para tres enzimas de tipo IIS Bsal, Bpil y Lgul comúnmente utilizadas y disponibles comercialmente, e identificamos metilasas adecuadas para cada una de ellas, a saber, M.Osp807II (para Bsal con secuencia solapada **GACNNGGCTCN**[^]NNNN), M2.NmeMC58II (para Bpil con secuencia solapada **GAAGACGC**[^]NNNN), y M.Xmnl (para Lgul con secuencia solapada **GAA CGCTCT**[^]CN[^]NNN, la base metilada o la base opuesta está subrayada, se requieren bases adicionales para especificar la actividad metilasa en negrita, secuencia de reconocimiento de enzimas de restricción en cursiva; las tres modificaciones son m6A N6-metiladenina). Desarrollamos cepas de *E. coli* que portan copias cromosómicas de casetes de expresión para la expresión constitutiva de estas metilasas y demostramos que el ADN propagado en estas cepas es refractario a la digestión por la enzima de restricción cuando el sitio de restricción se solapa con la secuencia de reconocimiento de metilasa correspondiente. Luego construimos sistemas MetClo de prueba de principio para cada par de metilasa/enzima de restricción para ensamblar cuatro fragmentos en un solo fragmento, que podría ser liberado por la misma enzima de restricción. Finalmente, demostramos la utilidad del sistema MetClo basado en M.Osp807II/Bsal para el ensamblaje jerárquico de un ADN de 218 kb a partir de ADN de 28 x 7,7 kb en 2 etapas utilizando Bsal.

Materiales y métodos

Construcción de plásmido

Los plásmidos se construyeron mediante técnicas de clonación basadas en ligación y enzimas de restricción estándar. Los fragmentos de ADN para la clonación se generaron mediante PCR o mediante síntesis de genes (IDT o Invitrogen). Para probar el ensamblaje modular principal de ADN de aproximadamente 1 kb a partir de 4 fragmentos, la columna vertebral del vector aceptor del ensamblaje modular se basa en el vector Moclo resistente a ampicilina pICH47732 (disponible en Addgene y descrito en A modular cloning system for standardized junction of multigene constructs). Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. PLoS One. 18 de febrero de 2011;6(2):e16765. doi: 10.1371/journal.pone.0016765) y la estructura principal del plásmido donante basada en un vector resistente a

kanamicina construido reemplazando el gen resistente a ampicilina en pICH47732 con un gen resistente a kanamicina. Para el ensamblaje jerárquico de ADN de ~200 kb a partir de 28 fragmentos, los vectores aceptores de ensamblaje modular se basan en la columna vertebral del vector BAC de baja copia resistente a gentamicina o kanamicina con origen de replicación F, y la columna vertebral del plásmido donante se basa en cadenas principales de vector de baja copia resistente a la kanamicina con Origen de replicación de p15A o F. Los plásmidos para la sobreexpresión de metilasa se construyeron basándose en un esqueleto de vector BAC de baja copia resistente a kanamicina con origen de replicación F. Los genes de metilasa estaban bajo el control del promotor sintético J23100 y flanqueados por un casete de selección de zeocina, así como secuencias homólogas del locus *arsB* para facilitar la integración estable en el genoma de *E. coli* mediante recombinación. Los plásmidos plantilla para el ensayo de metilasa se basan en pICH47732 resistente a ampicilina.

Generación de cepas que expresan metilasa.

Se generaron cepas de *E. coli* con expresión estable de metilasas específicas mediante recombinación. Brevemente, se generaron fragmentos de ADN que contienen el gen de la metilasa, el marcador de selección de zeocina y las secuencias homólogas flanqueantes de *arsB* mediante PCR utilizando el plásmido BAC que expresa la metilasa correspondiente como plantilla. Los fragmentos se usaron para la recombinación de rojo lambda en células DH10B seguido de selección con zeocina para insertar el casete de expresión de metilasa en el locus *arsB*. Los clones con inserciones cromosómicas correctas se verificaron mediante PCR y secuenciación. Las cepas que expresan metilasa pueden mantenerse estables sin selección de antibióticos.

Ensamblaje modular

Para probar el ensamblaje modular principal de ADN de ~ 1 kb a partir de 4 fragmentos, las reacciones de ensamblaje se configuraron utilizando 60 fmol de cada uno de los plásmidos donantes preparados a partir de células DH5alfa, 60 fmol de vector de ensamblaje preparado a partir de cepas DH10B que expresan metilasa compatible, 1000 U de ADN ligasa T4 (NEB), y 5U de Bsal (NEB) o Bpil (Fermentas) o 2,5U de Lgul (Fermentas) en 20 µl de tampón de ligasa T4 1x (NEB). La condición de reacción fue: 37°C 15 min, seguido de 45 ciclos de 37°C 2 min más 16°C 5 min, luego 37°C 20 min y 80°C 5 min. Las muestras ensambladas se transformaron en células químicamente competentes DH5alfa y se sembraron en placas LB con selección de AIX (ampicilina 100 µg/ml, IPTG 100 µM, X-Gal 50 µg/ml) a 37 °C durante la noche. Las colonias blancas se ampliaron y examinaron mediante secuenciación de ADN.

Para el ensamblaje jerárquico de ADN de aproximadamente 200 kb a partir de 28 fragmentos, el ensamblaje se llevó a cabo en dos etapas. En la etapa 1, los fragmentos se ensamblaron 7 por grupo usando 15 fmol de cada plásmido donante preparado a partir de células DH10B, 15 fmol de vector de ensamblaje preparado a partir de células DH10B-M.Osp807II, 1000 U de ligasa T4 (NEB) y 5 U de Bsal (NEB) en 20 µl. 1x tampón ligasa T4 (NEB). La condición de reacción fue la misma que la descrita anteriormente. Luego, las muestras ensambladas se dializaron gota a gota frente a 50 ml de dH2O utilizando un filtro de celulosa de 0,05 µm (Millipore) a temperatura ambiente durante 1 h. Se transformaron 5 µl de muestra dializada en 25 µl de células electrocompetentes NEB-10beta mediante electroporación a 0,9 kV, 100 Ω, 25 µF, utilizando Gene Pulser y cubetas de electroporación de 1 mm (Biorad). Después de la electroporación, las células se recuperaron añadiendo 1 ml de medio LB y se incubaron a 37 °C 23,03 rad/s (220 rpm) durante 1 h. Las células se sembraron en placas LB con selección GIX (gentamicina 2,5 µg/ml, IPTG 100 µM, X-gal 50 µg/ml) a 37°C durante la noche. Las colonias blancas se expandieron y seleccionaron mediante digestión de restricción usando XhoI.

En la etapa 2, el ADN de aproximadamente 200 kb se ensambló usando protocolos similares a los de la etapa 1 con las siguientes modificaciones: se usaron 30 fmol de cada plásmido donante y 15 fmol de vector de ensamblaje en la reacción de ensamblaje; Después del ensamblaje modular y antes del paso de diálisis por gota, se agregaron 10 U de Bsal a la reacción de ensamblaje para la incubación a 37 °C durante 45 min, seguido de inactivación por calor a 80 °C durante 5 min. Las células transformadas se sembraron en placas LB con selección KIX (kanamicina 30 µg/ml, IPTG 100 µM, X-gal 50 µg/ml). Las colonias blancas se examinaron mediante PCR utilizando cebadores que detectan la unión entre el segundo y el tercer fragmento de 54 kb (TACCCACGTGATTCACGCTG y ATCCTACAGACTCGCTGTGG). Los clones positivos para PCR seleccionados se examinaron mediante digestión de restricción usando XhoI o Bsal con electroforesis de campo pulsado.

Resultado

El sistema Metc10 se basa en el sistema de ensamblaje de ADN de un solo recipiente basado en enzimas de restricción tipo IIS estándar y agregó un par adicional de sitios de restricción tipo IIS en la columna vertebral del vector de ensamblaje para la liberación del ADN ensamblado para el ensamblaje de la siguiente etapa. A diferencia del sistema Moc10, que utiliza sitios de restricción diferentes de la enzima tipo IIS utilizada en el proceso de ensamblaje, Metc10 utiliza sitios de restricción para la misma enzima. Para evitar que esta enzima corte el plásmido ensamblado en la reacción de ensamblaje en un solo recipiente, Metc10 utiliza metilasa específica para metilar los sitios de restricción de tipo IIS que se solapan con las secuencias de reconocimiento de la metilasa, para bloquear estos sitios de restricción de tipo IIS en la columna vertebral del vector. La metilación se pierde cuando el plásmido ensamblado se transforma en cepas de *E. coli* que carecen de la metilasa específica, lo que permite la liberación del fragmento ensamblado del plásmido ensamblado mediante la misma enzima de restricción de tipo IIS utilizada en la reacción de

ensamblaje. Por lo tanto, el sistema Metclo resultante permite el ensamblaje jerárquico de fragmentos de ADN mediante una reacción en un solo recipiente con una única enzima de restricción de tipo IIS, entre varias aplicaciones.

La clave del sistema Metclo es la identificación de la metilasa adecuada para bloquear los sitios de restricción de tipo IIS que se solapan con la secuencia de reconocimiento de la metilasa en la columna vertebral del vector. Existen varios criterios para la selección de la metilasa adecuada. En primer lugar, la metilación del sitio de metilación/restricción solapado debe dar como resultado el bloqueo de la acción de la enzima de restricción de tipo IIS. El proceso de metilación podría ocurrir por metilación *in vitro* del plásmido vector purificado, o mediante la propagación del plásmido vector en cepas que expresan la metilasa *in vivo*. Se prefiere la metilación *in vivo* debido al costo y tiempo reducidos en el proceso de preparación del vector. En segundo lugar, la secuencia de reconocimiento de la metilasa no debe ser idéntica ni estar completamente rodeada por el sitio de restricción de tipo IIS; Se deben requerir pares de bases adicionales para especificar la acción de la metilasa en el sitio de metilación/restricción solapado. La razón de este requisito es que hay dos pares de sitios de restricción de tipo IIS para la misma enzima en el vector de ensamblaje. La metilasa debe bloquear un par dentro de la cadena principal del vector, mientras que el otro par dentro del inserto del marcador de selección negativa no debe bloquearse. Si la secuencia de reconocimiento de metilasa es idéntica al sitio de restricción de tipo IIS o está incluida en él, todos los sitios de restricción en el vector de ensamblaje para esta enzima de restricción de tipo II en particular se bloquearán mediante metilación, incluido el par dentro del inserto del marcador de selección negativa; el vector de ensamblaje metilado resultante no será cortado en absoluto por la enzima tipo IIS, por lo tanto, no se puede utilizar para el ensamblaje de ADN. En tercer lugar, los pares de bases adicionales para especificar la actividad metilasa no deben solaparse con la secuencia del extremo adhesivo generada por la enzima de restricción tipo IIS. Esto mantendrá la máxima flexibilidad en el diseño de la secuencia adaptadora para el ensamblaje del ADN.

Buscamos metilasas con capacidad predicha para bloquear los sitios de restricción de las enzimas de restricción de tipo IIS BsaI, BpiI y LguI. Se trata de enzimas de restricción de tipo IIS habituales en los sistemas de ensamblaje basados en IIS. Para elegir las metilasas candidatas utilizamos algunos criterios adicionales a los mencionados anteriormente. Nuestro objetivo era generar plásmidos metilados mediante metilación *in vivo*, por lo que lo ideal era que la metilasa fuera activa a la temperatura de crecimiento de *E. coli*. También preferimos metilasas que reconocieran secuencias más largas (5-6 pb), porque con secuencias de reconocimiento más largas, menos bases en el genoma huésped de *E. coli* y en el esqueleto del vector de ensamblaje servirían como sustrato de metilación por casualidad. Esto aumenta potencialmente la eficacia de la metilación y reduce la posibilidad de que la metilación heteróloga del ADN tenga efectos adversos sobre la función del gen huésped. Además, preferimos las metilasas de subunidad única de los sistemas de restricción/modificación de tipo II que carecen de actividad de restricción hacia la secuencia no metilada para evitar posibles problemas de restricción durante la etapa de construcción del vector. Por último, para aumentar la probabilidad de que la metilación bloquee los sitios de metilación/restricción solapados, preferimos metilasas que modifiquen bases cercanas al centro del sitio de restricción.

Con base en estos criterios, seleccionamos las siguientes metilasas para las pruebas: M.Osp807II (GACNNGGTCTC m6A) para BsaI, M.Rsp7740I (TTCAAGAC m6A) y M2.NmeMC58II (GAAGACGC m6A) para BpiI, y M.SacI (GAGCTCTTC m5C). M.AspJHL3I (GAGCTCTTC m4C) y M.XmnI (GAAGCTCTTC m6A) para LguI (base metilada o base complementaria subrayada). Examinamos la actividad de la metilasa preparando un plásmido plantilla de alto número de copias basado en ColE (es decir, derivado de ColE) que lleva sitios de metilación/restricción solapados en células DH10B que expresan la metilasa correspondiente bajo el promotor sintético constitutivo J23100 de un vector BAC de baja copia (estructura del vector BAC de baja copia). con origen de replicación F). Esto identificó a M.Osp807II (para BsaI), M2.NmeMC58II (para BpiI) y M.XmnI (para LguI) como candidatos adecuados para análisis adicionales (Figura 1A). Luego integramos el casete de expresión de metilasa con un marcador de selección de zeocina en el locus *arsB* usando recombinación de rojo lambda para generar cepas estables que sobreexpresan la metilasa a partir de una copia cromosómica (Hecho de células microbianas. 24 de junio de 2013; 12:60. doi: 10.1186/1475-2859-12-60 (PMID 23799955)), y probó la capacidad de estas metilasas para bloquear sitios de metilación/restricción solapados *in vivo*. El vector de prueba lleva un sitio de restricción que no se solapa con la secuencia de metilación, por lo que es permisible para la restricción tanto en la cepa de tipo salvaje como en la cepa que expresa metilasa, y un sitio con dos sitios de restricción de tipo IIS cabeza a cabeza, ambos de los cuales se solapan con la secuencia de metilación correspondiente y, por lo tanto, estarían bloqueados en la cepa que expresa la metilasa. La disposición de cabeza a cabeza permite la evaluación de la actividad de la nickasa, así como la actividad de restricción de doble hebra en el mismo ensayo (Figura 1B). Después de 1 hora de digestión de 60 fmol de plásmido de prueba (pMOP_BsaINC, pMOP_BpiINC y pMOP_LguINC) en 10 µl de reacción usando 5 U (BsaI, BpiI) o 2,5 U (LguI) de enzima a 37 °C, se analizó el ADN mediante electroforesis en gel de agarosa con tinción con EtBr. El resultado muestra que, en condiciones de prueba, las tres metilasas bloquean la restricción de los sitios solapados, como lo demuestra la falta de generación del fragmento de -600 pb (Figura 1C).

Luego desarrollamos sistemas de ensamblaje de ADN Metclo de prueba de principio utilizando estas metilasas para el ensamblaje de ADN en un solo recipiente (la Figura 2 muestra BsaI-MOsp807II como ejemplo). En estos sistemas, los insertos de ADN del donante que se van a ensamblar están flanqueados por un sitio de metilación/restricción solapado. Debido a que los plásmidos donantes que contienen los insertos de ADN que se van a ensamblar se preparan en una cepa de *E. coli* normal que carece de la metilasa correspondiente, el sitio de restricción no está bloqueado y la enzima podría liberar el inserto donante en la reacción de ensamblaje. El vector de ensamblaje lleva un marcador de selección diferente del plásmido donante y un marcador de selección LacZ flanqueado por sitios de

restricción tipo IIS cabeza a cabeza. Para cada sitio de cabeza a cabeza, los dos sitios de restricción de tipo IIS generan el mismo extremo adhesivo. El sitio exterior cercano al esqueleto del vector se solapa con la secuencia de metilación, y el sitio interior cercano al inserto LacZ no. Los vectores de destino se preparan a partir de una cepa de *E coli* que expresa la metilasa correspondiente. Como resultado, sólo los sitios internos de tipo IIS son funcionales en la reacción de ensamblaje. Durante la reacción, el inserto LacZ se libera de la columna vertebral del vector debido a la acción de la enzima en el sitio interior tipo IIS; la ligadura del inserto donante con el vector de destino da como resultado sitios de restricción de tipo IIS dentro de la estructura principal del vector de destino, que se bloquea mediante metilación; La religación del inserto LacZ liberado con la columna vertebral del vector de destino da como resultado la regeneración del sitio tipo IIS cabeza a cabeza, que es sensible a la restricción debido a que el sitio interior no está bloqueado. Como resultado, la reacción favorece la ligación de los fragmentos del donante al esqueleto del vector de destino. Luego, la reacción moclo se transforma en una cepa de *E coli* normal que carece de actividad metilasa específica para detectar la expresión de LacZ utilizando placas LB que contienen IPTG y X-Gal. Las colonias blancas que carecen de actividad LacZ deben contener plásmidos ensamblados con éxito. El fragmento de ADN ensamblado tiene sitios de restricción de tipo IIS flanqueantes que se solapan con secuencias de reconocimiento de metilasa, que son idénticas al vector donante. Como resultado, esto permite una ronda adicional de ensamblaje modular en el siguiente nivel usando la misma enzima de restricción y así sucesivamente durante múltiples rondas según sea necesario.

Probamos los sistemas para las tres enzimas para el ensamblaje a partir de cuatro fragmentos (Tabla 1). Los vectores donantes son resistentes a la kanamicina y los vectores de destino, resistentes a la ampicilina. Los vectores de destino se prepararon a partir de las correspondientes cepas que expresan metilasa. Después de la selección, se seleccionaron 8 colonias blancas para la minipreparación y se analizaron mediante secuenciación Sanger. El resultado de la secuenciación muestra que el 100% (8/8) de los clones son correctos para cada uno de los tres sistemas Moclo (Tabla 1). Esto confirma la utilidad de Moclo para el ensamblaje modular utilizando una única enzima de restricción como Bsal, Bpil o Lglul.

Para explorar más a fondo la utilidad de Moclo para el ensamblaje jerárquico de construcciones de ADN grandes, ensamblamos un fragmento de ADN de 218 kb a partir de fragmentos de 28 x 7,7 kb utilizando la enzima de restricción Bsal en dos etapas. En la etapa uno, se ensamblaron 28 fragmentos de ADN de 7,7 kb que estaban libres de sitios Bsal, 7 por grupo en 4 fragmentos de 54 kb utilizando el sistema Moclo basado en Bsal. En la segunda etapa, las 4 piezas de fragmentos de 54 kb se ensamblaron en un único fragmento de 218 kb. El fragmento final de 218 kb está compuesto por una secuencia no repetitiva con -48 % de contenido de GC. El resumen de restricción muestra una tasa de éxito del -50% para el ensamblaje de la etapa 1 (2/4 a 3/4 clones correctos) y del -20% para el ensamblaje de la etapa 2 (14 de 27 verificados por PCR, de los cuales 3 de 8 verificados mediante restricción) digerir) (Tabla 2). Esto confirma que el sistema Moclo se puede utilizar para el ensamblaje jerárquico de construcciones de ADN grandes utilizando una única enzima de restricción de tipo IIS.

Discusión

Aquí presentamos Moclo, un sistema de ensamblaje de ADN basado en enzimas de restricción de tipo IIS modificado para el ensamblaje de ADN en un solo recipiente utilizando enzimas de restricción de tipo IIS comunes como Bsal, Bpil o Lglul, que genera ADN ensamblado que se puede liberar mediante la misma enzima de restricción. Esto permite el ensamblaje jerárquico de construcciones de ADN grandes utilizando una única enzima de restricción de tipo IIS. El sistema utiliza cepas que sobreexpresan metilasa durante el paso de preparación del vector para *in vivo* metilación de sitios de metilación/restricción solapados para bloquear sitios de restricción selectiva en el vector. El proceso de ensamblaje es el mismo que el de los sistemas de ensamblaje de ADN tipo IIS mediante el uso de reacción en un solo recipiente, por lo que se preserva la simplicidad de los sistemas existentes. Los pasos de preparación y transformación de piezas utilizan cepas normales similares al sistema Moclo estándar, lo que permite el uso de cepas de clonación de rutina para el ensamblaje de ADN, similar a los sistemas de ensamblaje de tipo IIS actuales.

El sistema Moclo tiene una serie de ventajas sobre los sistemas de ensamblaje de ADN jerárquicos basados en enzimas de restricción tipo IIS existentes. En primer lugar, el enfoque de Moclo reduce la carga de eliminar los sitios de restricción internos durante la construcción de los componentes. En comparación con el sistema Moclo estándar que utiliza dos o más enzimas, el sistema Moclo de enzima única basado en Bsal y Bpil utiliza un cortador de 6 pb que reduce la frecuencia de los sitios de restricción prohibidos de 1 por ~1kb a 1 por ~2kb para ADN aleatorio con 50% de contenido de GC, y el sistema Moclo basado en Lglul utiliza un cortador de 7 pb que reduce aún más la frecuencia a 1 por ~8 kb.

Moclo mejora la flexibilidad del sistema de ensamblaje jerárquico de tipo IIS. Dado que los fragmentos de ADN ensamblados procedentes de diferentes etapas de ensamblaje pueden ser liberados por la misma enzima de restricción, con un diseño adecuado de los adaptadores y la elección de un vector resistente a los antibióticos, el sistema permite potencialmente el diseño de varios esquemas de ensamblaje que no pueden conseguirse utilizando el sistema existente. Entre ellos se incluyen el salto de nivel (el ensamblaje de ADN utilizando partes de diferentes etapas de ensamblaje) y la inserción de nivel (el ensamblaje de fragmentos de ADN a través de una etapa de ensamblaje intermedia). Estos esquemas son útiles para la modificación de grandes construcciones de ADN utilizando las piezas existentes, y para el subensamblaje de piezas complejas, como los marcos de lectura abiertos de varias partes, para una mayor reutilización.

También podrían idearse esquemas adicionales para la adición lineal de partes de ADN durante el proceso de ensamblaje. Por ejemplo, durante el ensamblaje del ADN podrían usarse partes de ADN que contengan marcadores de selección flanqueados por sitios de restricción de tipo IIS bloqueados por metilación interna. El plásmido ensamblado podría usarse entonces como vector de ensamblaje para la inserción de partes de ADN adicionales en lugar del marcador de selección en la siguiente etapa de ensamblaje. Esto es particularmente útil para la construcción de vectores de ensamblaje comunes que contienen módulos funcionales.

Metclo también aumenta la intercambiabilidad de piezas de diferentes estándares de ensamblaje. El desarrollo independiente de bibliotecas de piezas y estándares de ensamblaje conduce a sistemas de ensamblaje que utilizan diferentes enzimas de restricción de tipo IIS. Por ejemplo, para biología vegetal, Moclo usa Bsal/Bpil (y Esp3I), Goldenbraid usa Bsal y BsmBI (y BtgZI); para E.coli, Ecoflex utiliza Bsal/BsmBI, CIDAR utiliza Bsal/Bpil. Esto impone limitaciones a la intercambiabilidad de piezas entre diferentes colecciones. El sistema Metclo que utiliza una única enzima de restricción tipo IIS es compatible con cualquier sistema que utilice esa enzima de restricción. En particular, el sistema basado en Bsal es compatible con todos los sistemas existentes que utilizan Bsal. Dado que Bsal se ha utilizado en la mayoría de los sistemas de ensamblaje modular jerárquico basados en enzimas de restricción tipo IIS hasta la fecha (ACS Sintetizador Biol. 21 de octubre de 2016;5(10):1059-1069 (PMID 27096716); ACS Sintetizador Biol. 15 de enero de 2016;5(1):99-103. doi: 10.1021/acssynbio.5b00124 (PMID 26479688; ACS Sintetizador Biol. 18 de septiembre de 2015;4(9):975-86. doi: 10.1021/sb500366v (PMID25871405)); y Más uno. 18 de febrero de 2011;6(2):e16765. doi: 10.1371/journal.pone.0016765 (PMID 21364738), el sistema Metclo basado en Bsal tiene potencial para usarse como estándar para ensamblaje modular que es compatible con la mayoría de las bibliotecas de piezas existentes.

Existen sistemas de ensamblaje de ADN que utilizan la metilación del ADN para bloquear los sitios de restricción internos de tipo IIS durante el proceso de ensamblaje. Por ejemplo, el sistema de ensamblaje TNT (Sci Rep. 13 de enero de 2016; 6: 19278. doi: 10.1038/srep19278 (PMID 26758940)) utiliza M.TaqI para metilar el sitio de reconocimiento de Earl solapado, lo que da como resultado un sistema de ensamblaje que utiliza un cortador LguI de 7 pb (GCTCTTCN^NNN) y un cortador Earl de 6 pb (CTCTTCN^NNN), ambos con extremos adhesivos de 3 pb, por lo que requieren la eliminación de un solo 6 pb. secuencia de partes de ADN (CTCTTC). Sin embargo, debido a que la secuencia de reconocimiento M.TaqI (TCGA) que se solapa con el sitio de restricción Earl interfiere con la secuencia del extremo adhesivo (CTCTTCG^ANN), el sistema impone limitaciones en el diseño de los adaptadores, por lo que solo puede ensamblar fragmentos de tres a la vez. y no se pudo ampliar a un sistema de ensamblaje exclusivo de LguI. Como otro ejemplo, Greengate utilizó ADN conector metilado como partes durante el proceso de ensamblaje del ADN para introducir nuevos sitios Bsal para rondas de ensamblaje posteriores. Sin embargo, el sistema resultante sólo podría usarse para la adición lineal de partes de ADN, limitando así el uso del sistema a la construcción de construcciones de ADN con un pequeño número de unidades de transcripción. En comparación con estos sistemas, el sistema Metclo, con el uso de *in vivo* La metilación del sitio de metilación/restricción solapado fuera de la secuencia adaptadora preserva la ventaja del esquema de ensamblaje multiplicativo jerárquico del sistema Moclo y mantiene la libertad de elección de secuencias adaptadoras para el ensamblaje de ADN. En la práctica, se podrían lograr resultados similares agregando ADN conector metilado como primera y última parte durante el proceso de ensamblaje. Sin embargo, esto aumentaría el número de piezas en el ensamblaje en dos, lo que reduce la tasa de éxito del ensamblaje en comparación con el enfoque de metilación de vectores *in vivo* en el sistema Metclo.

En resumen, el sistema Metclo, mediante el uso de metilación *in vivo* para bloquear sitios específicos de restricción de tipo IIS dentro del vector de ensamblaje, permite el ensamblaje jerárquico de grandes fragmentos de ADN utilizando una única enzima de restricción de tipo IIS. El sistema simplifica el método de ensamblaje basado en enzimas de restricción tipo IIS existente y aumenta la flexibilidad del diseño de esquemas de ensamblaje. El sistema Metclo basado en Bsal en particular es compatible con todas las bibliotecas de piezas existentes que utilizan Bsal y podría usarse como un sistema de ensamblaje estándar tipo IIS.

Tabla 1: Prueba de principio de ensamblaje de fragmentos de ADN mediante sistemas Metcloc basados en BsaI, BpiI o Lgui

Tabla 1: Prueba de principio de ensamblaje de fragmentos de ADN mediante sistemas Metcloc basados en BsaI, BpiI o Lgui

Plásmido	Vector	Inserciones					Metilasa	Encima	Tamaño	Ratio de éxito
pMXP_test	pMXK2_ASE	pMXP_FragA	pMXP_FragB	pMXP_FragC	pMXP_FragD		M.Dsp807II	BsaI	908bp	8/8
pMYP_test	pMYK2_ASE	pMYP_FragA	pMYP_FragB	pMYP_FragC	pMYP_FragD		M2.NimelKCSIII	BpiI	1152bp	8/8
pMZP_test	pMZK2_P6T	pMZP_FragA	pMZP_FragB	pMZP_FragC	pMZP_FragD		M.XmiI	Lgui	1148bp	8/8

Tabla 2: Ensamblaje jerárquico de un fragmento de ADN de 218kb mediante el sistema Metcloc basado en BsaI

Plásmido	Vector	Inserciones					Tamaño			Ratio de éxito
pMXBG_3A1	pMXBG_A7I	pMOLK_2A1	pMOLK_2B1	pMOLK_2C1	pMOLK_2D1	pMOLK_2E1	pMOLK_2F1	pMOLK_2G1	54kb	3/4
pMXBG_3I1	pMXBG_I7G	pMOLK_2I1	pMOLK_2A2	pMOLK_2B2	pMOLK_2C2	pMOLK_2D2	pMOBK_2E2	pMOLK_2F2	55kb	2/4
pMXBG_3G1	pMXBG_G7F	pMOLK_2G2	pMOLK_2I2	pMOLK_2A3	pMOLK_2B3	pMOLK_2C3	pMOBK_2D3	pMOLK_2E3	54kb	2/4
pMXBG_3F1	pMXBG_F7E	pMOLK_2F3	pMOLK_2G3	pMOLK_2I3	pMOLK_2A4	pMOLK_2B4	pMOLK_2C4	pMOLK_2D4	54kb	2/4
pMXBK_4A1	pMXBK_A7E	pMXBG_3A1	pMXBG_3I1	pMXBG_3G1	pMXBG_3F1				218kb	3/6*

* 14/27 positivo PCR, de los cuales 3/8 verificados por digestión de restricción

Ejemplo de un elemento compuesto insercional

En otro escenario, se podrían incluir secuencias de ADN más largas entre el sitio de restricción protegible por metilación y el sitio de restricción no protegible opuesto en el vector de ensamblaje, de modo que después del ensamblaje del ADN, estas secuencias se agregarán en el extremo 5' o 3' del plásmido de ADN ensamblado. Este diseño es conveniente para agregar elementos comunes al plásmido ensamblado mediante el uso de vectores especializados.

En la figura 3 se muestra un ejemplo para el ensamblaje de una unidad de transcripción para la expresión de una proteína con tres dominios diferentes, a partir de plásmidos insertados que contienen los tres dominios individuales D1, D2 y D3 usando Metcloc basado en BsaI. La Figura A muestra el ensamblaje de Metcloc utilizando un vector Metcloc especial que contiene un promotor entre el sitio de restricción protegible por metilación y el sitio de restricción no protegible en 5' de LacZ (secuencia de descarte), y un elemento poliA en 3' ("pre- vector incrustado". El uso de Metcloc con este vector especial preincrustado permite el ensamblaje de los tres dominios en una unidad de transcripción en una reacción en un solo recipiente a partir de tres plásmidos insertados, y la unidad de transcripción ensamblada que consta de un promotor, una región codificante (D1, D2 y D3), y se podría liberar una señal poliA del plásmido ensamblado para la siguiente ronda de ensamblaje de Metcloc. Esto contrasta con el Metcloc estándar que utiliza "elemento compuesto de mantenimiento", mientras que la unidad de transcripción debe ensamblarse a partir de 5 plásmidos de inserción diferentes (Figura B). En las figuras XXXX y YYYY representan la secuencia adaptadora en los extremos 5' y 3' del fragmento de ADN ensamblado (unidad de transcripción) que podría liberarse para la siguiente ronda de ensamblaje de ADN. Para el ensamblaje de ADN utilizando "vectores preincrustados", estas dos secuencias adaptadoras pueden ser las mismas que las secuencias adaptadoras internas porque no están involucradas en el ensamblaje de tres fragmentos de un recipiente. Sin embargo, para el ensamblaje de ADN utilizando el vector Metcloc estándar, las secuencias adaptadoras XXXX y YYYY deben ser diferentes de las secuencias adaptadoras internas (AATG, ACTA, ATCT y GCTT). Además, el método que utiliza vectores preincrustados (Figura A) tiene la ventaja de que es necesario ensamblar una menor cantidad de partes de ADN en una sola reacción, lo que aumenta la tasa de éxito. La desventaja es que el vector especializado es más difícil de preparar.

Ejemplo 2 - Ensamblaje de ADN sin cicatrices (Scarless Metcloc)

Una aplicación única de Metcloc es un sistema para el ensamblaje jerárquico sin cicatrices de construcciones de ADN grandes utilizando un conjunto limitado de vectores de clonación universales. El ensamblaje de ADN basado en tipo

IIS que utiliza vectores universales se basa en una disposición específica de sitios de restricción de tipo IIS cabeza a cabeza en el vector de ensamblaje que cambia el extremo adhesivo del fragmento clonado. El ensamblaje sin cicatrices se logra mediante el uso de estos vectores universales para generar fragmentos de ADN ensamblados con extremos adhesivos (cicatrices) definidos por el fragmento de ADN para la ronda posterior de ensamblaje, que enterró la cicatriz del ensamblaje en la secuencia ensamblada. El uso del sistema Metcloc permite el mantenimiento de los mismos extremos adhesivos universales para los fragmentos exteriores durante las diferentes etapas del ensamblaje jerárquico. Una combinación de estas tres características da como resultado un sistema de ensamblaje universal y sin cicatrices.

Aquí describimos los principios del sistema de ensamblaje Metcloc sin cicatrices utilizando la enzima de restricción Bsal. El sistema se basa en un conjunto de tres tipos de vectores de clonación universales que cambian los extremos adhesivos de los fragmentos clonados (Figura 4A). Los tres tipos de vectores se denominan VL, VM y VR. Los vectores están diseñados para clonar fragmentos de ADN con el extremo adhesivo CTCC en el extremo 5' y CGAG en el extremo 3' que podrían clonarse en los vectores anteriores. El fragmento clonado, una vez liberado por Bsal del vector, tendrá diferentes extremos adhesivos según el vector utilizado. Vector VL está diseñado para mantener el extremo adhesivo 5' y cambiar el extremo adhesivo 3' a las cuatro bases dentro de la secuencia CGAG 3'; Vector VR para mantener el extremo adhesivo 3' y cambiar el extremo adhesivo 5' a las cuatro bases dentro de la secuencia CTCC 5'; Vector VM para cambiar ambos extremos a las cuatro bases del interior. De esta manera se podría utilizar un conjunto de vectores universales para generar fragmentos de ADN con extremos adhesivos definidos por la secuencia del inserto.

El detalle del proceso de cambio del extremo adhesivo se describe en la Figura 4B utilizando el vector VL como ejemplo. Cuando el vector se prepara en células competentes que expresan metilasa, la metilación de adenina (en negrita) debido a la solapamiento de la secuencia de reconocimiento de metilasa da como resultado el bloqueo de los sitios Bsal externos (encuadrados en línea de puntos), cuando los sitios Bsal internos (encuadrados en línea continua) Están disponibles para retirar el inserto LacZ. Tenga en cuenta que el brazo izquierdo del vector VL es similar al Metcloc estándar, y que tanto el sitio Bsal exterior como el interior dan lugar al mismo extremo adhesivo CTCC. Sin embargo, el brazo derecho de VL es diferente en que el extremo adhesivo CGAG generado por el sitio interior de Bsal se solapa con el sitio exterior de Bsal. El ADN del donante lleva un inserto que podría liberarse flanqueando sitios Bsal con extremos adhesivos universales compatibles CTCC y CGAG. El inserto de doble hebra lleva la secuencia XXXX de 4 pb en el extremo 5' y YYYY en el extremo 3' dentro del extremo adhesivo universal. Después del ensamblaje Metcloc del ADN donante en el vector VL, el plásmido resultante lleva el inserto correspondiente que podría ser liberado por los sitios Bsal flanqueantes (encuadrados en línea continua) en el vector. El fragmento liberado llevaría un extremo adhesivo 5' CTCC idéntico al usado para la clonación, pero el extremo adhesivo 3' sería YYYY según lo definido por la secuencia de inserción, que puede ser diferente del extremo adhesivo 3' universal CGAG. De esta manera, el vector VL cambia el extremo adhesivo derecho del fragmento clonado del extremo adhesivo universal CGAG al extremo adhesivo definido por secuencia YYYY.

El proceso de conversión utilizando los tres tipos de vectores de clonación universal podría describirse utilizando un sistema de representación simbólico simple (Figura 41C). Aquí usamos la letra u para representar el extremo adhesivo universal 5' CTCC y la letra v para representar el extremo adhesivo universal 3' CGAG. x y y representan la secuencia de 4 pb en los extremos 5' y 3' del fragmento de inserción dentro del extremo adhesivo universal. El inserto (u,v) representa un plásmido que porta fragmentos de inserto flanqueados por sitios Bsal que generan extremos adhesivos u y v. La reacción de ensamblaje con diferentes vectores da como resultado la conversión de diferentes extremos adhesivos.

El mismo conjunto de vectores universales podría usarse para ensamblar múltiples fragmentos con extremos adhesivos compatibles, siempre que el extremo 5' del primer fragmento y el extremo 3' del último fragmento tengan extremos adhesivos universales u y v (Figura 4D). El fragmento ensamblado lleva diferentes extremos adhesivos según el tipo de vector de ensamblaje utilizado. El ensamblaje de fragmentos en el vector VL mantiene el extremo adhesivo universal 5' y convierte el extremo 3' en el 4 pb interno. El ensamblaje en vector VM convierte ambos extremos y el ensamblaje en vector VR convierte el extremo de 5' y mantiene el extremo adhesivo universal de 3'. Tenga en cuenta que el fragmento ensamblado en el vector VL se parece al fragmento A en que sólo lleva el extremo adhesivo universal 5' u; el fragmento ensamblado en el vector VM se parece al fragmento B; y el fragmento ensamblado en el vector VR se parece al fragmento C. Por lo tanto, el proceso es autosimilar, de modo que los fragmentos ensamblados podrían usarse para la siguiente ronda de ensamblaje de ADN usando el mismo principio dependiendo de la posición del fragmento en el orden de ensamblaje. Esto permite el ensamblaje jerárquico sin cicatrices de grandes construcciones de ADN utilizando vectores de clonación universales.

Todo el proceso de ensamblaje jerárquico y sin cicatrices de ADN se describe en la Figura 2, utilizando como ejemplo el ensamblaje de ADN a partir de 9 fragmentos en 2 etapas. Comenzando con una secuencia de ADN libre de sitios Bsal internos, con los 4 pb en el extremo 5' denominado x y en el extremo 3' denominado y, el objetivo es ensamblar el ADN de 9 fragmentos en un único fragmento que, una vez liberado del plásmido ensamblado, porta extremos adhesivos x e y (Figura 5A). La secuencia de ADN se divide primero en 9 fragmentos denominados A-I con secuencias solapadas de 4 pb, y las secuencias solapadas se denominan a-h (Figura 5B). Las posiciones de los solapamientos de 4 pb podrían elegirse libremente, siempre que sean diferentes de los extremos adhesivos universales u y v, y entre sí dentro de cada reacción de subconjunto. Para la configuración actual, los siguientes pares de solapamientos de

4 pb deben ser diferentes dentro de cada par: (a,b), (d,e), (g,h) y (c,f). Se agregaron secuencias adaptadoras apropiadas a los extremos de las secuencias de los fragmentos A-I, y se sintetizó ADN doble mediante síntesis génica o PCR, de modo que después de la digestión de restricción con Bsal, el ADN bicatenario pudiera recortarse en fragmentos A-I que portaban extremos adhesivos universales u y v (Figura 5C). Estos fragmentos de ADN sintético de doble hebra podrían luego clonarse en tipos apropiados de vectores de clonación universal para cambiar el extremo adhesivo del fragmento dependiendo de la posición del fragmento en el ensamblaje de la siguiente etapa (Figura 5D). Los fragmentos A, D y G, que se ubicarán en la posición final 5' en la siguiente ronda de ensamblaje, se clonan en un vector de tipo VL; B, E, H en la posición media en la siguiente ronda se clonan en el tipo de vector VM; C, F, I en la posición final 3' en el tipo vectorial VR. En esta etapa se podrían secuenciar fragmentos clonados para comprobar si hay errores en la síntesis de genes o en la PCR. Estos fragmentos podrían luego usarse para una nueva ronda de ensamblaje y, nuevamente, el tipo de vector de ensamblaje utilizado depende de la posición del fragmento ensamblado en la siguiente ronda de ensamblaje de ADN (Figura 5E). Los fragmentos A, B, C se ensamblaron en el vector VL, porque el fragmento ensamblado ABC se usará como fragmento 5' en la siguiente ronda de ensamblaje; fragmenta D, E, F en el vector VM y fragmenta G, H, I en el vector VR. El paso final es similar de modo que los tres fragmentos ensamblados del paso 2 se ensamblan aún más en el vector VM, que genera el fragmento completamente ensamblado que lleva los extremos adhesivos x y definidos por los extremos de la secuencia inicial (Figura 5F).

Se requiere un conjunto de al menos 6 vectores de clonación universales que representen 3 tipos (VL, VM, VR) y 2 marcadores de selección de antibióticos para el ensamblaje jerárquico. Idealmente, 12 vectores que representan 2 números de copias diferentes son óptimos para hacer frente a diferentes tamaños de fragmentos de ADN en diferentes etapas de ensamblaje. Se podría desarrollar un sistema similar para Metclo basado en Lgcl, pero no para Bpil, porque el sistema sin cicatrices requiere que los adaptadores universales u y v sean diferentes. Para Bsal y SapI, las bases entre la secuencia de reconocimiento de la enzima y el extremo adhesivo generado por la enzima son flexibles, por lo tanto, es posible diseñar diferentes adaptadores u y v. Sin embargo, para Bpil, estas bases son fijas porque ambas bases son necesarias para especifique la secuencia de reconocimiento de metilasa en Metclo basado en Bpil. Además, la base metilada de Metclo basado en Bpil se encuentra en la cadena superior del sitio de restricción. La disposición cabeza a cabeza de los sitios Bpil para el cambio del extremo adhesivo inevitablemente da como resultado que la base metilada sea transportada por el marcador de selección LacZ, no por la columna vertebral del vector de ensamblaje después de la digestión con Bpil, lo que hace imposible el ensamblaje de un recipiente Golden Gate.

Las relaciones entre Metclo sin cicatrices y Metclo

Metclo se inventó para resolver el problema del uso de una única enzima de restricción para el ensamblaje jerárquico de construcciones de ADN grandes basadas en enzimas de restricción tipo IIS. Scarless Metclo es un desarrollo posterior de Metclo que utiliza componentes clave del sistema Metclo (cepas que expresan metilasa, sitio de metilación/restricción solapado) para lograr un ensamblaje jerárquico sin cicatrices de grandes construcciones de ADN utilizando un conjunto de vectores de ensamblaje universal.

Scarless Metclo se basa en una combinación de varios conceptos diferentes de sistema de ensamblaje basado en tipo IIS:

1. Un diseño especial de sitio de restricción IIS tipo cabeza a cabeza para cambio de extremo adhesivo, que se ha utilizado en el sistema Moclo para el diseño de vectores de clonación universal "nivel -1". El diseño vectorial "nivel -1" de Moclo utiliza un sitio de restricción de cabeza a cabeza Bpil-Bsal, que es diferente del diseño de cabeza a cabeza solo Bsal en Metclo, pero el concepto de cambio de extremo adhesivo se basa en el diseño vectorial Moclo "nivel -1".

2. La eliminación de la cicatriz de ensamblaje mediante el diseño de la secuencia de la cicatriz alrededor de la secuencia de ADN ensamblada. Esto no se ha demostrado antes en un ensamblaje jerárquico basado en IIS.

3. El uso del sistema Metclo para mantener las mismas secuencias de extremos adhesivos para los bloques externos de fragmentos de ADN para cada etapa de ensamblaje.

El uso del sistema Metclo es clave para conseguir un ensamblaje jerárquico multietapa utilizando la misma secuencia de extremo adhesivo universal. Los dos primeros conceptos por sí solos son suficientes para el ensamblaje sin cicatrices de una sola etapa de ADN utilizando un vector de ensamblaje universal. Sin embargo, el sistema de ensamblaje sin cicatrices de una sola etapa no podría adaptarse fácilmente para el ensamblaje jerárquico multietapa utilizando el sistema Moclo tradicional, porque la secuencia adaptadora universal del diseño del vector de cambio del extremo adhesivo para diferentes enzimas son diferentes (Bsal sería CTCN y NGAG, y Bpil ACNN y NNGT). Esto significa que para conseguir un ensamblaje jerárquico sin cicatrices utilizando el Moclo basado en dos enzimas, es necesario añadir secuencias adaptadoras repetitivas muy largas a los extremos de los bloques exteriores en función del número de etapas de ensamblaje, que serán "masticadas" después de cada etapa de ensamblaje (Figura 6). Dado que el sistema Metclo utiliza una única enzima para diferentes etapas de ensamblaje, se podría utilizar la misma secuencia adaptadora universal para los bloques exteriores durante diferentes etapas de ensamblaje, lo que simplifica enormemente el diseño del proceso de ensamblaje.

Ejemplo 3 - Combinaciones identificadas de metilasa y enzima de restricción de tipo IIS

La Tabla 3 contiene una lista de combinaciones de metilasa/enzimas de restricción que pueden dar como resultado el bloqueo del sitio de metilación/restricción solapado para todas las enzimas de restricción de tipo IIS conocidas (78 enzimas que representan 19 especificidades diferentes).

- 5 La lista se generó utilizando un script Perl personalizado basado en una lista de enzimas de restricción de tipo IIS conocidas y una lista de metilasa del ADN de tipo II conocidas, ambas de Rebase (PMID 25378308). Ácidos nucleicos res. 2015 43: D298-9).

- 10 Las metilasas enumeradas se limitan a metilasas de sistemas de restricción/modificación de tipo II con secuencias de reconocimiento de al menos 4 pb y para las cuales se conoce la posición de la base metilada. Se excluyen las metilasas del sistema de restricción/modificación tipo IIG que contienen actividad tanto de metilasa como de enzima de restricción en la misma proteína. Además de la secuencia de reconocimiento de la enzima de restricción, se deben requerir bases adicionales para especificar la actividad metilasa hacia el sitio de metilación/restricción solapado. Se elimina cualquier metilasa cuya secuencia de reconocimiento sea idéntica o esté incluida en el sitio de restricción. Además, la secuencia de metilación no debe invadir la secuencia del extremo adhesivo generada por la enzima de restricción de tipo IIS.

Las secuencias de ADN se enumeran utilizando los códigos de ambigüedad de la IUPAC. M representa A o C; R representa A o G; W representa A o T; S representa C o G; Y representa C o T; K representa G o T; V representa A o C o G; H representa A o C o T; D representa A o G o T; B representa C o G o T; N representa A o T o C o G.

- 20 La “enzima de restricción” enumera el nombre de la enzima de restricción tipo IIS. Las enzimas de restricción que reconocen la misma secuencia se agrupan.

El “sitio de restricción” enumera la secuencia de reconocimiento de la enzima de restricción.

La “metilasa” enumera el nombre de la metilasa que podría bloquear el sitio de metilación/restricción solapado. Se agrupan las metilasas que reconocen la misma secuencia y generan el mismo patrón de metilación.

La “secuencia de metilación” enumera la secuencia de reconocimiento de la metilasa.

- 25 La “posición de la base metilada” enumeraba la posición de la base metilada o de la base opuesta en la secuencia de reconocimiento de la metilasa.

El “tipo de metilación” enumera los tipos de metilación de las bases metiladas en el orden de las posiciones de las bases metiladas enumeradas en la columna anterior. m6A representa N6-metiladenina; m5C representa metilcitosina C5; m4C representa N4-metilcitosina; “Desconocido” representa un tipo desconocido de metilación.

- 30 El “sitio de metilación/restricción solapado” enumera la secuencia del sitio de metilación/restricción solapado que puede resultar en el bloqueo de la acción de la enzima de restricción enumerada a través de la metilación por la metilasa enumerada. Para algunos pares de combinaciones de enzima de restricción/metilasa, hay varios sitios de metilación/restricción solapados diferentes para elegir. Las bases entre corchetes en “[]” representan los extremos adhesivos generados por la enzima de restricción.

Tabla 3:

Enzima de restricción	Lugar de restricción	Metilasa	Secuencia de metilación	Posición de la base metilada	Tipo de metilación	Sitio de metilación/restricción solapado
Sbh132I	CCCG	M.AciI	CCCG	1, 3	m5C	CCCGC[NNNN]
Sbh132I	CCCG	M.Ama53I, M.ChaPQ3II, M.Eco1926II, M.SacCI763IV, M.SspCII, M.XorPXB	CCGGCG	1, 6	m5C	CCGGCGCG[NNNN]
Sbh132I	CCCG	M.ApaI	GGGCCC	3, 4	m5C	GGGGCCG[NNNN]
Sbh132I	CCCG	M.AqiI	CYGGCG	1, 6	m5C	CCGGCGNN[NNNN]
Sbh132I	CCCG	M.AspDU12II, M.BstYI, M.CagII, M.MpaI757V, M.MruI279V, M.KsaII, M.XhoII	RGATCY	2, 5	m4C	RGATCGC[NNNN]
Sbh132I	CCCG	M.AspWA102I, M.AvaVI, M.BceII, M.CbeI, M.Ckr177II, M.CmaIM2IV, M.CpeAVII, M.Csp323II, M.PhoI, M.Spy0714II, M.SnaI	GGCC	2, 3	m4C	GGCCG[NNNN]
Sbh132I	CCCG	M.AvaV, M.SbfI	GATC	1, 4	m4C	GATCG[NNNN], CCGATC[NNNN]
Sbh132I	CCCG	M.BamHI	GGATCC	2, 5	m4C	GGATCCG[NNNN]
Sbh132I	CCCG	M.BanII, M.EcoT30I	GGGCG	3, 4	m5C	GGGGCG[NNNN]
Sbh132I	CCCG	M.BanII, M.CpaII, M.SnaI	GGWCC	2, 4	m5C	GGWCCG[NNNN]
Sbh132I	CCCG	M.BceSI	GGWCC	1, 5	m5C	GGWCCG[NNNN], GGWCCG[NNNN]
Sbh132I	CCCG	M.BspI, M.BspIY, M.Gel16401II, M.PdI0503I	CCGG	1, 4	m5C	CCGGCG[NNNN], CCGGCG[NNNN]

Sth132I	CCCG	M.Bgl2196III	CGTACG	2, 5	m4C	CCCGTACG[NNNN]
Sth132I	CCCG	M.BhaII, M.BspRI, M.BsuRI, M.CthVI, M.DthLII, M.HaeIII, M.MthII, M.NgoAI, M.NgoPII, M.Pod1539II	GGCC	2, 3	m5C	GGCGCGNNNN[NNNN]
Sth132I	CCCG	M.BsoBI	CYCGRG	1, 6	m4C	CCCGRGNN[NNNN]
Sth132I	CCCG	M.Bsp14III, M.Bsp3003I, M.Bsp3004I, M.CalB3III, M.CpeAIII, M.Esu35585I, M.FpsIIV, M.Lba2031I, M.Lbo2004I, M.LlaKR2I, M.Lmo1042I, M.Lmo7956II, M.Lsp10393II, M.MspI7I, M.Sau3AI, M.Sba3I, M.Sin395I, M.Spn23FI, M.Sth368I	GATC	1, 4	m5C	GATCCCGNNNN[NNNN], CCCGATCN[NNNN]
Sth132I	CCCG	M.BsuFI, M.CpeAIV, M.MspI	CCGG	1, 4	m5C	CCCGGNNN[NNNN]
Sth132I	CCCG	M.Gac8I	GCNNGC	2, 5	m5C	GGCGGNNN[NNNN]
Sth132I	CCCG	M.Cdi630III	CCSSGG	2, 5	m4C	CCCGGGNN[NNNN], CCCGGGNN[NNNN]
Sth132I	CCCG	M.Cfr9I, M.SmaI, M.Sma36365II, M.Xca85III, M.XcyI, M.XmaI, M.XveII	CCCGGG	2, 5	m4C	CCCGGGNN[NNNN]
Sth132I	CCCG	M.CocII, M.Fsi4947II, M.TdeII, M.Tph12270I	GGNCC	1, 5	m4C	GGNCCCCGNNNN[NNNN], GGNCCCCGNNNN[NNNN], GGCGCGNNNN[NNNN], GGCGGNNC[NNNN]
Sth132I	CCCG	M.CviII	RGCB	2, 3	m5C	CCCGCYN[NNNN], VGGCCGNNNN[NNNN]
Sth132I	CCCG	M.Dac11109II, M.Hfo6591I, M.MhuII, M.MjaV, M.MlaZII, M.MspLW5I, M.RsaI	GTAC	1, 4	m4C	GTACCGNNNN[NNNN], CCCGTACN[NNNN]
Sth132I	CCCG	M.Djo10085II, M.Ssp6803II	GGCC	2, 3	Unkno wn	GGCGCGNNNN[NNNN]
Sth132I	CCCG	M.DsaV, M.Ecl18kI, M.FpsjVI, M.NlaX,	CCNGG	2, 4	m5C	CCCGGNNN[NNNN]

		M.SenPI, M.SsoII					
Sth132I	CCCG	M.Eco29kI, M.NgoAIII, M.SacII	CCGGGG	2, 5	m5C	CCCGCGGN[NNNN]	
Sth132I	CCCG	M.EcoO109I	RGNGCCY	3, 5	m5C	RGNGCCCGNNNN[NNNN]	
Sth132I	CCCG	M.EfaRFII, M.Mae7806II	CCGG	2, 3	m4C	CCCGGNNN[NNNN]	
Sth132I	CCCG	M.FisAW1I, M.Thal, M.Tph12270II	CGCG	1, 4	m4C	CCCGCGNN[NNNN]	
Sth132I	CCCG	M.FnoB1II, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VII, M.HpyfV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspJDRI, M.TmeBII	CCGG	1, 4	m4C	CCCGGNNN[NNNN]	
Sth132I	CCCG	M.FpsII, M.HpaII, M.Nme18II, M.NmcAI	CCGG	2, 3	m5C	CCCGGNNN[NNNN]	
Sth132I	CCCG	M.HgiCI	GGYRCC	2, 5	m5C	GGYRCCCGNNNN[NNNN]	
Sth132I	CCCG	M.Hpy66I, M.Hpy99I	CGWCG	2, 4	m4C	CCCGWCGN[NNNN]	
Sth132I	CCCG	M.Hpy99IV, M.HpyFORFX	CCNNGG	1, 6	m4C	CCCGGGNN[NNNN]. CCGGNGGN[NNNN]	
Sth132I	CCCG	M.Mae7806III, M.PspPI, M.Sau96I	GGNCC	2, 4	m5C	GGNCCCGNNNN[NNNN]. GGCCCGNNNN[NNNN]	
Sth132I	CCCG	M.MjaII	GGNCC	1, 5	m5C	GGNCCCGNNNN[NNNN]. GGNCNCCGNNNN[NNNN]. GGCCCGNNNN[NNNN]. CCGGNGCC[NNNN]	
Sth132I	CCCG	M.MpelI	RGY	2, 3	m5C	RGCCCGNNNN[NNNN]	
Sth132I	CCCG	M.NgoAXV, M.NgoMV	GGNNCC	2, 5	m5C	GGNNCCCGNNNN[NNNN]. GGNCNCCGNNNN[NNNN]	
Sth132I	CCCG	M.NheI	GCTAGC	1, 6	m4C	GCTAGCCCGNNNN[NNNN]	
Sth132I	CCCG	M.NmeDI	VCCGGY	2, 5	m5C	CCCGGYNN[NNNN]	
Sth132I	CCCG	M.Pfa23572II	CGATCG	1, 6	m5C	CCCGATCG[NNNN]	

Sth132I	CCCG	M.Psp737I	CCCCCD	4	m4C	CCCCCGNNNN[NNNN]
Sth132I	CCCG	M.RpaV	GCCCGCC	4	m4C	GCCCGCCNN[NNNN]
Sth132I	CCCG	M.SuaII	RGATCY	2, 5	m5C	RGATCCCGNNNN[NNNN]
Sth132I	CCCG	M1.Bag18758II, M2.Bag18758II	CCCGTG	2, 4	Unkno wn	CCCGTGNN[NNNN]
Sth132I	CCCG	M1.BenI, M2.BenI, M.NciI	CCSGG	2, 4	m4C	CCCGGNNN[NNNN]
Sth132I	CCCG	M1.BspACI	CCGC	1	m5C	CCGCGNNN[NNNN]
Sth132I	CCCG	M1.CspNS6II	TCCGCC	5	m5C	TCCGCCCGNNNN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.AauI, M.Lsp6406III	GAGGTC	2, 5	m6A	GAGGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.AbaEs12I, M.Awo1030I, M.Cdu23823I, M.Csa8155II, M.Cun9529I, M.EasL1Dem, M.Eba57Dcm, M.Ecl34399Dcm, M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.EsaSP1II, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFdcm, M.Kpn928Dcm, M.KpnNIH1Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I, M.PmiI, M.SpaD2III, M.SplRp8II, M.SptADcm, M.Tsu2489III	CCWGG	1, 5	m5C	CCWGGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.AbrSp7I, M.Apa101655II, M.Ath21654I, M.BdiNKK6I, M.Bel76I, M.Bsp460I, M.BspJG11I, M.BspLA11I, M.BspRAC04II,	GANTC	2, 4	m6A	GAGTCTCN[NNNN]

		M.CagII, M.CcrMI, M.CcrNAI, M.CnaPQ2I, M.Csp16704III, M.CspNS6I, M.CviBI, M.CviQVI, M.Eli8509I, M.Fnu419I, M.Fsp71I, M.HbaI, M.HhaII, M.HinfI, M.HinHiaII, M.Hpy298IV, M.Hpy299IV, M.Hpy32IV, M.Hpy57IV, M.Hpy66IV, M.Hpy99IX, M.HpyAIV, M.HpyFIV, M.HpyUM032IV, M.HpyUM037IV, M.Mbo45III, M.Msp10I, M.Msp73BI, M.MspAL11I, M.MspLW5II, M.NhaXII, M.NpeUS61I, M.OspHL35I, M.Pin17395II, M.PlaSIII, M.PspLM6I, M.RbaNRL2I, M.Rgi1451I, M.Rmo1926I, M.RpaIII, M.RsaI, M.RshII, M.Rsp8I, M.SmeC3I, M.SmoLV, M.Sna24252I, M.SseLI, M.SspM41I, M.SstE37II, M.ThaIII, M.Xsp91I, M.ZmoCP4I				
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.AciI	GCGG	2, 4	m5C	GCGGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.AflIII	ACRYGT	2, 5	m4C	ACRYGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.AhdI	GACNNNN NGTC	2, 10	m6A	GACNNNNGTCTCN[NNNN]
Alw26I, BcoDI	GTCTC	M.Ama55I, M.CnaPQ2III, M.Rmo1926II	CGGCCG	1, 6	m5C	CGGCCGTCTCN[NNNN]

BsmAI, BsoMAI, BstMAI		M.SenC1763IV, M.SspGI1, M.XorPXII					
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Aql	CYCGRG	1, 6	m5C	CYCGRGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.AvaIX, M.Chr10I, M.Nme18VI, M.VchAI, M.Vch0395I	RCCGGY	2, 5	m5C	RCCGGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.BcoSlacII, M.MspLW5III, M.Nbr128I, M.Psy18884II, M.Tam77409II, M.TbrGE1II, M.TpaRLII	CTCGAG	1, 6	m4C	CTCGAGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Bepl, M.PpsIV, M.Gel16401III, M.Pdi8503I	CGCG	1, 4	m5C	CGCGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.BsoBI	CYCGRG	1, 6	m4C	CYCGRGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.BsuFI, M.CpeAIV, M.Mspl	CCGG	1, 4	m5C	CCGGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Cab13497I, M.Dac11109III, M.Dco13527I, M.Djo10085I, M.Hbo11551I, M.Hka19301III, M.Hme33500I, M.HsaRII, M.Hsi33800I, M.HspKM1VII, M.HspNII, M.Htul, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.Mjal, M.MthZI, M.Nas12278I, M.Rf13010I,	CTAG	1, 4	m4C	CTAGTCTCN[NNNN]	

Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.SsmMI M.CagVII	ACGCGT	2, 5	m4C	ACGCGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.CfrBI, M.Sen1735IV, M.SenAnaIV	CCWWGG	1, 6	m4C	CCWWGGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Cmu51329I, M.RorR1Dcm, M.Sen8391Dcm, M.ZalSM2II	CCWGG	1, 5	Unkno wn	CCWGGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Ddel	CTNAG	1, 5	m5C	CTNAGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Esp3I	CGTCTC	4, 3	m5C, m6A	CGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.FatI	CATG	1, 4	m5C	CATGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.FisAW1I, M.Thal, M.Tph12270II	CGCG	1, 4	m4C	CGCGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.FnoB1II, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VIII, M.HpyFV, M.HpyUM032IX, M.HpyUM037IX,	CGCG	1, 4	m4C	CGCGTCTCN[NNNN]	

Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.MjaVI, M.PspIDRI, M.TmeBII M.GsuRedIIV, M.Osp807II, M.Tth111I	GACNNNG TC	2, 8	m6A	GACNNNGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Hpy298II, M.Hpy299II, M.Hpy66II, M.HpyUM032II, M.HpyUM037II	ACNGT	2, 4	m4C	ACNGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.MhuI	CTNAG	1, 5	m4C	CTNAGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Hpy99IV, M.HpyFORFX	CCNNGG	1, 6	m4C	CCNNGGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Hpy99XI	ACGT	2, 3	m5C	ACGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.HpyD27I	GAGG	2, 4	m6A, m5C	GAGGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Lph145I, M.Lph93I	TCWCC	1	m6A	GTCTCC[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.MlyI, M.Pen113I	GASTC	2, 4	m6A	GAGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Msp1II	CCAGG	5	m4C	CCAGGTCTCN[NNNN]

BsmAI, BsoMAI, BstMAI		GTCTC				CCWGG	1, 5	m4C	CCWGGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI		GTCTC		M.Msp42II						
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI		GTCTC		M.NmeDI		VCCGGY	2, 5	m5C	VCCGGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI		GTCTC		M.NmeMC58III		TCTGG	5	m5C	TCTGGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI		GTCTC		M.Nsp209III, M.Nsp51109II		CTCGAG	1, 6	m5C	CTCGAGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI		GTCTC		M.NspI, M.NspHI		RCATGY	2, 5	m5C	RCATGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI		GTCTC		M.Pam7686I		CCATGG	1, 6	m5C	CCATGGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI		GTCTC		M.Pfa23572II		CGATCG	1, 6	m5C	CGATCGTCTCN[NNNN]	
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI		GTCTC		M.Pla5I		CTAG	1, 4	Unkno wn	CTAGTCTCN[NNNN]	

Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Sam53668II		CCATGG	1, 6	Unkno wn	CCATGGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.Sen1781III		CCWWGG	1, 6	Unkno wn	CCWWGGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.SonIII		GCANNNN GTC	3, 9	m6A	GCANNNNGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.TdelI		GAAGAG	5, 6	m6A, m4C	GAAGAGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M.TspRI		CASTG	1, 5	m5C	CASTGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M1.AgoG2I, M2.AgoG2I		CCAGT	1, 4	m4C	ACTGGTCTCN[NNNN], CCAGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M1.BceSIII		GCCGT	4	m4C	GCCGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M1.BspACI		GCGG	4	m5C	GCGGTCTCN[NNNN]
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M1.Dsh12II, M2.Dsh12II		GCCGTC	3, 4	m4C	GCCGTCTCN[NNNN]

BstMAI	GTCTC		GGTCTC						
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M1.Eco31I, M1.EcoMI	GGTCTC	3	m6A	GGTCTCN[NNNN]			
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M1.Hpy32II	TCYYC	1	m6A	GTCTCC[NNNN]			
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M1.LpIAG30I	GAAGG	5	m5C	GAAGGTCTCN[NNNN]			
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M2.BeeSII, M1.Pru9I, M2.Pru9I	GCCGT	2, 4	m4C	GCCGTCTCN[NNNN]			
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M2.BfiI	CCCAGT	5	m4C	CCCAGTCTCN[NNNN]			
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M2.BfuAI	GCAGGT	5	m5C	GCAGGTCTCN[NNNN]			
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M2.Eco31I	GGTCTC	4	m5C	GGTCTCN[NNNN]			
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M2.HgaI, M1.LlaII	GGGTC	3	m5C	GGGTCTCN[NNNN]			
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M2.HpyAVI, M1.MnlI	GAGG	4	m5C	GAGGTCTCN[NNNN]			

BsmAI, BsoMAI, BstMAI									
Alw26I, BcoDI, BsmAI, BsoMAI, BstMAI	GTCTC	M2.Nme579IV, M2.NmeMC58II	GGTTC	4	m6A	GGTCTCTCN[NNNN]			
AlwXI, BbvI, BseKI, Bsp423I, Bst12I, Bst71I, BstV1I, GeoICI, Lsp1109I	GCAG C	M.AbaBs12I, M.Awo1030I, M.Cdu23823I, M.Csa8155II, M.Cun9529I, M.EasL1Dcm, M.Eba57Dcm, M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.EsaSPIII, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFDcm, M.Kpn928Dcm, M.KpnNIH1Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I, M.Pmi1I, M.SpaD2III, M.SpiRp8II, M.SptADcm, M.Tsu2489III	CCWGG	1, 5	m5C	CCWGGCAGCNCNNNNNNNN[NNNN], GCAGCCWGGNNNNN[NNNN]			
AlwXI, BbvI, BseKI, Bsp423I, Bst12I, Bst71I, BstV1I, GeoICI, Lsp1109I	GCAG C	M.Acc65VI, M.SenAZIII	CTKMAG	2, 5	m6A	CTGCAGCNCNNNNNNNN[NNNN]			
AlwXI, BbvI, BseKI, Bsp423I, Bst12I, Bst71I, BstV1I, GeoICI, Lsp1109I	GCAG C	M.AciI	CCGC	1, 3	m5C	GCGCAGCNCNNNNNNNN[NNNN], GCAGCGGNNNNNN[NNNN], CCGCAGCNCNNNNNNNN[NNNN], GCAGCCGCNCNNNN[NNNN]			

GeolCI, Lsp1109I	GCAG	M.Ade56712I, M.Asp588III, M.AveCB51II, M.BbrUIII, M.Blo30698I, M.BsuBI, M.CpaYL7III, M.Dba7044I, M.Dfa45958II, M.Eco12581I, M.Eco2012I, M.EcoGIII, M.EcoG089I, M.Eho16691II, M.Hal5920II, M.Lsp6406II, M.MspNR41II, M.Pae490I, M.Pox23570I, M.PstI, M.RpaII, M.Sca3403II, M.Tsp61I, M.XmnII, M.YenWI, M.YinY228I	CTGCAG	2, 5	m6A	CTGCAGCNNNNNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst7II, BstVII, GeolCI, Lsp1109I	GCAG C					
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst7II, BstVII, GeolCI, Lsp1109I	GCAG C	M.AluBI	AGCT	1, 4	m6A	GCAGCTNNNNNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst7II, BstVII, GeolCI, Lsp1109I	GCAG C	M.AluI, M.CspCY2I	AGCT	2, 3	m5C	GCAGCTNNNNNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst7II, BstVII, GeolCI, Lsp1109I	GCAG C	M.Ama55I, M.CnaPQ2III, M.Rmo1926II, M.SenC1765IV, M.SspG1I, M.XorPXII	CGGCGG	1, 6	m5C	CGGCGGCGACNNNNNNNN[NNNN], GCAGCGGCGCGNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst7II, BstVII, GeolCI, Lsp1109I	GCAG C	M.AplI	CTGCAG	3, 4	m5C	CTGCAGCNNNNNNNN[NNNN]

Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.AquI		CYCGRG	1, 6	m5C	CYCGRGCAGCNNNNNNNN[NNNN], GCAGCYCGRGNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.AsiSI		GGGATCGC	2, 7	m5C	GGGATCGCAGCNNNNNNNN[NNNN], GCAGGGATCGCNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.AvaIX, M.Cfr10I, M.Nme18VI, M.VchAI, M.Vch0395I		RCCGGY	2, 5	m5C	RCCGGCAGCNNNNNNNN[NNNN], GCAGCCGGYNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.BcoSIacII, M.MspLW5III, M.Nbr128I, M.Psy18884II, M.Tam77409II, M.TbrGE1II, M.TpaRLII		CTCGAG	1, 6	m4C	CTCGAGCAGCNNNNNNNN[NNNN], GCAGCTCGAGNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.BepI, M.PpsIV, M.Gel16401III, M.Pdi8503I		GGCG	1, 4	m5C	GGCGCAGCNNNNNNNN[NNNN], GCAGGGCGNNNN[NNNN]

Bst71I, BstV1I, GeoICl, Lsp1109I	GCAG	M.BglI, M.CagI, M.DdeVCDI, M.GsuRedII	GCNNNN NGGC	2, 10	m4C	GCNNNNNGGCAGCNNNNNNNN[NNNN]
AlwXI, BbvI, BseXI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeoICl, Lsp1109I	GCAG	M.Ble402III, M.Psp10HI	SAGCTS	3, 4	m4C	GCAGCTSNNNNN[NNNN]
AlwXI, BbvI, BseXI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeoICl, Lsp1109I	GCAG	M.BsoBI	CYCGRG	1, 6	m4C	CYCGRGCAGCNNNNNNNN[NNNN], GCAGCYCGRGNN[NNNN]
AlwXI, BbvI, BseXI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeoICl, Lsp1109I	GCAG	M.BsuFI, M.CpeAIV, M.MspI	CCGG	1, 4	m5C	CCGGCAGCNNNNNNNN[NNNN], GCAGCCGGNNNN[NNNN]
AlwXI, BbvI, BseXI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeoICl, Lsp1109I	GCAG	M.CabI3497I, M.Dac11109III, M.Dco13527I, M.Djo10085I, M.Hbo11551I, M.Hka19301III, M.Hme33500I, M.HsaR1I, M.Hst33800I, M.HspKM1VII, M.HspNII,	CTAG	1, 4	m4C	CTAGCAGCNNNNNNNN[NNNN], GCAGCTAGNNNN[NNNN]

GeolCI, Lsp1109I		M.Huul, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.Mjal, M.MthZI, M.Nas12278I, M.Rfi3010I, M.SsmMI				
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Cac8I	GCNNGC	2, 5	m5C	GCNNGCAGCNCNNNNNN[NNNN], GCAGCNGCNCNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.CorNAV	YGCCGGGR	3, 6	m5C	YGCCGGCAGCNCNNNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.CfrBI, M.Sen1735IV, M.SenAnaIV	CCWWGG	1, 6	m4C	CCWWGGCAGCNCNNNNNN[NNNN], GCAGCCWWGGNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Cmu51329I, M.EcoIA03PDcm, M.RorR1Dcm, M.Sen8391Dcm, M.ZalSM2II	CCWGG	1, 5	Unkno wn	CCWGGCAGCNCNNNNNN[NNNN], GCAGCCWGGNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.CstfSMII, M.MfoC'I6IV, M.NgoAIV, M.NgoMIV, M1.Sgr13350II	GCCGGC	2, 5	m5C	GCCGGCAGCNCNNNNNN[NNNN], GCAGCCGGCNCNN[NNNN]

Bst71I, BstV41I, GeolCI, Lsp1109I							
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.CviII		RGB	2, 3	m5C	GCAGCBNNNNNN[NNNN]
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.CviRI, M.Hpy298VIII, M.Hpy299VIII, M.Hpy30VI, M.Hpy32V, M.Hpy57II, M.Hpy66VIII, M.HpyFVI, M.HpyUM032VIII, M.HpyUM037VIII		TGCA	1, 4	m6A	TGCAGCNNNNNNNN[NNNN]
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.DdeI		CTNAG	1, 5	m5C	CTNAGCAGCNNNNNNNN[NNNN], GCAGCTNAGNNNN[NNNN]
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Dfel, M.Fnu419II, M.Pen113II		CTNNAG	2, 5	m6A	CTGCAGCNNNNNNNN[NNNN]
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.FatI		CATG	1, 4	m5C	CATGCAGCNNNNNNNN[NNNN], GCAGCATGNNNN[NNNN]
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C						

GeolCl, Lsp1109I	GCAG	M.FisAW1I, M.Thal, M.Tph12270II	CGCG	1, 4	m4C	CGCGCAGCNNNNNNNN[NNNN], GCAGCGGNNNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCl, Lsp1109I	GCAG C	M.FnoB1II, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VII, M.HpyFV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspIDRI, M.TmeBII	CCGG	1, 4	m4C	CCGGCAGCNNNNNNNN[NNNN], GCAGCGGNNNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCl, Lsp1109I	GCAG C	M.Fsi4947I, M.Mru1279IV, M.NsoJS138I, M1.Psp320WII, M.PvuII, M.SenA46IV, M.SptAI	CAGCTG	3, 4	m4C	GCAGCTGNNNNNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCl, Lsp1109I	GCAG C	M.Hhal, M.HinP1I, M.Hpy99III, M.HpyAVIII, M.HpyUM032XV	GGCG	2, 3	m5C	GGCGCAGCNNNNNNNN[NNNN], GCAGCGCNNNNNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCl, Lsp1109I	GCAG C	M.Hme33500II	HGCWGCK	3	m4C	MGCWGCAGCNNNNNNNN[NNNN], MGCAGCDNNNNNNNN[NNNN], HGCAGCKNNNNNNNN[NNNN], GCAGCWGCKNNNNNN[NNNN]

AlwXI, BbvI, BseXI, Dsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.MhuI	CTNAG	1, 5	m4C	CTNAGCAGCNNNNNNNN[NNNN], GCAGCTNAGNNNN[NNNN]
AlwXI, BbvI, BseXI, Dsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Hpy57X	CTRYAG	2, 5	m6A	CTGCAGCNNNNNNNN[NNNN]
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Dst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Hpy99IV, M.HpyFORFX	CCNNGG	1, 6	m4C	CCNNGGCAGCNNNNNNNN[NNNN], GCAGCCNNGGNNN[NNNN]
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.HpyD27I	GAGG	2, 4	m6A, m5C	GCAGCCTCNNNN[NNNN], GAGGCAGCNNNNNNNN[NNNN]
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Mma5219II, M.TarMY3I	AGCT	2, 3	m4C	GCAGCTNNNNNNNN[NNNN]
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG	M.MpelI	RGCY	2, 3	m5C	GCAGCYNNNNNNNN[NNNN]

BseKI, BseXI, Bsp423I, Bst12I, Dst71I, DstV1I, GeolCI, Lsp1109I	C						
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Dst71I, DstV1I, GeolCI, Lsp1109I	GCAG C	M.Msp1II	CCAGG	5	m4C	GCAGCTGGNNNN[NNNN], CCAGGCAGCNNNNNNNN[NNNN]	
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Msp42II	CCWGG	1, 5	m4C	CCWGGCAGCNNNNNNNN[NNNN], GCAGCCWGGNNNN[NNNN]	
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.NgoAI	RGCGCY	3, 4	m5C	GCAGCGCYNNNNNN[NNNN]	
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.NheI	GCTAGC	1, 6	m4C	GCTAGCAGCNNNNNNNN[NNNN], GCAGCTAGCNNNN[NNNN]	
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Nme18V	GGCGGC	2, 5	m5C	GGCGGCAGCNNNNNNNN[NNNN], GCAGCGGCNNNN[NNNN]	

Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.NmeDI		VCCGGY	2, 5	m5C	VCCGGCAGCNNNNNNNN[NNNN], GCAGCGGBNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.NmeMC58III		CCAGA	1	m5C	TCTGGCAGCNNNNNNNN[NNNN], GCAGCCAGANNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Nsp209III, M.Nsp51109II		CTCGAG	1, 6	m5C	CTCGAGCAGCNNNNNNNN[NNNN], GCAGCTCGAGNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Nspl, M.NspHI		RCATGY	2, 5	m5C	RCATGCAGCNNNNNNNN[NNNN], GCAGCATGYNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Pae10332II, M2.Pae50071I	M1.Pae50071I,	CAGCTC	3, 4	m4C	GCAGCTCNNNNNN[NNNN]

Bst71I, BstV1I, GeolCI, Lsp1109I		M.Pam7686I		CCATGG	1, 6	m5C	CCATGGCAGCNNNNNNNN[NNNN], GCAGCCATGNNN[NNNN]
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C						
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Pfa23572II		CGATCG	1, 6	m5C	CGATCGCAGCNNNNNNNN[NNNN], GCAGCGATCGNNN[NNNN]
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Pla5I		CTAG	1, 4	Unkno wn	CTAGCAGCNNNNNNNN[NNNN], GCAGCTAGNNNN[NNNN]
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.Sam53568II		CCATGG	1, 6	Unkno wn	CCATGGCAGCNNNNNNNN[NNNN], GCAGCCATGNNN[NNNN]
AlwXI, BbvI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCI, Lsp1109I	GCAG C	M.SbaUII		CAGCTG	2, 5	m6A	GCAGCTGNNNNNN[NNNN]

GeolCl, Lsp1109I	GCAG	M.Sen1781III		CCWWGG	1, 6	Unkno wn	CCWWGGCAGCNNNNNNNN[NNNN], GCAGCCWWGGNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCl, Lsp1109I	GCAG C	M.SonIII		GCANNNN GTC	3, 9	m6A	GCAGCNGTCNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCl, Lsp1109I	GCAG C	M.TdelI		CTCTTC	1, 2	m4C, m6A	GAAGAGCAGCNNNNNNNN[NNNN], GCAGCTCTTCNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCl, Lsp1109I	GCAG C	M.TspRI		CASTG	1, 5	m5C	CASTGCAGCNNNNNNNN[NNNN], GCAGCASTGNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeolCl, Lsp1109I	GCAG C	M.Yen2516I		GCGCGC	2, 5	Unkno wn	GCGCGCAGCNNNNNNNN[NNNN], GCAGCGCGCNNNN[NNNN]

AlwXI, BseKI, Dsp423I, Bst71I, GeolCI, Lsp1109I	BbvI, BseXI, C	GCAG	M1.AgoG2I, M2.AgoG2I	CCAGT	1, 4	m4C	ACTGGCAGCNNNNNNNN[NNNN], GCAGCCAGTNNNN[NNNN]
AlwXI, BseKI, Dsp423I, Bst71I, GeolCI, Lsp1109I	BbvI, BseXI, C	GCAG	M1.BspACI	CCGC	1	m5C	GCGGCAGCNNNNNNNN[NNNN], GCAGCCGCNNNNNN[NNNN]
AlwXI, BseKI, Bsp423I, Dst71I, GeolCI, Lsp1109I	BbvI, BseXI, C	GCAG	M1.HgaI, M2.LlaII	GGCTC	2	m5C	GACGAGCNNNNNNNN[NNNN], GCAGCGTCNNNNNN[NNNN]
AlwXI, BseKI, Bsp423I, Bst71I, GeolCI, Lsp1109I	BbvI, BseXI, C	GCAG	M1.LplAG30I	GAAGG	5	m5C	GCAGCCTTCNNNN[NNNN], GAAGGCAGCNNNNNNNN[NNNN]
AlwXI, BseKI, Bsp423I, Bst71I, GeolCI, Lsp1109I	BbvI, BseXI, C	GCAG	M1.Sgr13350III, M2.Sgr13350III	GCTCTTC	2, 3	m4C, m6A	GAAGAGCAGCNNNNNNNN[NNNN], GCAGCTTTCNNNN[NNNN]
AlwXI, BbvI,		GCAG	M2.BceSIII, M1.Pru9I, M2.Pru9I	ACGGC	2, 4	m4C	GCAGCCGTNNNNNN[NNNN],

BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeoICI, Lsp1109I	C					ACGGCAGCNCNNNNNNNN[NNNN]
AlwXI, BbvI, BseKI, BseXI, Bsp423I, Bst12I, Bst71I, BstV1I, GeoICI, Lsp1109I	GCAG C	M2.HpyAVI, M1.MnII	GAGG	4	m5C	GCAGCCTCNCNNNN[NNNN], GAGGAGCNCNNNNNNNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.AbaBs12I, M.Awo1030I, M.Cdu23823I, M.Csa8155II, M.Cun9529I, M.Eas11Dcm, M.Eba57Dcm, M.Ecl34399Dcm, M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.EsaSP1II, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFDcm, M.Kpn928Dcm, M.KpnNIH1Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I, M.Pmi1I, M.SpaD2III, M.SpiRp8II, M.SptADcm, M.Tsu2489III	CCWGG	1, 5	m5C	CCWGGCATCNCNNNN[NNNN], GCATCCWGGN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.AciI	CCGC	1, 3	m5C	GGGGCATCNCNNNN[NNNN], CCGCATCNCNNNN[NNNN], GCATCGGCNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Ama55I, M.CnaPQ2III, M.Rmo1926II, M.SenC1765IV, M.SspG1I, M.XorPXII	CGGCCG	1, 6	m5C	CGGGCCGATCNCNNNN[NNNN], GCATCGGGCCG[NNNN]

Bmsl, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.Aqui	CYCGRG	1, 6	m5C	CYCGRGCATCNNNNN[NNNN], GCATCYCGRG[NNNN]
Bmsl, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.AsiSI	GGGATCGC	2, 7	m5C	GGGATCGCATCNNNNN[NNNN]
Bmsl, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.AvalX, M.Cfr10I, M.Nme18VI, M.VchAI, M.VchO395I	RCCGGY	2, 5	m5C	RCCGGCATCNNNNN[NNNN]
Bmsl, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.BamWS8I, M.CthV, M.EfaRFL, M.Saf8902II, M.TacII, M.TcoKWC4III	GCWGC	2, 4	m5C	GCWGCATCNNNNN[NNNN]
Bmsl, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.Bbvl, M.BceSIV	GCAGC	2, 4	m5C	GCTGCATCNNNNN[NNNN], GCAGCATCNNNNN[NNNN]
Bmsl, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.BcoSlacII, M.MspLW5III, M.Nbr128I, M.Psy18884II, M.Tam77409II, M.TbrGE1II, M.TpaRLII	CTCGAG	1, 6	m4C	CTCGAGCATCNNNNN[NNNN], GCATCTCGAG[NNNN]
Bmsl, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.Bepl, M.FpsIV, M.GeI1640I, M. M.Pdi8503I	CGCG	1, 4	m5C	CGCGCATCNNNNN[NNNN], GCATCGCGNN[NNNN]
Bmsl, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.BfaSIII, M.Ccl10114I, M.CfrB38I, M.CfrH1I, M.EcoGVI, M.EcoKII, M.Fnu10562II, M.SbaUIV, M.Sbo12419III, M.Sbo268II, M.Sen10384II, M.Sen10708II, M.Sen1080II, M.Sen1175III, M.Sen13IV, M.Sen13311IV, M.Sen1387II, M.Sen1427II, M.Sen14882III, M.Sen15791III, M.Sen158III, M.Sen16I, M.Sen1655III, M.Sen1676III, M.Sen1677III, M.Sen1728II, M.Sen1735III, M.Sen1736IV, M.Sen1764II, M.Sen1766III, M.Sen1781IV, M.Sen1783III,	ATGCAT	2, 5	m6A	ATGCATCNNNNN[NNNN]

		M.Sen18569II, M.Sen1878III, M.Sen1880III, M.Sen1896III, M.Sen1898III, M.Sen1899III, M.Sen1903III, M.Sen1906III, M.Sen1908III, M.Sen1910III, M.Sen1921III, M.Sen1927II, M.Sen195II, M.Sen2030II, M.Sen2064III, M.Sen2069III, M.Sen22462I, M.Sen255II, M.Sen318III, M.Sen3387III, M.Sen3890III, M.Sen483III, M.Sen624III, M.Sen641III, M.Sen7378II, M.Sen8391III, M.Sen8692III, M.SenA46III, M.SenAbaII, M.SenAboII, M.SenAnaIII, M.SenC1736III, M.SenC1765III, M.SenC1808IV, M.SenC1810III, M.SenJIII, M.SenL1II, M.SenL15III, M.SenN1660II, M.SenSARA26II, M.SenSPBIII, M.SenTFIII				
BmsI, BspST5I, LweI, PhaI, SfaNI	GCAT C	M.BglII, M.CagI, M.DdeVCDI, M.GsuRedII	GCCNNNN NGGC	2, 10	m4C	GCCNNNNNGGCATCNNNNN[NNNN]
BmsI, BspST5I, LweI, PhaI, SfaNI	GCAT C	M.BseCI, M.ClaI, M.MfoCT6III, M.Nfa11134I, M.Nsp209I, M.Nsp51109I, M.Pfu17724II, M.SonI	ATCGAT	2, 5	m6A	GCATCGATNN[NNNN]
BmsI, BspST5I, LweI, PhaI, SfaNI	GCAT C	M.BsoBI	CYCGRG	1, 6	m4C	CYCGRGCATCNNNNN[NNNN], GCATCYCGRG[NNNN]
BmsI, BspST5I, LweI, PhaI, SfaNI	GCAT C	M.BsuFI, M.CpeAIV, M.MspI	CCGG	1, 4	m5C	CCGGCATCNNNNN[NNNN], GCATCCGGNN[NNNN]
BmsI, BspST5I, LweI, PhaI, SfaNI	GCAT C	M.Cab13497I, M.Dac11109III, M.Dco13527I, M.Djo10085I, M.Hbo11551I,	CTAG	1, 4	m4C	CTAGCATCNNNNN[NNNN], GCATCTAGNN[NNNN]

			M.Hka19301III, M.Hmc33500I, M.HsaR1I, M.Hsi33800I, M.HspKM1WII, M.HspNII, M.Htul, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.Mjal, M.MthZI, M.Nas12278I, M.RII3010I, M.SsmMI					
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.Cac8I		GCNNGC	2, 5	m5C	GCNNGCATCNNNNN[NNNN]	
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.CagVI		TCCGGA	1, 6	m6A	GCATCCGGAN[NNNN]	
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.CcrNAV		YGCCGGGR	3, 6	m5C	YGCCGGCATCNNNNN[NNNN]	
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.Cdi630IV, M.Ckr177III, M.CmaLM2II, M.CocII, M.Fnu4HI, M.Fsp4HI		GCNGC	2, 4	m5C	GCNGCATCNNNNN[NNNN]	
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.CfrBI, M.Sen1735IV, M.SenAnaIV		CCWWGG	1, 6	m4C	CCWWGGCATCNNNNN[NNNN], GCATCCWWGG[NNNN]	
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.Cmu51329I, M.EcoIA03PDcm, M.RorR1Dcm, M.Sen8391Dcm, M.ZaiSM2II		CCWGG	1, 5	Unkno wn	CCWGGCATCNNNNN[NNNN], GCATCCWGGN[NNNN]	
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.CsiFSMII, M.MfoCT6IV, M.NgoAIV, M.NgoMIV, M1.Sgr13350II		GCCGGC	2, 5	m5C	GCCGGCATCNNNNN[NNNN]	
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.CviQIII, M.CviSIII, M.Hpy299XI, M.Hpy30II, M.Hpy32X, M.Hpy99VII, M.HpyAX, M.HpyFII, M.HpyUM032XVII, M.Tam77409III, M.TaqI, M.TpaRLI, M.TthHB8I		TCGA	1, 4	m6A	GCATCGANN[NNNN]	
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M.CviRI, M.Hpy298VIII, M.Hpy299VIII, M.IIpy30VI, M.Hpy32V, M.Hpy57II,		TCGA	1, 4	m6A	TGCATCNNNNN[NNNN]	

BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Hpy66VIII, M.HpyFVI, M.HpyUM032VIII, M.HpyUM037VIII	CTNAG	1, 5	m5C	CTNAGCATCNNNNN[NNNN], GCATCTNAGN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.FatI	CATG	1, 4	m5C	CATGCATCNNNNN[NNNN], GCATCATGNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.FisAW1I, M.Thal, M.Tph12270II	CGCG	1, 4	m4C	CGCGCATCNNNNN[NNNN], GCATCGCGNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.FnoB1II, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VIII, M.HpyFV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspIDRI, M.TneBII	CCGG	1, 4	m4C	CCGGCATCNNNNN[NNNN], GCATCCGGNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Hhal, M.HimP1I, M.Hpy99III, M.HpyAVIII, M.HpyUM032XY	GGGC	2, 3	m5C	GGGCATCNNNNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Hka19301IV	TCGCGA	2, 5	m4C	GCATCGCGAN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Hme33500II	MGCWGGC	5	m4C	MGCWGCATCNNNNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Hpy188I, M.Hpy298IX, M.Hpy299V, M.Hpy66V, M.HpyUM032V	TCNGA	1, 5	m6A	GCATCNGANN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Hpy30VIII, M.Hpy32VIII, M.Hpy57III, M.Hpy66VI, M.Hpy99XVIII, M.HpyFVII, M.HpyPVIII	TCNNGA	1, 6	m6A	GCATCNGAN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.Mhul	CTNAG	1, 5	m4C	CTNAGCATCNNNNN[NNNN], GCATCTNAGN[NNNN]

BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Hpy99IV, M.HpyFORFX		CCNNGG	1, 6	m4C	CCNNGGCATCNNNNN[NNNN], GCATCCNNGG[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.HpyD27I		GAGG	2, 4	m6A, m5C	GCA'CC'YC'NN[NNNN], GAGGCATCNNNNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Lph145I, M.Lph93I		TCWCC	1	m6A	GCATCWCCNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.MspIII		CCAGG	5	m4C	GCATCCTCGN[NNNN], CCAGGCATCNNNNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Msp42II		CCWGG	1, 5	m4C	CCWGGCATCNNNNN[NNNN], GCATCCWGGN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.NheI		GCTAGC	1, 6	m4C	GCTAGCATCNNNNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Nme18III		TCTGG	2, 4	m5C	GCATCTGGNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Nme18V		GCGCGC	2, 5	m5C	GCGCGCATCNNNNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.NmeDI		VCCGGY	2, 5	m5C	VCCGGCATCNNNNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.NmeMC58III		CCAGA	1	m5C	TCTGGCATCNNNNN[NNNN], GCATCCAGAN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Nsp209III, M.Nsp51109II		CTCGAG	1, 6	m5C	CTCGAGCATCNNNNN[NNNN], GCATCTCGAG[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Nspl, M.NspHI		RCATGY	2, 5	m5C	RCATGCATCNNNNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Pam7686I		GCATGG	1, 6	m5C	GCATGGCATCNNNNN[NNNN], GCATCCATGG[NNNN]

BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Pla23572II		CGATCG	1, 6	m5C	CGATCGCATCNNNNN[NNNN], GCATCGATCG[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Pla5I		CTAG	1, 4	Unkno wn	CTAGCATCNNNNN[NNNN], GCATCTAGNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Sam53668II		GCATGG	1, 6	Unkno wn	GCATGGCATCNNNNN[NNNN], GCATCCCATGG[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Sen1781III		CCWWGG	1, 6	Unkno wn	CCWWGGCATCNNNNN[NNNN], GCATCCWWGG[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.SmoLV1, M.Spn23FII, M.Spn7465I, M.SpnRI, M.XbaI	M.Spn556I,	TCTAGA	1, 6	m6A	GCATCTAGAN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.SonII		GCANNNN GTC	3, 9	m6A	GCATCNGTGC[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.TdelI		CTCTTC	1, 2	m4C, m6A	GAAGAGCATCNNNNN[NNNN], GCATCTCTTC[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.TspRI		CATG	1, 5	m5C	CATGGCATCNNNNN[NNNN], GCATCASTGN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M.Yen2516I		GGCCGC	2, 5	Unkno wn	GGCCGCATCNNNNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M1.AgoG2I, M2.AgoG2I		CCAGT	1, 4	m4C	ACTGGCATCNNNNN[NNNN], GCATCCAGTN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M1.BspACI		CCGC	1	m5C	GGGCATCNNNNN[NNNN], GCATCCGCNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M1.BstF5I, M3.BstF5I, M2.Csp12AI		CATCC	3	m6A	GCATCCNNNN[NNNN]
BmsI, BspST5I, LweI, PfaI, SfaNI	GCAT C	M1.HgaI, M2.LlaII		GACGC	4	m5C	GACGCATCNNNNN[NNNN]

BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M1.HphI	TCACC	2	m5C	GCATCACCN[NNNN]
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M1.Hpy32II	TCYYC	1	m6A	GCATCYGCNN[NNNN]
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M1.HpyAII, M1.MboII, M1.NcuI	TCTTC	1	m6A	GCATCTTCNN[NNNN]
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M1.LplAG30I	GAAGG	5	m5C	GCATCCTTCN[NNNN], GAAGGATC[NNNN]
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M1.Sgr13350III, M2.Sgr13350III	GAAGAGC	5, 6	m6A, m4C	GAAGAGCATC[NNNN]
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M2.RceSIII, M1.Pru9I, M2.Pru9I	ACGGC	2, 4	m4C	ACGGCATC[NNNN]
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M2.BstF5I	CATCC	2	m6A	GCATCCNNNN[NNNN]
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M2.ClaSBI, M1.Fli6794IV, M2.HphI, M2.Ngo3502II, M1.Ngo6140II, M1.NgoAXVI, M1.NgoBVIII, M1.NgoFA19I, M2.NgoMVI,II, M2.PspHUN102I, M1.Ssp10II	TCACC	1	m6A	GCATCACCN[NNNN]
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M2.Hpy32II, M2.HpyAII, M2.MboII, M2.NcuI	TCTTC	2	m4C	GCATCTTCNN[NNNN]
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M2.HpyAVI, M1.MnII	GAGG	4	m5C	GCATCCTC[NNNN], GAGGATC[NNNN]
BmsI, BspST5I, Lwel, Phal, SfaNI	GCAT C	M4.BstF5I, M1.Cin755I, M2.Cin755I, M1.CmeSR3I, M2.CmeSR3I, M1.Csp12AI, M1.Csp323I, M2.Csp323I, M1.Csp410I, M2.Csp410I, M.FokI, M1.Msp315I,	CATCC	2, 3	m6A	GCATCCNNNN[NNNN]

BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, GGGA C	M2.Msp315I, M.Sau13435I, M.StsI, M1.Tsu2489IV, M2.Tsu2489IV, M3.Tsu2489IV	GACGTC	2, 5	m6A	GGGACGTCNNNNNNNN[NNNN]
		M.AatII, M.Lsp6406III				
		M.AbaBs12I, M.Awo1030I, M.Gdu23823I, M.Csa8155II, M.Cun9529I, M.EasL1Dcm, M.Fba57Dcm, M.Ecl34399Dcm, M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.EsaSP1II, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFDcm, M.Kpn928Dcm, M.KpnNIH10Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I, M.PmiI, M.SpaD2III, M.SplRp8II, M.SptADcm, M.Tsu2489III				CCWGGGACNNNNNNNNNN[NNNN], CCWGGGACNNNNNNNNNN[NNNN], GGGACCCWGGNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, GGGA C	M.AbrSp7I, M.Apa101655II, M.Ath21654I, M.BdiNK6I, M.Bel76I, M.Bsp460I, M.BspJGK1II, M.BspLA1II, M.BspRAC04II, M.CagII, M.CcrMI, M.CcrNAL, M.CnaPQ2I, M.Csp16704III, M.CspNS6I, M.CviBI, M.CviQVI, M.Eli8509I, M.Fnu419I, M.Fsp71I, M.HbaI, M.HhaII, M.Hinfi, M.HinHiaII,	GANTC	2, 4	m6A	GGGACTCNNNNNNNNNN[NNNN]

		M.Hpy298IV, M.Hpy299IV, M.Hpy32IV, M.Hpy57IV, M.Hpy66IV, M.Hpy99IX, M.HpyAIV, M.HpyFIV, M.HpyUM032IV, M.HpyUM037IV, M.Mbo45III, M.Msp10I, M.Msp73BI, M.MspAL11I, M.MspLW5II, M.NhaXII, M.NpeUS6II, M.OspHL35I, M.Pin17395II, M.Pla5III, M.PspLM6I, M.RbaNRL2I, M.Rgi145II, M.Rmo1926I, M.RpaIII, M.RsaII, M.RshII, M.Rsp8I, M.SmeC3I, M.SmoLV, M.Sna24252I, M.SscL1I, M.SspM41II, M.Ste37II, M.ThalII, M.Xsp91I, M.ZmoCP4I				
BpuSI, BsmFI, RspLU11III, BstOZ616I, FagI	GGGA C	M.AciI	CCGC	1, 3	m5C	GGGGGACNNNNNNNNNN[NNNN], GGGGGACNNNNNNNNNN[NNNN], GGGACCGCNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.AflIII	ACRYGT	2, 5	m4C	GGGACRYGTNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.AhdI	GACNNNN NGTC	2, 10	m6A	GGGACNNNNNGTCNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.Ama55I, M.CnaPQ2III, M.Rmo1926II, M.SenC1765IV, M.SspG1I, M.XorPXII	CGGCCG	1, 6	m5C	CGGGCGGGACNNNNNNNNNN[NNNN], GGGACGGCGCNNNNNN[NNNN]

BspLU11III, DstOZ616I, FagI							
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.Aql	CYCGRG	1, 6	m5C	CYCGRGGGACNNNNNNNNNN[NNNN], CYCGGGGACNNNNNNNNNN[NNNN], CYCGGGGACNNNNNNNNNN[NNNN], GGGACYCGRGNNNNNN[NNNN]	
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.AspU41II, M.BisIII, M.BsuW23II, M.GnaPQ2II, M.Eco06Dcm, M.Eco1258IDcm, M.Eco1655Dcm, M.Eco29787Dcm, M.Eco3024Dcm, M.Eco3317Dcm, M.Eco3325Dcm, M.Eco3740Dcm, M.Eco3911Dcm, M.Eco4211Dcm, M.Eco4465Dcm, M.Eco6000Dcm, M.Eco644Dcm, M.Eco86Dcm, M.Eco9001Dcm, M.Eco9387Dcm, M.EcoE1140Dcm, M.EcoE455Dcm, M.EcoGDcm, M.EcoNU14Dcm, M.EcoRII, M.Kpn223II, M.Kpn35657II, M.Kpn36Dcm, M.Kpn555I, M.Kpn62629II, M.KpnAATV, M.Kqu603Dcm, M.Psp345I, M.Sen158Dcm, M.Sen641Dcm, M.SenA46Dcm, M.SenAboDcm, M.TcoKWC4II	CCWGG	2, 4	m5C	CCWGGGACNNNNNNNNNN[NNNN]	
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.AvaIX, M.Cfr10I, M.Nme18VI, M.VchAI, M.VchO395I	RCCGY	2, 5	m5C	GGGACCGGYNNNNNN[NNNN]	

BpuSI, BsmFI, BspLU11III, BstOZ616I, FaqI	BsiFI, C	GGGA	M.BanLI, M.CpeAI, M.SinI	GgWCC	2, 4	m5C	GGGACNNNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FaqI	BsiFI, C	GGGA	M.BceSI	GgWCC	1, 5	m5C	GGGACNNNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FaqI	BsiFI, C	GGGA	M.BcoSIacII, M.MspLW5III, M.Nbr128I, M.Psy18884II, M.Tam77409II, M.ThrGE1II, M.TpaRLII	CTCGAG	1, 6	m4C	CTCGAGGGACNNNNNNNNNN[NNNN], GGGACTCGAGNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FaqI	BsiFI, C	GGGA	M.BepI, M.FpsIV, M.GelI6401III, M.Pdi8503I	CCCG	1, 4	m5C	CGCGGGACNNNNNNNNNN[NNNN], GGGACGCGNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FaqI	BsiFI, C	GGGA	M.Bhy27164IV, M.BstNI, M.EcoDHB4Dcm, M.EcoK54Dcm, M.Gth3570I, M.Hla49239I, M.MvaI, M.Pdi8503II, M.Sso30807Dcm, M.YpeI, M.Ype41I, M.YpeShI	CCWGG	2, 4	m4C	CCWGGGACNNNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FaqI	BsiFI, C	GGGA	M.DsoBI	CYCGRG	1, 6	m4C	CYCGRGGGACNNNNNNNNNN[NNNN], CYCGGGACNNNNNNNNNN[NNNN], CYCGGGACNNNNNNNNNN[NNNN], GGGACYCGGNNNNNN[NNNN]
BpuSI, BsmFI,	BsiFI, C	GGGA	M.BsuFI, M.CpeAIV, M.MspI	CCGG	1, 4	m5C	CCGGGGACNNNNNNNNNN[NNNN], CCGGGACNNNNNNNNNN[NNNN]

BspLU11III, BstOZ616I, FagI						GGGACCGGNNNNNN[NNNN]
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.CabI3497I, M.Dac11109III, M.Dco13527I, M.Djo10085I, M.Hbo11551I, M.Hka19301III, M.Hme33500I, M.HsaRII, M.Hsi33800I, M.HspKM1WII, M.HspNII, M.HluI, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.MjaI, M.MthZI, M.Nas12278I, M.Rfi3010I, M.SsmMI	CTAG	1, 4	m4C	CTAGGGACNNNNNNNNNN[NNNN], GGGACTAGNNNNNNNN[NNNN]
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.CagVIII	ACGCGT	2, 5	m4C	GGGACGCGTNNNNNN[NNNN]
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.Cdi630III	CCSSGG	2, 5	m4C	CCSSGGGACNNNNNNNNNN[NNNN], CCSGGGACNNNNNNNNNN[NNNN]
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.Cfr9I, M.SmaI, M.Sma36365II, M.Xca85II, M.XcyI, M.XmaI, M.XvuII	CCCGGG	2, 5	m4C	CCCGGGGACNNNNNNNNNN[NNNN], CCCGGGACNNNNNNNNNN[NNNN]
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.CfrBI, M.Sen1735IV, M.SenAnaIV	CCWWGG	1, 6	m4C	CCWWGGGACNNNNNNNNNN[NNNN], CCWWGGGACNNNNNNNNNN[NNNN], GGGACCCWWGGNNNNNN[NNNN]
BpuSI, BsiFI,	GGGA	M.Cmu51329I, M.Eco/A03PDcm,	CCWWGG	1, 5	Unkno	CCWWGGGACNNNNNNNNNN[NNNN],

BsmFI, BspLU11III, BstOZ616I, FagI	C	M.RorR1Dcm, M.Sen8391Dcm, M.ZaiISM2II			wn	CCWGGGACNNNNNNNNNN[NNNN], GGGACCGWGGNNNNNN[NNNN]
BpuSI, BsiFI, BstOZ616I, FagI	GGGA C	M.CocIII, M.Fsi4947II, M.TdeIII, M.Tph12270I	GGNCC	1, 5	m4C	GGGACCCNNNNNNNNNN[NNNN]
BpuSI, BsiFI, BstOZ616I, FagI	GGGA C	M.Cpa1869GIII, M.Ype19I, M.YpeDodI, M.YpeEDI, M.YpeJ9I, M.YpePGI	CCWGG	2, 4	Unkno wn	CCWGGGACNNNNNNNNNN[NNNN]
BpuSI, BsiFI, BstOZ616I, FagI	GGGA C	M.DdeI	CTNAG	1, 5	m5C	CTNAGGGACNNNNNNNNNN[NNNN], GGGACTNAGNNNNNN[NNNN]
BpuSI, BsiFI, BstOZ616I, FagI	GGGA C	M.DsaV, M.Ecl18kI, M.FpsjVI, M.NlaX, M.SenPI, M.SsoII	CCNGG	2, 4	m5C	CCNGGGACNNNNNNNNNN[NNNN], CCGGGACNNNNNNNNNN[NNNN]
BpuSI, BsiFI, BstOZ616I, FagI	GGGA C	M.Eco29kI, M.NgoAIII, M.SacII	CCGCGG	2, 5	m5C	CCGCGGGACNNNNNNNNNN[NNNN]
BpuSI, BsiFI, BstOZ616I, FagI	GGGA C	M.EcoNI	CCTNNNN NAGG	2, 10	m4C	CCTNNNNNAGGGACNNNNNNNNNN[NNNN]]

BstOZ616I, FagI									
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.Eco0109I	RGGNCCY	3, 5	m5C	GGGACCYNNNNNNNNN[NNNN]			
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.EfaRFII, M.Mae7806II	CCGG	2, 3	m4C	CCGGGACNNNNNNNNNN[NNNN]			
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.FatI	CATG	1, 4	m5C	CATGGGACNNNNNNNNNN[NNNN], GGGACATGNNNNNNNN[NNNN]			
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.FisAW1I, M.Thal, M.Tph12270II	CGCG	1, 4	m4C	CGCGGGACNNNNNNNNNN[NNNN], GGGACGCGNNNNNNNN[NNNN]			
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.FnoB1II, M.Hal2458I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VII, M.IIpyFV, M.IIpyUM032IX, M.IIpyUM037IX, M.MjaVI, M.PspIDRI, M.TneBII	CCGG	1, 4	m4C	CCGGGGACNNNNNNNNNN[NNNN], CCGGGACNNNNNNNNNN[NNNN], GGGACCGGNNNNNNNN[NNNN]			
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M.FpsII, M.HpaII, M.NmeI8II, M.NmeAI	CCGG	2, 3	m5C	CCGGGACNNNNNNNNNN[NNNN]			

BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, GGGA C	M.Gel16401IV	CTGG	3	m5C	CTGGGACNNNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, GGGA C	M.GsuRedIIIV, M.Osp807II, M.Tth111I	GACNNNG TC	2, 8	m6A	GGGACNNNGTCNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, GGGA C	M.Hpy298II, M.Hpy299II, M.Hpy66II, M.HpyUM032II, M.HpyUM037II	ACNGT	2, 4	m4C	GGGACNGTNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, GGGA C	M.Hpy30VIII, M.Hpy32VIII, M.Hpy57III, M.Hpy66VI, M.Hpy99XVIII, M.HpyFVII, M.HpyPVIII	TCNNGA	1, 6	m6A	TCGGGACNNNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, GGGA C	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.Mhul	CTNAG	1, 5	m4C	CTNAGGGACNNNNNNNNNN[NNNN], GGGACTNAGNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, GGGA C	M.Hpy99IV, M.HpyFORFX	CCNNGG	1, 6	m4C	CCNNGGGGACNNNNNNNNNN[NNNN], CCNNGGGACNNNNNNNNNN[NNNN], CCNGGGACNNNNNNNNNN[NNNN], GGGACCNNGGNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, GGGA C	M.Hpy99XI	ACGT	2, 3	m5C	GGGACGTNNNNNNNN[NNNN]

BspLU11III, BstOZ616I, FaeI							
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FaeI	GGGA C	M.HpyAV, M.NmeI8IV	GAAGG	3, 4	m6A, m5C	GAAGGGACNNNNNNNN[NNNN]	
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FaeI	GGGA C	M.HpyD27I	GAGG	2, 4	m6A, m5C	GGGACCTCNNNNNNNN[NNNN], GAGGGGACNNNNNNNNNN[NNNN], GAGGGACNNNNNNNNNN[NNNN]	
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FaeI	GGGA C	M.Mac7806III, M.PspPI, M.Sau96I	GGNCC	2, 4	m5C	GGGACCNNNNNNNNNN[NNNN]	
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FaeI	GGGA C	M.MjaII	GGNCC	1, 5	m5C	GGGACCNNNNNNNNNN[NNNN]	
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FaeI	GGGA C	M.MlyI, M.Pen113I	GATC	2, 4	m6A	GGGACTCNNNNNNNNNN[NNNN]	
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FaeI	GGGA C	M.Msp1II	CCAGG	5	m4C	GGGACCTGGNNNNNNNN[NNNN], CCAGGGGACNNNNNNNNNN[NNNN], CCAGGGACNNNNNNNNNN[NNNN]	

BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M.Msp42II		CCWGG	1, 5	m4C	CCWGGGACNNNNNNNNNN[NNNN], CCWGGGACNNNNNNNNNN[NNNN], GGGACCGGNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M.NgoAXV, M.NgoMV		GGNNCC	2, 5	m5C	GGGACNNNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M.Nme18III		TCTGG	2, 4	m5C	TCTGGGACNNNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M.NmeDI		VCCGGY	2, 5	m5C	RCCGGGACNNNNNNNNNN[NNNN], GGGACCGGBNNNNNN[NNNN], RCCGGGACNNNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M.NmeMC58III		CCAGA	1	m5C	TCTGGGACNNNNNNNNNN[NNNN], TCTGGGACNNNNNNNNNN[NNNN], GGGACCGAGNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M.Nsp209III, M.Nsp51109II		CTCGAG	1, 6	m5C	CTCGAGGACNNNNNNNNNN[NNNN], GGGACTCGACNNNNNN[NNNN]
BpuSI, BsmFI,	BsiFI, C	GGGA	M.Nspl, M.NspHI		RCATGY	2, 5	m5C	GGGACATGYNNNNNN[NNNN]

BspLU11III, BstOZ616I, FaeI									
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FaeI	GGGA C	M.Pam7686I		CCATGG	1, 6	m5C	CCATGGGGACNNNNNNNNNN[NNNN], CCATGGGACNNNNNNNNNN[NNNN], GGGACCATGGNNNNNN[NNNN]		
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FaeI	GGGA C	M.Pfa23572II		CGATCG	1, 6	m5C	CGATCGGGACNNNNNNNNNN[NNNN], GGGACGATCGNNNNNN[NNNN]		
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FaeI	GGGA C	M.Pla5I		CTAG	1, 4	Unkno wn	CTAGGGACNNNNNNNNNN[NNNN], GGGACTAGNNNNNNNN[NNNN]		
BspLU11III, BstOZ616I, FaeI									
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FaeI	GGGA C	M.Sam53668II		CCATGG	1, 6	Unkno wn	CCATGGGGACNNNNNNNNNN[NNNN], CCATGGGACNNNNNNNNNN[NNNN], GGGACCATGGNNNNNN[NNNN]		
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FaeI	GGGA C	M.Sen1781III		CCWWGG	1, 6	Unkno wn	CCWWGGGACNNNNNNNNNN[NNNN], CCWWGGGACNNNNNNNNNN[NNNN], GGGACCCWWGGNNNNNN[NNNN]		
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FaeI	GGGA C	M.SonIII		GACNNNN TGC	2, 8	m6A	GGGACNNNNNTGCNNNN[NNNN]		

BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M.TdeII		CTCTTC	1, 2	m4C, m6A	GAAGAGGGACNNNNNNNNNN[NNNN], GGGACTCTTCNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M.TspRI		CASTG	1, 5	m5C	CASTGGGACNNNNNNNNNN[NNNN], GGGACASTGCNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M1.AgoG2I, M2.AgoG2I		CCAGT	1, 4	m4C	ACTGGGGACNNNNNNNNNN[NNNN], ACTGGGACNNNNNNNNNN[NNNN], GGGACTGGNNNNNNNN[NNNN], GGGACCACTNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M1.Bag18758II, M2.Bag18758II		CACGGG	3, 5	Unkno wn	CACGGGGACNNNNNNNNNN[NNNN], CACGGGACNNNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M1.BceSIII		ACGGC	2	m4C	GGGACGGCNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M1.BcnI, M2.BcnI, M.NciI		CCSGG	2, 4	m4C	CCSGGGACNNNNNNNNNN[NNNN], CCGGGACNNNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M1.BfiI		ACTGGG	5	m4C	ACTGGGGACNNNNNNNNNN[NNNN], ACTGGGACNNNNNNNNNN[NNNN]

BspLU11III, BstOZ616I, FagI							
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M1.BspACI		CCGC	1	m5C	GGGGGACNNNNNNNNNN[NNNN], GCGGGACNNNNNNNNNN[NNNN], GGGACCGCNNNNNNNN[NNNN]
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M1.Dsh12II, M2.Dsh12II		GACGGC	3, 4	m4C	GGGACGGCNNNNNNNN[NNNN]
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M1.Hpy32II		GRRGA	5	m6A	GGGGACNNNNNNNNNN[NNNN]
BspLU11III, BstOZ616I, FagI							
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M1.LpIAC30I		GAAGG	5	m5C	GGGACCTTCNNNNNN[NNNN], GAAGGGGACNNNNNNNNNN[NNNN], GAAGGGACNNNNNNNNNN[NNNN]
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M2.BceSIII, M1.Pru9I, M2.Pru9I		ACGGC	2, 4	m4C	GGGACGGCNNNNNNNN[NNNN]
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C	M2.BfiI		ACTGGG	2	m4C	GGGACTGGGNNNNNN[NNNN]
BpuSI, BsiFI, BsmFI, BspLU11III, BstOZ616I, FagI	GGGA C						

BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M2.BfuAI	ACCTGC	2	m5C	GGGACCTGCNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M2.HgaI, M1.LlaII	GAGGC	3	m5C	GGGAGGCNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M2.HpyAVI, M1.MnII	GAGG	4	m5C	GGGACCTCNNNNNN[NNNN], GAGGGACNNNNNNNN[NNNN], GAGGACNNNNNNNNNN[NNNN]
BpuSI, BsmFI, BspLU11III, BstOZ616I, FagI	BsiFI, C	GGGA	M2.Nme579IV, M2.NmeMC58II	GAGGC	2	m6A	GGGAGGCNNNNNNNN[NNNN]
CseI, HgaI	GACG	C	M.AbaBs12I, M.Awo1030I, M.Cdu23823I, M.Csa8155II, M.Cun9529I, M.EasL1Dcm, M.Eba57Dcm, M.Ecl34399Dcm, M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.EsaSPIII, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFDcm, M.Kpn928Dcm, M.KpnNIH1Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I,	CCWGG	1, 5	m5C	CCWGGAGGCNNNNNN[NNNNN], GAGGCCWGGN[NNNNN]

		M.PmiII, M.SpaD2III, M.SplRp8II, M.SptADcm, M.Tsu2489III					
Csel, Hgal	GACG C	M.AclI, M.FpsFPG3II, M.SdeAIV	GTMKAC	2, 5	m6A	GTMGACGCNNNNN[NNNNN]	
Csel, Hgal	GACG C	M.AclI	CCGC	1, 3	m5C	GCGGACGCNNNNN[NNNNN], GACGGGNNN[NNNNN], GACGCCGCNN[NNNNN]	
Csel, Hgal	GACG C	M.AHI	ACRYGT	2, 5	m4C	GACGGTNNN[NNNNN]	
Csel, Hgal	GACG C	M.Alw26I	GAGAC	4, 3	m6A, m5C	GAGACGCNNNNN[NNNNN]	
Csel, Hgal	GACG C	M.Ama55I, M.CnaPQ2III, M.Rmo1926II, M.SenC1765IV, M.SspG1I, M.XorPXII	CGGCCG	1, 6	m5C	CGGCCGACGCNNNNN[NNNNN], GACGGCGCG[NNNNN]	
Csel, Hgal	GACG C	M.AorT14III, M.GsuRedIII, M.SalI	GTCGAC	2, 5	m6A	GTCGACGCNNNNN[NNNNN]	
Csel, Hgal	GACG C	M.Aql	CYGGRG	1, 6	m5C	CYGGRGACGCNNNNN[NNNNN], GACGCYCGRG[NNNNN]	
Csel, Hgal	GACG C	M.AvaIX, M.Cfr10I, M.Nme18VI, M.VchAI, M.VchO395I	RCCGGY	2, 5	m5C	GACGCCGGYN[NNNNN]	
Csel, Hgal	GACG C	M.BamWS8I, M.CthV, M.EfaRFI, M.Saf8902II, M.TacII, M.TcoKWC4III	GCWGC	2, 4	m5C	GACGCWGCNN[NNNNN]	
Csel, Hgal	GACG C	M.BbvI, M.BceSIV	GCAGC	2, 4	m5C	GACGCTGCNN[NNNNN], GACGCAGC[NNNNN]	
Csel, Hgal	GACG C	M.BcoSlaII, M.MspLW5III, M.Nbr128I, M.Psy18884II, M.Tam77409II, M.ThrGF1II, M.TpaRII	CTCGAG	1, 6	m4C	CTCGAGACGCNNNNN[NNNNN], GACGCTCGAG[NNNNN]	

Csel, Hgal	GACG C	M.BepI, M.FpsIV, M.Pdi8503I	M.GelI6401III,	CGCG	1, 4	m5C	GACGCGNNNN[NNNNN], CGCGACGCGNNNN[NNNNN]
Csel, Hgal	GACG C	M.BsoBI		CYCGRG	1, 6	m4C	CYCGRGACGCGNNNN[NNNNN], GACGCGGCG[NNNNN]
Csel, Hgal	GACG C	M.BsuFI, M.CpeAIV, M.MspI		CCGG	1, 4	m5C	CCGGACGCGNNNN[NNNNN], GACGCGGCG[NNNNN]
Csel, Hgal	GACG C	M.CabI3497I, M.DacI11109III, M.DcoI3527I, M.DjoI0085I, M.HboI1551I, M.HkaI9301III, M.Hme33500I, M.HsaR1I, M.Hsi33800I, M.HspKM1VII, M.HspNII, M.Htul, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.Mjal, M.MthZI, M.NasI2278I, M.Rfi3010I, M.SsmMI		CTAG	1, 4	m4C	CTAGACGCGNNNN[NNNNN], GACGCTAGNN[NNNNN]
Csel, Hgal	GACG C	M.Cac8I		GCNNGC	2, 5	m5C	GACGCGNNGC[NNNNN]
Csel, Hgal	GACG C	M.CagVI		TCCGGA	1, 6	m6A	TCCGGACGCGNNNN[NNNNN]
Csel, Hgal	GACG C	M.CagVIII		ACGCGT	2, 5	m4C	GACGCGTNNN[NNNNN]
Csel, Hgal	GACG C	M.CcrNAV		YGCCGGCR	3, 6	m5C	GACGCGGCR[NNNNN]
Csel, Hgal	GACG C	M.Cdi630IV, M.Ckr177III, M.CmaIM2II, M.CocII, M.Fnu4HI, M.Fsp4HI		GCNGC	2, 4	m5C	GACGCGCINN[NNNNN]
Csel, Hgal	GACG C	M.CfrBI, M.SenI735IV, M.SenAnalV		CCWWGG	1, 6	m4C	CCWWGGACGCGNNNN[NNNNN], GACGCGWWGG[NNNNN]
Csel, Hgal	GACG	M.Cmu51329I, M.EcoIA03PDcm,		CCWGG	1, 5	Unkno	CCWGGACGCGNNNN[NNNNN]

	C	M.RorR1Dcm, M.Sen8391Dcm, M.ZaISM2II			wn	GACGCCWGCN[NNNNN]
Csel, Hgal	GACG C	M.CocI, M.Fli6794II, M.HincII, M.HindII, M.HinHiaIII	GTyrAC	2, 5	m6A	GTYGACGCNNNNN[NNNNN]
Csel, Hgal	GACG C	M.CsiFSMII, M.MfoCT6IV, M.NgoAIV, M.NgoMIV, M1.Sgr13350II	GCGGCG	2, 5	m5C	GACGCCGGCN[NNNNN]
Csel, Hgal	GACG C	M.CviJI	VGCY	2, 3	m5C	GACGCYNNNN[NNNNN]
Csel, Hgal	GACG C	M.CviQIII, M.CviSIII, M.Hpy299XI, M.Hpy30II, M.Hpy32X, M.Hpy99VII, M.HpyAX, M.HpyFII, M.HpyUM032XVII, M.Tam77409III, M.TaqI, M.TpaRLI, M.TthHB8I	TCGA	1, 4	m6A	TCGACGCNNNNN[NNNNN]
Csel, Hgal	GACG C	M.DdeI	CTNAG	1, 5	m5C	CTNAGACGCNNNNN[NNNNN], GACGCTNAGN[NNNNN]
Csel, Hgal	GACG C	M.Esp3I	GAGACG	4, 3	m6A, m5C	GAGACGCNNNNN[NNNNN]
Csel, Hgal	GACG C	M.FatI	CATG	1, 4	m5C	CATGACGCNNNNN[NNNNN], GACGCATGNN[NNNNN]
Csel, Hgal	GACG C	M.FisAW1I, M.Thal, M.Tph12270II	GCGG	1, 4	m4C	GACGCNNNN[NNNNN], GCGGACGCNNNN[NNNNN]
Csel, Hgal	GACG C	M.FnoB1II, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VIII, M.HpyFV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspJDR1, M.TmeBII	CCGG	1, 4	m4C	CCGGACGCNNNNN[NNNNN], GACGCCGGNN[NNNNN]
Csel, Hgal	GACG	M.GsuRcd1IV, M.Osp807II, M.Tth111I	GACNNNG	2, 8	m6A	GACGNGTCN[NNNNN]

	C	TC			
Csel, Hgal	GAGG C	M.HhaI, M.HinPII, M.Hpy99III, M.HpyAVIII, M.HpyUM032XV	GCGC	2, 3	m5C [NNNNN]
Csel, Hgal	GAGG C	M.HkaI930IIIV	TCGCGA	2, 5	m4C [NNNNN]
Csel, Hgal	GAGG C	M.Hme33500II	HGCWGCK	3	m4C [NNNNN]
Csel, Hgal	GAGG C	M.Hpy188I, M.Hpy298IX, M.Hpy299V, M.Hpy66V, M.HpyUM032V	TCNGA	1, 5	m6A [NNNNN]
Csel, Hgal	GAGG C	M.Hpy30IX, M.Hpy32IX, M.Hpy57IX, M.Hpy66XIII, M.Hpy8I, M.HpyFIX, M.HpyUM032XI, M.HpyUM037V, M.MjaIV, M.MspRAC08II	GTNNAC	2, 5	m6A [NNNNN]
Csel, Hgal	GAGG C	M.Hpy30VIII, M.Hpy32VIII, M.Hpy57III, M.Hpy66VI, M.Hpy99XVIII, M.HpyFVIII, M.HpyPVIII	TCNNGA	1, 6	m6A [NNNNN]
Csel, Hgal	GAGG C	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.MhuI	CTNAG	1, 5	m4C [NNNNN] GACGCTNAGN[NNNNN]
Csel, Hgal	GAGG C	M.Hpy66I, M.Hpy99I	CCWCG	2, 4	m4C [NNNNN]
Csel, Hgal	GAGG C	M.Hpy99II, M.Tsp45I, M.Tsu2489II	GTSAC	2, 4	m6A [NNNNN]
Csel, Hgal	GAGG C	M.Hpy99IV, M.HpyFOREX	CCNNGG	1, 6	m4C [NNNNN] GACGCNNGG[NNNNN]
Csel, Hgal	GAGG C	M.HpyD27I	GAGG	2, 4	m6A, m5C [NNNNN] GAGGACGCNNNNN[NNNNN]

Csel, Hgal	GACG C	M.Lph145I, M.Lph93I		GGWGA	5	m6A	GGWGACGCGNNNNN[NNNNN]
Csel, Hgal	GACG C	M.Msp1II		CCAGG	5	m4C	GACGCTGGN[NNNNN], CCAGGACGCGNNNNN[NNNNN]
Csel, Hgal	GACG C	M.Msp42II		CCWGG	1, 5	m4C	CCWGGACGCGNNNNN[NNNNN], GACGCGWGGN[NNNNN]
Csel, Hgal	GACG C	M.NheI		GCTAGC	1, 6	m4C	GACGCTAGCN[NNNNN]
Csel, Hgal	GACG C	M.Nme18III		CCAGA	2, 4	m5C	CCAGACGCGNNNNN[NNNNN]
Csel, Hgal	GACG C	M.Nme18V		GCGGGC	2, 5	m5C	GACGCGGGCN[NNNNN]
Csel, Hgal	GACG C	M.NmeDI		VCCGGY	2, 5	m5C	GACGCGGGBN[NNNNN]
Csel, Hgal	GACG C	M.NmeMC58III		GCAGA	1	m5C	TCTGGACGCGNNNNN[NNNNN], GACGCGAGAN[NNNNN]
Csel, Hgal	GACG C	M.Nsp209III, M.Nsp51109II		CTCGAG	1, 6	m5C	CTCGACGCGNNNNN[NNNNN], GACGCTCGAG[NNNNN]
Csel, Hgal	GACG C	M.NspI, M.NspHI		RCATGY	2, 5	m5C	GACGCATGYN[NNNNN]
Csel, Hgal	GACG C	M.Pam7686I		CCATGG	1, 6	m5C	CCATGGACGCGNNNNN[NNNNN], GACGCCATGG[NNNNN]
Csel, Hgal	GACG C	M.Pla23572II		CCATCG	1, 6	m5C	CGATCGACGCGNNNNN[NNNNN], GACGGGATCG[NNNNN]
Csel, Hgal	GACG C	M.Pla5I		CTAG	1, 4	Unkno wn	CTAGACGCGNNNNN[NNNNN], GACGCTAGNN[NNNNN]

Csel, Hgal	GACG C	M.Sam53668I]		CCATGG	1, 6	Unkno wn	CCATGGACGCNNNNN[NNNNN], GACGCCATGG[NNNNN]
Csel, Hgal	GACG C	M.Sen1781III		CCWWGG	1, 6	Unkno wn	CCWWGGACGCNNNNN[NNNNN], GACGCCWWGG[NNNNN]
Csel, Hgal	GACG C	M.SmoLVI, M.Spn23FII, M.Spn7465I, M.SpnRI, M.XbaI	M.Spn556I,	TCTAGA	1, 6	m6A	TCTAGACGCNNNNN[NNNNN]
Csel, Hgal	GACG C	M.SonIII		GACNNNN TGC	2, 8	m6A	GACGCNNTGC[NNNNN]
Csel, Hgal	GACG C	M.TdelI		CTCTTC	1, 2	m4C, m6A	GAAGAGACGCNNNNN[NNNNN], GACGCTCTTC[NNNNN]
Csel, Hgal	GACG C	M.TspRI		CAATG	1, 5	m5C	CAATGACGCNNNNN[NNNNN], GACGCAATGN[NNNNN]
Csel, Hgal	GACG C	M.Yen2516I		GCGGCG	2, 5	Unkno wn	GACGCGGCGN[NNNNN]
Csel, Hgal	GACG C	M1.AgoG2I, M2.AgoG2I		GCAGT	1, 4	m4C	ACTGGACGCNNNNN[NNNNN], GACGCCAGTN[NNNNN]
Csel, Hgal	GACG C	M1.BspACI		CCGC	1	m5C	GCGACGCNNNNN[NNNNN], GACGCCGCNN[NNNNN]
Csel, Hgal	GACG C	M1.HphI		GgTGA	4	m5C	GgTGACGCNNNNN[NNNNN]
Csel, Hgal	GACG C	M1.Hpy32II		GRRGA	5	m6A	GRRGACGCNNNNN[NNNNN]
Csel, Hgal	GACG C	M1.HpyAII, M1.MboII, M1.NcuI		GAAGA	5	m6A	GAAGACGCNNNNN[NNNNN]
Csel, Hgal	GACG C	M1.LpIAG30I		GAAGG	5	m5C	GACGCCTTCN[NNNNN], GAAGGACGCNNNNN[NNNNN]

Csel, Hgal	GACG C	M1.Sgr13350III, M2.Sgr13350III	GCTCTTC	2, 3	m4C, m6A	GACGCTCTTC[NNNNN]
Csel, Hgal	GACG C	M2.BeeSIII, M1.Pru9I, M2.Pru9I	GCCGT	2, 4	m4C	GACGCCGTNN[NNNNN]
Csel, Hgal	GACG C	M2.ClaSBI, M1.Fli6794IV, M2.HphI, M2.Ngo3502II, M1.Ngo6140II, M1.NgoAXVI, M1.NgoBVIII, M1.NgoFA19I, M2.NgoMVIII, M2.PspHUN102I, M1.Ssp10II	GGTGA	5	m6A	GGTGAGGCNNNNN[NNNNN]
Csel, Hgal	GACG C	M2.Hpy32II, M2.HpyAII, M2.MboII, M2.NcuI	GAAGA	4	m4C	GAAGAGGCNNNNN[NNNNN]
Csel, Hgal	GACG C	M2.HpyAVI, M1.MnII	GAGG	4	m5C	GACGCTCINN[NNNNN], GAGGACGCNNNNN[NNNNN]
Csel, Hgal	GACG C	M2.RspA76I	GGGAC	4	m6A	GGGAGGCNNNNN[NNNNN]
FokI	GCAT G	M.AbaBs12I, M.Awo1030I, M.Cdu23823I, M.Csa8155II, M.Cun9529I, M.EasL1Dcm, M.Eba57Dcm, M.Ecl34399Dcm, M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.EsaSPIII, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770REFDcm, M.Kpn928Dcm, M.KpnNIH1Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I, M.Pmi1I, M.SpaD2III, M.SplRp8II, M.SptADcm, M.Tsu2489III	CCWGG	1, 5	m5C	CCWGGGATGNNNNNNNNN[NNNN], CCWGGGATGNNNNNNNNN[NNNN]

FokI	GGAT G	M.AciI	GGCG	2, 4	m5C	GGCGGATGNNNNNNNNNN[NNNN], GGCGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.Ama55I, M.CnaPQ2III, M.Rmo1926II, M.SenC1765IV, M.SspG1I, M.XorPXII	GGGCGG	1, 6	m5C	CGGCGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.AquI	CYCGRG	1, 6	m5C	CYCGRGGATGNNNNNNNNNN[NNNN], CYCGGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.AspU4II, M.BisIII, M.BsuW23II, M.CnaPQ2II, M.Eco06Dcm, M.Eco1258IDcm, M.Eco1655Dcm, M.Eco29787Dcm, M.Eco3024Dcm, M.Eco3317Dcm, M.Eco3325Dcm, M.Eco3740Dcm, M.Eco3911Dcm, M.Eco4211Dcm, M.Eco4465Dcm, M.Eco6000Dcm, M.Eco644Dcm, M.Eco86Dcm, M.Eco9001Dcm, M.Eco9387Dcm, M.EcoE1140Dcm, M.EcoE455Dcm, M.FcnGDcm, M.FcnII14Dcm, M.FcoRII, M.Kpn223III, M.Kpn35657II, M.Kpn36Dcm, M.Kpn555I, M.Kpn62629II, M.KpnAATV, M.Kqu603Dcm, M.Psp345I, M.Sen158Dcm, M.Sen641Dcm, M.SenA46Dcm, M.SenAboDcm, M.TcoKWC4II	CCWGG	2, 4	m5C	CCWGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.AvaV, M.SliSI	GATC	1, 4	m4C	GGATGATCNNNNNNNN[NNNN]
FokI	GGAT G	M.BceSII	GGWCC	1, 5	m5C	GGATGGWCCNNNNNNNN[NNNN]

FokI	GGAT G	M.BcoSIacII, M.MspI ^{W5} III, M.NbrI28I, M.Psy18684II, M.Tam77409II, M.TbrGE1II, M.TpaRLII	CTCGAG	1, 6	m4C	CTCGAGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.RepI, M.FpsIV, M.GelI6401II, M.Pdi8503I	GGCG	1, 4	m5C	GGCGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.BfaSIII, M.Ccl10114I, M.CfrB38I, M.CfrH1I, M.EcoGVI, M.EcoKII, M.FnuI0562II, M.SbaUIV, M.SboI2419III, M.Sbo268II, M.Sen10384II, M.Sen10708II, M.Sen1080II, M.Sen1175III, M.Sen13IV, M.Sen13311IV, M.Sen1387II, M.Sen1427II, M.Sen14882III, M.Sen15791III, M.Sen158III, M.Sen16I, M.Sen1655III, M.Sen1676III, M.Sen1677III, M.Sen1728II, M.Sen1735III, M.Sen1736IV, M.Sen1764II, M.Sen1766III, M.Sen1781IV, M.Sen1783III, M.Sen18569II, M.Sen1878III, M.Sen1880II, M.Sen1896III, M.Sen1898III, M.Sen1899II, M.Sen1903III, M.Sen1906III, M.Sen1908III, M.Sen1910III, M.Sen1921III, M.Sen1927II, M.Sen195II, M.Sen2030II, M.Sen2064II, M.Sen2069III, M.Sen22462I, M.Sen255II, M.Sen318III, M.Sen3387III, M.Sen3890II, M.Sen483III, M.Sen624III, M.Sen641II, M.Sen7378II, M.Sen8391III, M.Sen8692II, M.SenA46III, M.SenAball, M.SenAboII,	ATGCAT	2, 5	m6A	GGATGCATNNNNNNNN[NNNN]

		M.SenAnaIII, M.SenC1736III, M.SenC1765III, M.SenC1808IV, M.SenC1810III, M.SenJIII, M.SenL1I, M.SenL15III, M.SenN1660II, M.SenSARA26II, M.SenSPBIII, M.SenTFIII					
FokI	GGAT G	M.Bhy27164IV, M.BstNI, M.EcoDHB4Dcm, M.EcoK54Dcm, M.Gth3570I, M.Hla49239I, M.MvaI, M.Pdi8503II, M.Sso30807Dcm, M.YpeI, M.Ype41I, M.YpeShI	CCWGG	2, 4	m4C	CCWGGATGNNNNNNNNNN[NNNN]	
FokI	GGAT G	M.BsoBI	CYCGRG	1, 6	m4C	CYCGRGGATGNNNNNNNNNN[NNNN], CYCGGGATGNNNNNNNNNN[NNNN]	
FokI	GGAT G	M.Bsp14III, M.Bsp3003I, M.Bsp3004I, M.CalB3III, M.CpeAIII, M.Esu35585I, M.FpsIV, M.Lba2031I, M.Lbo2004I, M.LiaKR2I, M.Lmo1042I, M.Lmo7956II, M.Lsp10393II, M.Msp17I, M.Sau3AI, M.Sba3I, M.Sin395I, M.Spn23FI, M.Sth368I	GATC	1, 4	m5C	GGATGATCNNNNNN[NNNN]	
FokI	GGAT G	M.BstXI	CCANNNN NNTGG	3, 10	m6A	CCANNGGATGNNNNNNNNNN[NNNN]	
FokI	GGAT G	M.BsuFI, M.CpeAIV, M.Mspl	CCGG	1, 4	m5C	CCGGGATGNNNNNNNNNN[NNNN], CCGGATGNNNNNNNNNN[NNNN]	
FokI	GGAT G	M.Cab13497I, M.Dac11109III, M.Dco13527I, M.Djo10085I, M.Hbo11551I, M.Hka19301III, M.Hme33500I, M.HsaR1I, M.Hsi33800I, M.HspKM1WII, M.HspNII, M.HtuI, M.HvoDSI, M.Lba2029II, M.Mbo45II,	CTAG	1, 4	m4C	CTAGGATGNNNNNNNNNN[NNNN]	

		M.MjaI, M.Mth7I, M.Nas12278I, M.RP3010I, M.SsmII						
FokI	GGAT G	M.CagVI	TCCGGA	1, 6	m6A	TCCGGATGNNNNNNNNNN[NNNN]		
FokI	GGAT G	M.Cdi630III	GCSSGG	2, 5	m4C	GCSSGGATGNNNNNNNNNN[NNNN]		
FokI	GGAT G	M.Cfr9I, M.SmaI, M.Sma36365II, M.Xca85II, M.XcyI, M.XmaI, M.XveII	CCCGGG	2, 5	m4C	CCCGGGATGNNNNNNNNNN[NNNN]		
FokI	GGAT G	M.CfrBI, M.Sen1735IV, M.SenAnaIV	CCWWGG	1, 6	m4C	CCWWGGGATGNNNNNNNNNN[NNNN], CCWWGGGATGNNNNNNNNNN[NNNN]		
FokI	GGAT G	M.Cmu51329I, M.EcoIA03PDcm, M.RorR1Dcm, M.Sen8391Dcm, M.ZalSM2II	CCWGG	1, 5	Unkno wn	CCWGGGATGNNNNNNNNNN[NNNN], CCWGGGATGNNNNNNNNNN[NNNN]		
FokI	GGAT G	M.CocII, M.Fsi4947II, M.TdeII, M.Tph12270I	GGNCC	1, 5	m4C	GGATGGNCCNNNNNN[NNNN]		
FokI	GGAT G	M.CpaI8696III, M.Ype19I, M.YpeDodI, M.YpeEDI, M.YpeI9I, M.YpePGI	CCWGG	2, 4	Unkno wn	CCWGGATGNNNNNNNNNN[NNNN]		
FokI	GGAT G	M.CviRI, M.Hpy298VIII, M.Hpy299VIII, M.Hpy30VI, M.Hpy32V, M.Hpy57II, M.Hpy66VIII, M.HpyFVI, M.HpyUM032VIII, M.HpyUM037VIII	TGCA	1, 4	m6A	GGATGCANNNNNNN[NNNN]		
FokI	GGAT G	M.DacII1109II, M.Hfo659II, M.MhuII, M.MjaV, M.MlaZII, M.Msp1LW5I, M.RsaI	GTAC	1, 4	m4C	GGATGTACNNNNNN[NNNN]		
FokI	GGAT G	M.DdeI	CTNAG	1, 5	m5C	CTNAGGATGNNNNNNNNNN[NNNN]		
FokI	GGAT G	M.DsaV, M.Ecl18kI, M.FpsjVI, M.NlaX, M.SenPI, M.SsoII	CCNGG	2, 4	m5C	CCNGGATGNNNNNNNNNN[NNNN]		

FokI	GGAT G	M.Eco29kI, M.NgoAIII, M.SacII	CCGCGG	2, 5	m5C	CCGCGGATCNNNNNNNNNN[NNNN]
FokI	GGAT G	M.EcoNI	CCTNNNN NAGG	2, 10	m4C	CCTNNNNNAGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.EfaRFII, M.Mae7806II	CCGG	2, 3	m4C	CCGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.Esp1396I	CCANNNN NTGG	3, 9	m6A	CCANNGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.FatI	CATG	1, 4	m5C	CATCGATCNNNNNNNNNN[NNNN]
FokI	GGAT G	M.FisAW1I, M.Thal, M.Tph12270II	CGCG	1, 4	m4C	CGCGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.FnuBIII, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VII, M.HpyFV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspIDRI, M.TmeBII	CCGG	1, 4	m4C	CCGGGATGNNNNNNNNNN[NNNN], CCGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.FpsJI, M.HpaII, M.Nme18II, M.NmeAI	CCGG	2, 3	m5C	CCGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.Gel16401IV	CTGG	3	m5C	CTCGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.Hpy188I, M.Hpy298IX, M.Hpy299V, M.Hpy66V, M.HpyUM032V	TCNGA	1, 5	m6A	TCGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.Hpy30VIII, M.Hpy32VIII, M.Hpy57III, M.Hpy66VI, M.Hpy99XVIII, M.HpyFVIII, M.HpyPVIII	TCNNGA	1, 6	m6A	TCNCGATGNNNNNNNNNN[NNNN]

FokI	GGAT G	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.Mhul	CTNAG	1, 5	m4C	CTNAGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.Hpy99IV, M.HpyFORFX	CCNNGG	1, 6	m4C	CCNNGGATGNNNNNNNNNN[NNNN], CCNNGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.HpyAV, M.Nme18IV	GAAGG	3, 4	m6A, m5C	GAAGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.HpyD27I	GAGG	2, 4	m6A, m5C	GAGGATGNNNNNNNNNN[NNNN], GAGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.MjaII	GCNCC	1, 5	m5C	GCATCGGCCNNNNNNNN[NNNN]
FokI	GGAT G	M.Msp1II	CCAGG	5	m4C	CCAGGGATGNNNNNNNNNN[NNNN], CCAGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.Msp42II	CCWGG	1, 5	m4C	CCWGGGATGNNNNNNNNNN[NNNN], CCWGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.NheI	GCTAGC	1, 6	m4C	GGATGCTAGCNNNNNN[NNNN]
FokI	GGAT G	M.Nme18III	TCTGG	2, 4	m5C	TCTGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.NmeDI	RCCGGB	2, 5	m5C	RCCGGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.NmeMC58III	TCTGG	5	m5C	TCTGGGATGNNNNNNNNNN[NNNN], TCTGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.Nsp209III, M.Nsp51109II	CTCGAG	1, 6	m5C	CTCGAGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.Pam7686I	CCATGG	1, 6	m5C	CCATGGGATGNNNNNNNNNN[NNNN], CCATGGATGNNNNNNNNNN[NNNN]

FokI	GGAT G	M.Pla23572II		CGATCG	1, 6	m5C	CGATCGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.Pla5I		CTAG	1, 4	Unkno wn	CTAGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.Sam53668II		CCATGG	1, 6	Unkno wn	CCATGGGATGNNNNNNNNNN[NNNN], CCATGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.Sen1781III		CCWWGG	1, 6	Unkno wn	CCWWGGGATGNNNNNNNNNN[NNNN], CCWWGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.SonIII		GACNNNN TGC	2, 8	m6A	GACNGGATGCNNNNNNNNNN[NNNN]
FokI	GGAT G	M.TdelI		GAAGAG	5, 6	m6A, m4C	GAAGAGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M.TspRI		CASTG	1, 5	m5C	CASTGGATCNNNNNNNNNN[NNNN]
FokI	GGAT G	M.XcmI		GCANNNN NNNNNTG G	3, 13	m6A	CCANNNNNGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M1.AgoG2I, M2.AgoG2I		ACTGG	2, 5	m4C	ACTGGGATGNNNNNNNNNN[NNNN], ACTGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M1.Bag18758II, M2.Bag18758II		CACGGG	3, 5	Unkno wn	CACGGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M1.BclI, M1.Hpy30XI, M2.Hpy66X, M1.Jsp2502I, M2.MmyCVI		GATGG	2	m6A	GGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M1.BclI, M2.BclI, M.NclI		CCSGG	2, 4	m4C	CCSGGATGNNNNNNNNNN[NNNN]
FokI	GGAT G	M1.BfiI		ACTGGG	5	m4C	ACTGGGATGNNNNNNNNNN[NNNN]

	G								
FokI	GGAT	M1.BhaI, M1.Bst19I, M.Fnu10562III, M1.Mbo45IV, M2.Mbo45IV	GATGC	3	m6A	GGATGCNNNNNNNNNN[NNNN]			
FokI	G	M1.BspACI	GCGG	4	m5C	GCGGGATGNNNNNNNNNN[NNNN], GCGGATGNNNNNNNNNN[NNNN]			
FokI	GGAT	M1.Csp16704IV, M2.Csp16704IV, M.Lsp10393I, M.Mha183II, M.Mha185II, M.Mha807II, M.Mva1312II	GATGC	2, 3	m6A	GGATGCNNNNNNNNNN[NNNN]			
FokI	G	M1.Hpy299X, M2.Hpy32XIII, M1.HpyFXIII, M2.HpyFXIII, M2.HpyUM037VI, M1.McaCI, M2.McaCI, M1.PspHUN102II, M2.PspHUN102II	GATGG	2, 3	m6A	GGATGCNNNNNNNNNN[NNNN]			
FokI	GGAT	M1.Hpy32II	GRRGA	5	m6A	GRGGATGNNNNNNNNNN[NNNN]			
FokI	G	M1.LpLAG30I	GAAGG	5	m5C	GAAGGGATGNNNNNNNNNN[NNNN], GAAGGATGNNNNNNNNNN[NNNN]			
FokI	GGAT	M2.BclI, M2.Hpy299X, M1.Hpy66X, M1.HpyC1I, M2.HpyC1I, M1.HpyUM032X, M2.Jsp2502I, M1.MmyCVI	GATGG	3	m6A	GGATGCNNNNNNNNNN[NNNN]			
FokI	G	M2.BhaI	GATGC	2	m6A	GGATGCNNNNNNNNNN[NNNN]			
FokI	GGAT	M2.HpyAVI, M1.MnII	GAGG	4	m5C	GAGGGATGNNNNNNNNNN[NNNN], GAGGATGNNNNNNNNNN[NNNN]			
SstI	G	M.AbaBs12I, M.Awo1030I, M.Cdu23823I, M.Csa8155II, M.Cun9529I, M.EasL1Dcm, M.Eba57Dcm, M.Ecl34399Dcm	CCWGG	1, 5	m5C	CCWGGGATGNNNNNNNNNN[NNNN], CCWGGATGNNNNNNNNNN[NNNN]			

		M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.EsaSPIII, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFDcm, M.Kpn928Dcm, M.KpnNIH1Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I, M.Pmi1I, M.SpaD2III, M.SplRp8II, M.SptADcm, M.Tsu2489III				
StsI	GGAT G	M.AciI	GCGG	2, 4	m5C	GCGGGATGNNNNNNNNNNNN[NNNN], GCGGATGNNNNNNNNNNNN[NNNN]
StsI	GGAT G	M.Ama55I, M.CnaPQ2III, M.Rmo1926II, M.SenC1765IV, M.SspG1I, M.XorPXII	GCGCGG	1, 6	m5C	GCGCCGGATGNNNNNNNNNNNN[NNNN]
StsI	GGAT G	M.AquI	CYCGRG	1, 6	m5C	CYCGRGGATGNNNNNNNNNNNN[NNNN], CYCGGGATGNNNNNNNNNNNN[NNNN]
StsI	GGAT G	M.AspU41II, M.BisIII, M.BsuW23II, M.CnaPQ2II, M.Eco06Dcm, M.Eco1258IDcm, M.Eco1655Dcm, M.Eco29787Dcm, M.Eco3024Dcm, M.Eco3317Dcm, M.Eco3325Dcm, M.Eco3740Dcm, M.Eco3911Dcm, M.Eco4211Dcm, M.Eco4465Dcm, M.Eco600Dcm, M.Eco644Dcm, M.Eco86Dcm, M.Eco9001Dcm, M.Eco9387Dcm, M.EcoE1140Dcm, M.EcoE455Dcm, M.EcoGDcm, M.EcoNU14Dcm, M.EcoRII,	CCWGG	2, 4	m5C	CCWGGATGNNNNNNNNNNNN[NNNN]

			M.Kpn223III, M.Kpn35637II, M.Kpn36Dcm, M.Kpn555I, M.Kpn62629II, M.KpnAATV, M.Kqu603Dcm, M.Psp345I, M.Sen158Dcm, M.Sen641Dcm, M.SenA46Dcm, M.SenAboDcm, M.TcoKWC4II						
StsI	GGAT G		M.AvaV, M.SliSI	GATC	1, 4	m4C	GGATGATC	NNNNN	
StsI	GGAT G		M.BceSII	GGWGC	1, 5	m5C	GGATGGWCC	NNNNN	
StsI	GGAT G		M.BcoSlaII, M.MspLW5III, M.Nbr128I, M.Psy18884II, M.Tam77409II, M.TbrGE1II, M.TpaRII	CTCGAG	1, 6	m4C	CTCGAGGATG	NNNNN	NNNNN
StsI	GGAT G		M.BepI, M.FpsIV, M.Gel16401II, M.Pdi8503I	CGCG	1, 4	m5C	CGCGGATG	NNNNN	NNNNN
StsI	GGAT G		M.BfaSIII, M.Ccl10114I, M.CfrB38I, M.CfrH1I, M.EcoGV, M.EcoKII, M.Fnu10562II, M.SbaUIV, M.Sbo12419II, M.Sbo268II, M.Sen10384II, M.Sen10708II, M.Sen1080II, M.Sen1175III, M.Sen13IV, M.Sen13311IV, M.Sen1387II, M.Sen1427III, M.Sen14882III, M.Sen15791II, M.Sen158III, M.Sen16I, M.Sen1655III, M.Sen1676III, M.Sen1677III, M.Sen1728III, M.Sen1735III, M.Sen1736IV, M.Sen1764II, M.Sen1766III, M.Sen1781IV, M.Sen1783II, M.Sen18569II, M.Sen1878III, M.Sen1880III,	ATGCAT	2, 5	m6A	GGATGCAT	NNNNN	NNNNN

		M.Sen1896III, M.Sen1898III, M.Sen1899III, M.Sen1903III, M.Sen1906III, M.Sen1908III, M.Sen1910III, M.Sen1921III, M.Sen1927II, M.Sen195II, M.Sen2030II, M.Sen2064III, M.Sen2069III, M.Sen22462I, M.Sen255II, M.Sen318III, M.Sen3387III, M.Sen3890II, M.Sen483III, M.Sen624III, M.Sen641III, M.Sen7378II, M.Sen8391III, M.Sen8692III, M.SenA46III, M.SenAbalI, M.SenAbolI, M.SenAnallI, M.SenC1736III, M.SenC1765III, M.SenC1808IV, M.SenC1810III, M.SenJIII, M.SenLII, M.SenL15III, M.SenN1660II, M.SenSARA26II, M.SenSPBIII, M.SenTFIII				
StsI	GGAT G	M.Bhy27164IV, M.BstNI, M.EcoDHB4Dcm, M.EcoK54Dcm, M.Gth3570I, M.Hla49239I, M.MvaI, M.Pdi8503II, M.Sso30807Dcm, M.Ypel, M.Ype41I, M.YpeShI	CCWGG	2, 4	m4C	CCWGGATGNNNNNNNNNN[NNNN]
StsI	GGAT G	M.BsoBI	CYCGRG	1, 6	m4C	CYCGRGGATGNNNNNNNNNN[NNNN], CYCGGGATGNNNNNNNNNN[NNNN]
StsI	GGAT G	M.Bsp14III, M.Bsp3003I, M.Bsp3004I, M.CalB3III, M.CpeAIII, M.Esu35585I, M.FpsIV, M.Lba2031I, M.Lbo2004I, M.LlaKR2I, M.Lmo1042I, M.Lmo7956II, M.Lsp10393II, M.Msp17I, M.Sau3AI, M.Sba3I, M.Sim395I, M.Spn23FI, M.Sth368I	GATC	1, 4	m5C	GGATGATCNNNNNNNN[NNNN]

SstI	GGAT G	M.BstXI	CCANNNN NNTGG	3, 10	m6A	CCANNNGGATGNNNNNNNNNN[NNNN]
SstI	GGAT G	M.BsuFI, M.CpeAIV, M.MspI	CCGG	1, 4	m5C	CCGGGATGNNNNNNNNNN[NNNN], CCGGATGNNNNNNNNNN[NNNN]
SstI	GGAT G	M.CabI3497I, M.Dac111109III, M.Dco13527I, M.Djo10085I, M.Hbo11551I, M.Hka19301III, M.Hme33500I, M.HsaR1I, M.Hsi33800I, M.HspKM1VII, M.HspNII, M.Htul, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.Mjal, M.MthZI, M.Nas12278I, M.Rfi3010I, M.SsmMI	CTAG	1, 4	m4C	CTAGGATGNNNNNNNNNN[NNNN]
SstI	GGAT G	M.CagVI	TCCGGA	1, 6	m6A	TCCGGATGNNNNNNNNNN[NNNN]
SstI	GGAT G	M.Cdi630III	CCSSGG	2, 5	m4C	CCSSGGATGNNNNNNNNNN[NNNN]
SstI	GGAT G	M.Cfr9I, M.SmaI, M.Sma36365II, M.Xca85III, M.XcyI, M.XmaI, M.XveII	CCCGGG	2, 5	m4C	CCCGGGATGNNNNNNNNNN[NNNN]
SstI	GGAT G	M.CfrBI, M.Sen1735IV, M.SenAnaIV	CCWWGG	1, 6	m4C	CCWWGGGATGNNNNNNNNNN[NNNN], CCWWGGATGNNNNNNNNNN[NNNN]
SstI	GGAT G	M.Cmu51329I, M.EcoIA03PDcm, M.RorR1Dcm, M.Sen8391Dcm, M.ZaiSM2II	CCWGG	1, 5	Unkno wn	CCWGGGATGNNNNNNNNNN[NNNN], CCWGGATGNNNNNNNNNN[NNNN]
SstI	GGAT G	M.CocII, M.Fsi4947II, M.TdeII, M.Tph12270I	GGNCC	1, 5	m4C	GGATGGNCCNNNNNN[NNNN]
SstI	GGAT G	M.CpaI8696III, M.Ype19I, M.YpeDodI, M.YpeEDI, M.YpeI9I, M.YpePGI	CCWGG	2, 4	Unkno wn	CCWGGATGNNNNNNNNNN[NNNN]
SstI	GGAT	M.CviRI, M.Hpy298VIII, M.Hpy299VIII,	TGCA	1, 4	m6A	GGATGCANNNNNNNN[NNNN]

	G	M.Hpy30VI, M.Hpy32V, M.Hpy57II, M.Hpy66VIII, M.HpyFVI, M.HpyUM032VIII, M.HpyUM037VIII					
StsI	GGAT G	M.DacII109II, M.Hfo659II, M.MhuII, M.MjaV, M.MlaZII, M.Msp1W5I, M.RsaI	GTAC	1, 4	m4C	GGATGTACNNNNNNNN[NNNN]	
StsI	GGAT G	M.DdeI	CTNAG	1, 5	m5C	CTNAGGATGNNNNNNNNNN[NNNN]	
StsI	GGAT G	M.DsaV, M.EclI18kI, M.FpsVI, M.NlaX, M.SenPI, M.SsoII	CCNGG	2, 4	m5C	CCNGGATGNNNNNNNNNN[NNNN]	
StsI	GGAT G	M.Eco29kI, M.NgoAIII, M.SacII	CCGCGG	2, 5	m5C	CCGCGGATGNNNNNNNNNN[NNNN]	
StsI	GGAT G	M.EcoNI	CCTNNNN NAGG	2, 10	m4C	CCTNNNNAGGATGNNNNNNNNNN[NNNN] J	
StsI	GGAT G	M.EfaRFII, M.Mae7806II	CCGG	2, 3	m4C	CCGGATGNNNNNNNNNN[NNNN]	
StsI	GGAT G	M.Esp1396I	CCANNNN NTGG	3, 9	m6A	GCANNGGATGNNNNNNNNNN[NNNN]	
StsI	GGAT G	M.FatI	CATG	1, 4	m5C	CATGGATGNNNNNNNNNN[NNNN]	
StsI	GGAT G	M.FisAW1I, M.Thal, M.Tph12270II	CGCG	1, 4	m4C	CGCGGATGNNNNNNNNNN[NNNN]	
StsI	GGAT G	M.FnuBII, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VIII, M.HpyFV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspIDRI, M.TneBII	CCGG	1, 4	m4C	CCGGGATGNNNNNNNNNN[NNNN], CCGGATGNNNNNNNNNN[NNNN]	

SstI	GGAT G	M.FpsI, M.HpaII, M.Nme18II, M.NmeAI	CCGG	2, 3	m5C	CCGGATGNNNNNNNNNNNN[NNNN]
SstI	GGAT G	M.Gel16401IV	CTGG	3	m5C	CTGGATGNNNNNNNNNNNN[NNNN]
SstI	GGAT G	M.Hpy188I, M.Hpy298IX, M.Hpy299V, M.Hpy66V, M.HpyUM032V	TCNGA	1, 5	m6A	TCGGATGNNNNNNNNNNNN[NNNN]
SstI	GGAT G	M.Hpy30VIII, M.Hpy32VIII, M.Hpy57III, M.Hpy66VI, M.Hpy99XVIII, M.HpyFVII, M.HpyPVIII	TCNNGA	1, 6	m6A	TCNGGATGNNNNNNNNNNNN[NNNN]
SstI	GGAT G	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.MhuI	CTNAG	1, 5	m4C	CTNAGGATGNNNNNNNNNNNN[NNNN]
SstI	GGAT G	M.Hpy99IV, M.HpyFORFX	CCNNGG	1, 6	m4C	CCNNGGATGNNNNNNNNNNNN[NNNN], CCNNGGATGNNNNNNNNNNNN[NNNN]
SstI	GGAT G	M.HpyAV, M.Nme18IV	GAAGG	3, 4	m6A, m5C	GAGGATGNNNNNNNNNNNN[NNNN]
SstI	GGAT G	M.HpyD27I	GAGG	2, 4	m6A, m5C	GAGGATGNNNNNNNNNNNN[NNNN], GAGGATGNNNNNNNNNNNN[NNNN]
SstI	GGAT G	M.MjaII	GGNCC	1, 5	m5C	GGATGGCCNNNNNNNN[NNNN]
SstI	GGAT G	M.Msp1II	CCAGG	5	m4C	CCAGGATGNNNNNNNNNNNN[NNNN], CCAGGATGNNNNNNNNNNNN[NNNN]
SstI	GGAT G	M.Msp42II	CCWGG	1, 5	m4C	CCWGGATGNNNNNNNNNNNN[NNNN], CCWGGATGNNNNNNNNNNNN[NNNN]
SstI	GGAT G	M.NheI	GCTAGC	1, 6	m4C	GGATGCTAGCNNNNNN[NNNN]
SstI	GGAT	M.Nme18III	TCTGG	2, 4	m5C	TCTGGATGNNNNNNNNNNNN[NNNN]

StsI	G	GGAT	M.NmcDI	RCCGGB	2, 5	m5C	RCCGGGATGNNNNNNNNNN[NNNN]
StsI	G	GGAT	M.NmeMC58III	TCTGG	5	m5C	TCTGGGATGNNNNNNNNNN[NNNN], TCTGGATGNNNNNNNNNN[NNNN]
StsI	G	GGAT	M.Nsp209III, M.Nsp51109II	CTCGAG	1, 6	m5C	CTCGAGGATGNNNNNNNNNN[NNNN]
StsI	G	GGAT	M.Pam7686I	CCATGG	1, 6	m5C	CCATGGGATGNNNNNNNNNN[NNNN], CCATGGATGNNNNNNNNNN[NNNN]
StsI	G	GGAT	M.Pfa23572II	CGATCG	1, 6	m5C	CGATCGGATGNNNNNNNNNN[NNNN]
StsI	G	GGAT	M.Pla5I	CTAG	1, 4	Unkno wn	CTAGGATGNNNNNNNNNN[NNNN]
StsI	G	GGAT	M.Sam53668II	CCATGG	1, 6	Unkno wn	CCATGGGATGNNNNNNNNNN[NNNN], CCATGGATGNNNNNNNNNN[NNNN]
StsI	G	GGAT	M.Sen1781III	CCWWGG	1, 6	Unkno wn	CCWWGGGATGNNNNNNNNNN[NNNN], CCWWGGATGNNNNNNNNNN[NNNN]
StsI	G	GGAT	M.SonII	GACNNNN TGC	2, 8	m6A	GACNGGATGNNNNNNNNNN[NNNN]
StsI	G	GGAT	M.TdelI	GAAGAG	5, 6	m6A, m4C	GAAGAGGATGNNNNNNNNNN[NNNN]
StsI	G	GGAT	M.TspRI	CASATG	1, 5	m5C	CASATGGATGNNNNNNNNNN[NNNN]
StsI	G	GGAT	M.XcmI	CCANNNN NNNNNTG G	3, 13	m6A	CCANNNNNGGATGNNNNNNNNNN[NNNN]]

StsI	GGAT G	M1.AgoG2I, M2.AgoG2I	ACTGG	2, 5	m4C	ACTGGGATGNNNNNNNNNNNN[NNNN], ACTGGATGNNNNNNNNNNNN[NNNN]
StsI	GGAT G	M1.Bag18758II, M2.Bag18758II	CACGGG	3, 5	Unkno wn	CACGGGATGNNNNNNNNNNNN[NNNN]
StsI	GGAT G	M1.BccI, M1.Hpy30XI, M2.Hpy66X, M1.Jsp2502I, M2.MmyCVI	GATGG	2	m6A	GGATGNNNNNNNNNN[NNNN]
StsI	GGAT G	M1.BcnI, M2.BcnI, M.NciI	CCSGG	2, 4	m4C	CCSGGATGNNNNNNNNNNNN[NNNN]
StsI	GGAT G	M1.BfiI	ACTGGG	5	m4C	ACTGGGATGNNNNNNNNNNNN[NNNN]
StsI	GGAT G	M1.BhaI, M1.Bst19I, M.Fnu10562III, M1.Mbo45IV, M2.Mbo45IV	GATGC	3	m6A	GGATGNNNNNNNNNN[NNNN]
StsI	GGAT G	M1.BspACI	GCGG	4	m5C	GCGGGATGNNNNNNNNNNNN[NNNN], GCGGATGNNNNNNNNNNNN[NNNN]
StsI	GGAT G	M1.Csp16704IV, M2.Csp16704IV, M.Lsp10393I, M.Mha183II, M.Mha185II, M.Mha807II, M.Mva1312II	GATGC	2, 3	m6A	GGATGNNNNNNNNNN[NNNN]
StsI	GGAT G	M1.Hpy299X, M2.Hpy32XIII, M1.HpyFXIII, M2.HpyFXIII, M2.HpyUM037VI, M1.McaCI, M2.McaCI, M1.PspHUN102II, M2.PspHUN102II	GATGG	2, 3	m6A	GGATGNNNNNNNNNN[NNNN]
StsI	GGAT G	M1.Hpy32II	GRRGA	5	m6A	GRGGATGNNNNNNNNNNNN[NNNN]
StsI	GGAT G	M1.LplAG:30I	GAAGG	5	m5C	GAAGGATGNNNNNNNNNNNN[NNNN], GAAGGATGNNNNNNNNNNNN[NNNN]
StsI	GGAT	M2.BccI, M2.Hpy299X, M1.Hpy66X,	GATGG	3	m6A	GGATGNNNNNNNNNN[NNNN]

	G	M1.HpyC1I, M2.HpyC1I, M1.HpyUM032X, M2.Jsp2502I, M1.MmyCVI					
StsI	GGAT G	M2.BhaI	GATGC	2	m6A	GGATGCNNNNNNNNNN[NNNN]	
StsI	GGAT G	M2.HpyAVI, M1.MnII	GAGG	4	m5C	GAGGATGNNNNNNNNNN[NNNN], GAGGATGNNNNNNNNNN[NNNN]	
Acc36I, BfuAI, BspMI, BveI	ACCT GC	M.AauTCI, M.AchA6II, M.Asp146I, M.Asp24188I, M.Asp31YII, M.AspU41I, M.Bsp KG1III, M.EsaDix4I, M.EsaDix5I, M.Llu27648II, M.Mch9957II, M.Mru1279VI, M.MseI, M.Msi9946I, M.Sam53668II, M.Sth20745I, M.Tam77409I	TTAA	1, 4	m6A	TTAAGCTGCNNNNNN[NNNN]	
Acc36I, BfuAI, BspMI, BveI	ACCT GC	M.AbaBs12I, M.Awo1030I, M.Cdu23823I, M.Csa8155II, M.Cun9529I, M.EasL1Dcm, M.Eba57Dcm, M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.EsaSPIII, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFDcm, M.Kpn928Dcm, M.KpnNIH1Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I, M.Pmi1I, M.SpaD2III, M.SplRp8II, M.SptADcm, M.Tsu2489III	CCWGG	1, 5	m5C	ACCTGCCWGG[NNNN]	
Acc36I, BfuAI, BspMI, BveI	ACCT GC	M.Acc65I, M.Eco86I, M.KpnI, M.Lai216II, M.Sen3387IV, M.Sen839IIV	GGTACC	3, 4	m6A	GGTACCTGCNNNNNN[NNNN]	

Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Acc65VI, M.SenAZIII		CTKMAG	2, 5	m6A	ACCTGCAGNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.AccI, M.FpsFFG3II, M.SdeAIV		GTMKAC	2, 5	m6A	GTMKACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.AciI		CCGC	1, 3	m5C	ACCTGGGGNN[NNNN], ACCTGGCGCN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Ade56712I, M.Asp588III, M.AveCB51II, M.BbrUIII, M.Blo30698I, M.BsuBI, M.CpaYL7III, M.Dba7044I, M.Dfa45958II, M.Eco1258II, M.Eco2012I, M.EcoGII, M.EcoG089I, M.Eho16691II, M.Hal5920II, M.Lsp6406II, M.MspNR41II, M.Pac490I, M.Pox23570I, M.PstI, M.RpaII, M.Sca3403II, M.Tsp61I, M.XmnII, M.YenWI, M.YinY228I		CTGCAG	2, 5	m6A	ACCTGCAGNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.AgsI		TTTAA	1, 5	m6A	TTTAAACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.AhdI		GACNNNN NGTC	2, 10	m6A	GACCTGCNGTC[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Alw26I		GAGAC	4, 3	m6A, m5C	GAGACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.AorT14III, M.GsuRed1III, M.SalI		GTCGAC	2, 5	m6A	GTCGACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.AplI		CTGCAG	3, 4	m5C	ACCTGCAGNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.AvaIX, M.Cfr10I, M.Nme18VI, M.VchAI, M.Vch0395I		RCGGY	2, 5	m5C	ACCTGCCGGY[NNNN]

Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.BamWS8I, M.CthV, M.EfaRFL, M.Saf8902II, M.TacII, M.TcoKWC4III	GCWGC	2, 4	m5C	ACCTGCWGCN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.BanLII, M.CpeAII, M.SimI	GGWCC	2, 4	m5C	GGACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.BbvI, M.BceSIV	GCAGC	2, 4	m5C	ACCTGCTGCN[NNNN], ACCTGCACGN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.BceI6656II, M.Bce22EII, M.Bce254I6II, M.Bce842II, M.Bce895II, M.BceJIV, M.BceK56II, M.Bgl2196I, M.BmaBMKII, M.BmaBMZII, M.Bps10134II, M.Bps1651II, M.Bps7894II, M.Bps840II, M.BpsR668II, M.BspRB39I, M.Bth1643II, M.Cba4G11II, M.MspNR41I, M.Oba14I, M.Pap16535I, M.Pfa23572I, M.Pox23570II, M.PpnI, M.PpnRB38I, M.Pth25325I	GTWWAC	2, 5	m6A	GTWWACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.BceSII	GGWCC	1, 5	m5C	GGACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.BepI, M.FpsIV, M.Gel16401II, M.Pdl8503I	CGCG	1, 4	m5C	ACCTGGCGCN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Bhy27164V, M.DraI, M.Lsp6406V, M.Tsu2489V	TTTAAA	1, 6	m6A	TTTAAACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.BspRAC04I, M.Ecal, M.Eco3936IV	GGTNACC	3, 5	m6A	GGTNACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.BstXI	CCANNNN NNTGG	3, 10	m6A	CCACCTGCNTGG[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.BsuFI, M.CpeAIV, M.Mspl	CCGG	1, 4	m5C	ACCTGCCGCN[NNNN]

BspMI, Bvel	GC		CTAG		1, 4	m4C	ACCTGCTAGN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Cab13497I, M.Dac11109III, M.Dco13527I, M.Djo10085I, M.Hbo1155II, M.Hka19301III, M.Hme33500I, M.HsaRII, M.Hsi33800I, M.HspKM1VII, M.HspNII, M.Htul, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.Mjal, M.MthZI, M.Nas12278I, M.Rfi3010I, M.SsmMI	CTAG				
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Cac8I	GCNNGC		2, 5	m5C	ACCTGCNNGC[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.CagV, M.Lsp6406I, M.Rsp7740I, M.RspPBTS2II, M.VbaLP2AI	TTCGAA		1, 6	m6A	TTCGAACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.CagVI	TCCGGA		1, 6	m6A	TCCGGACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Cdi630IV, M.Ckr177III, M.CmaLM2II, M.CocII, M.Fnu4HI, M.Fsp4HI	GCNGC		2, 4	m5C	ACCTGCNGCN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Cdi630V, M.Cdi81I, M.CdiG46II, M.CmaLM2I, M.CpaYI7II, M.Pat19335II, M.PbaVA2I, M.Pd19689I	CAAAAA		6	m6A	CAAAAAACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Cmu51329I, M.EcoJA03PDcm, M.RorR1Dcm, M.Sen8391Dcm, M.ZaiSM2II	CCWGG		1, 5	Unkno wn	ACCTGCCWGG[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.CocI, M.Fli6794II, M.IincII, M.IindII, M.HinHia1II	GTYRAC		2, 5	m6A	GTYRACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.CocII, M.Fsi4947II, M.TdeII, M.Tph12270I	GGNCC		1, 5	m4C	GGACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT	M.CpeAVI, M.TcoKWC4I	GTATAC		2, 5	m6A	GTATACCTGCNNNN[NNNN]

BspMI, Bvel	GC						
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.CsiI, MfoCT6IV, M.NgoAIV, M.NgoMIV, M1.Sgr13350II	GCCGGC	2, 5	m5C	ACCTGCCGGC[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.CviQI, M.Hpy32XII, M.Hpy57XI, M.Hpy66XII, M.Hpy99XII, M.HpyFXII, M.HpyPIV, M.HpyUM037XI, M.PabI	GTAC	2, 3	m6A	GTACCTGCNNNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.CviQIII, M.CviSIII, M.Hpy299XI, M.Hpy30II, M.Hpy32X, M.Hpy99VII, M.HpyAX, M.HpyFII, M.HpyUM032XVII, M.Tam77409III, M.TaqI, M.TpaRLI, M.TthHB8I	TCGA	1, 4	m6A	TCGACCTGCNNNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.CviRI, M.Hpy298VIII, M.Hpy299VIII, M.Hpy30VI, M.Hpy32V, M.Hpy57II, M.Hpy66VIII, M.HpyFVI, M.HpyUM032VIII, M.HpyUM037VIII	TGCA	1, 4	m6A	TGCACCTGCNNNN[NNNN], ACCTGCANNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.DacI1109II, M.Hfo6591I, M.MhuII, M.MjaV, M.MlaZII, M.MspLW5I, M.RsaI	GTAC	1, 4	m4C	GTACCTGCNNNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.DdeI	CTNAG	1, 5	m5C	ACCTGCTNAG[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.DfeI, M.Fnu419II, M.Pen113II	CTNNAG	2, 5	m6A	ACCTGCAGNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Eco0109I	RGNCCT	3, 5	m5C	RGACCTGCNNNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Esp1396I	CCANNNN NTGG	3, 9	m6A	CCACCTGCTGGN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.FatI	CATG	1, 4	m5C	ACCTGCATGN[NNNN]	

BspMI, Bvel	GC						
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.FisAW1I, M.Thal, M.Tph12270II	CGCG	1, 4	m4C	ACCTGGCGGN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.FnoB1II, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VII, M.HpyFV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspIDRI, M.TmeBII	CCGG	1, 4	m4C	ACCTGGCGGN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.HgiCI	GGYRCC	2, 5	m5C	GGYACCTGCNNNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Hhal, M.HinP1I, M.Hpy99III, M.HpyAVIII, M.HpyUM032XV	GGGC	2, 3	m5C	ACCTGGCGGN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Hme33500II	HGCWUCC	3	m4C	ACCTGCWGGCK[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Hpal, M.Mwh3033I, M.Rsp7I	GTTAAC	2, 5	m6A	GTTAACCTGCNNNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Hpy188I, M.Hpy298IX, M.Hpy299V, M.Hpy66V, M.HpyUM032V	TCNGA	1, 5	m6A	TCNGACCTGCNNNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Hpy30IX, M.Hpy32IX, M.Hpy57IX, M.Hpy66XIII, M.Hpy8I, M.HpyFIX, M.HpyUM032XI, M.HpyUM037V, M.MjaIV, M.MspRAC08II	GTNNAC	2, 5	m6A	GTNNACCTGCNNNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Hpy30VIII, M.Hpy32VIII, M.Hpy57III, M.Hpy66VI, M.Hpy99XVIII, M.HpyFVIII, M.HpyPVIII	TCNNGA	1, 6	m6A	TCNNGACCTGCNNNN[NNNN]	
Acc36I, BfuAI	ACCT	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.Mhul	CTNAG	1, 5	m4C	ACCTGCTNAG[NNNN]	

BspMI, Bvel	GC						
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Hpy57X		CTRYAG	2, 5	m6A	ACCTGCAGNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Hpy99II, M.Tsp45I, M.Tsu2489II		GTSAC	2, 4	m6A	GTSACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.HpyD27I		CCTC	1, 3	m5C, m6A	ACCTGGCTCN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Lph145I, M.Lph93I		GGWGA	5	m6A	GGWACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Mae7896III, M.PspPI, M.Sau96I		GGNCC	2, 4	m5C	GGACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.MjaII		GGNCC	1, 5	m5C	GGACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.MspIII		CCTGG	1	m4C	ACCTGCCTGG[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Msp42II		GCWGG	1, 5	m4C	ACCTGGCWGG[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.NgoAXV, M.NgoMV		GGNNCC	2, 5	m5C	GGNACCTGCNNNN[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.NheI		GCTAGC	1, 6	m4C	ACCTGCTAGC[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Nme18V		GGGCGC	2, 5	m5C	ACCTGGGCGC[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.NmeDI		VCCGGY	2, 5	m5C	ACCTGCCGGB[NNNN]
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.NmeMC58III		CCAGA	1	m5C	ACCTGCCAGA[NNNN]

BspMI, Bvel	GC						
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.NspI, M.NspHI	RCATGY	2, 5	m5C	ACCTGCATGY[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Pla5I	CTAG	1, 4	Unkno wn	ACCTGCTAGN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.SmoLV1, M.Spn23FII, M.Spn556I, M.Spn7465I, M.SpnRI, M.XbaI	TCTAGA	1, 6	m6A	TCTAGACCTGCNNNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.SonIII	GACNNNN TGC	2, 8	m6A	GACCTGCTGCN[NNNN], GACNACCTGCNNNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.TspRI	CASTG	1, 5	m5C	ACCTGCASGT[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M.Yen2516I	GCGGCG	2, 5	Unkno wn	ACCTGGGCGC[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M1.AgoG2I, M2.AgoG2I	CCAGT	1, 4	m4C	ACCTGCCAGT[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M1.RspACI	CCGC	1	m5C	ACCTGGCGCN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M1.Eco31I, M1.EcoMI	GAGACC	4	m6A	GAGACCTGCNNNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M1.HgaI, M2.LlaII	GCGTC	2	m5C	ACCTGGCTCN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M1.Hpy32II	GRRGA	5	m6A	GRRGACCTGCNNNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M1.HpyAII, M1.MbolI, M1.NcuI	GAAGA	5	m6A	GAAGACCTGCNNNN[NNNN]	
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M1.LplAG30I	CCTTC	1	m5C	ACCTGCCTTC[NNNN]	

BspMI, Bvel	GC							
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M2.BceSIII, M1.Pru9I, M2.Pru9I	GCCGT	2, 4	m4C	ACCTGCCGTN[NNNN]		
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M2.ClaSBI, M1.Fli6794IV, M2.HphI, M2.Ngo3502II, M1.Ngo6140II, M1.NgoAXVI, M1.NgoBVIII, M1.NgoFA19I, M2.NgoMVIII, M2.PspHUN102I, M1.Ssp10II	GGTGA	5	m6A	GGTGACCTGCNNNN[NNNN]		
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M2.HpyAVI, M1.MnII	CCTC	1	m5C	ACCTGCCTCN[NNNN]		
Acc36I, BfuAI, BspMI, Bvel	ACCT GC	M2.RspA76I	GGGAC	4	m6A	GGGACCTGCNNNN[NNNN]		
AccIII	CAGC TC	M.AbaBs12I, M.Awo1030I, M.Cdu23823I, M.Csa8155II, M.Cun9529I, M.EasL1Dcm, M.Fha57Dcm, M.Eco12761Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.FsaSP1II, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFDcm, M.Kpn928Dcm, M.KpnNIH1Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp40I, M.Pmi1I, M.SpaD2III, M.SpiRp8II, M.SptADcm, M.Tsu2489III	CCWGG	1, 5	m5C	CAGCTCCWGCNNN[NNNN]		
AccIII	CAGC TC	M.Acc65VI, M.SenAZIII	CTKMAG	2, 5	m6A	CTKCAGCTCNNNNNNNN[NNNN]		
AccIII	CAGC	M.AciI	CCGC	1, 3	m5C	CAGCTCCGCNNNN[NNNN]		

AceIII	TC	M.Ade56712I, M.Asp500II, M.AveCB51II, M.BbrUIII, M.Blo30698I, M.BsuBI, M.CpaYL7III, M.Dba7044I, M.Dfa45958II, M.Eco12581I, M.Eco2012I, M.EcoGIII, M.EcoG089I, M.Eho16691II, M.Hal5920II, M.Lsp6406II, M.MspNR41II, M.Pae490I, M.Pox23570I, M.PstI, M.RpaII, M.Sca3403II, M.Tsp61I, M.XmaII, M.YenWJ, M.YinY228I	CTGCAG	2, 5	m6A	CTGCAGCTCNNNNNNNN[NNNN]
AceIII	CAGC	M.Ama55I, M.CnaPQ2III, M.Rmo1926II, M.SenC1765IV, M.SspGII, M.XorPXII	CGGCCG	1, 6	m5C	CAGCTCGGCCGNN[NNNN]
AceIII	CAGC	M.AplII	CTGCAG	3, 4	m5C	CTGCAGCTCNNNNNNNN[NNNN]
AceIII	CAGC	M.Aql	CYCGRG	1, 6	m5C	CAGCTCGRGNNNN[NNNN], CAGCTCYCGRGNN[NNNN]
AceIII	CAGC	M.Asp588II, M.BstVI, M.Dsp964I, M.DvuII, M.GsuRed1II, M.Hka19301II, M.Mch9957IV, M.MfoCT6II, M.Mhy14884I, M.Mru1279II, M.MsmI, M.Msm8159I, M.Nal45188I, M.Osp807I, M.Pae63I, M.PaeR7I, M.XhoI	CTCGAG	2, 5	m6A	CAGCTCGAGNNNN[NNNN]
AceIII	CAGC	M.AvaV, M.SliSI	GATC	1, 4	m4C	GATCAGCTCNNNNNNNN[NNNN]
AceIII	CAGC	M.BamWS8I, M.CthV, M.EfaRFI, M.SafB902II, M.TacII, M.TcoKWC4III	GCWGC	2, 4	m5C	GCAGCTCNNNNNNNN[NNNN]
AceIII	CAGC	M.DbvI, M.DeeSIV	GCAGC	2, 4	m5C	GCAGCTCNNNNNNNN[NNNN]

AceIII	CAGC TC	M.BceSII	GGWCC	1, 5	m5C	GGWCCAGCTCNNNNNN[NNNN]
AceIII	CAGC TC	M.BcoSlaII, M.MspLW5III, M.Nbr128I, M.Psy18884II, M.Tam77409II, M.TbrGE1II, M.TpaRII	CTCGAG	1, 6	m4C	CAGCTCGAGNNNN[NNNN], CAGCTCTCGAGNN[NNNN]
AceIII	CAGC TC	M.BepI, M.FpsIV, M.Gel16401II, M.Pdi8503I	CGCG	1, 4	m5C	CAGCTCGCGNNNN[NNNN]
AceIII	CAGC TC	M.BsoBI	CYCGRG	1, 6	m4C	CAGCTCGRGNNN[NNNN], CAGCTCYCGRGNN[NNNN]
AceIII	CAGC TC	M.BspI4III, M.Bsp3003I, M.Bsp3004I, M.CalB3III, M.CpeAIII, M.Esu35585I, M.FpsIV, M.Lba2031I, M.Lbo2004I, M.LlaKR2I, M.Lmo1042I, M.Lmo7956II, M.Lsp10393II, M.MspI7I, M.Sau3AI, M.Sba3I, M.Sin395I, M.Spn23FI, M.Sth368I	GATC	1, 4	m5C	GATCAGCTCNNNNNN[NNNN]
AceIII	CAGC TC	M.BstXI	CCANNNN NNTGG	3, 10	m6A	CCAGCTCNNNTGGNN[NNNN]
AceIII	CAGC TC	M.BsuFI, M.CpeAIV, M.MspI	CCGG	1, 4	m5C	CAGCTCCGCGNNNN[NNNN]
AceIII	CAGC TC	M.CabI3497I, M.Dac11109III, M.Dco13527I, M.Djo10085I, M.Hbo11551I, M.Hka19301III, M.Hme33500I, M.HsaRII, M.Hsi33800I, M.HspKM1WII, M.HspNII, M.HtuI, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.MjaI, M.MthZI, M.Nas12278I, M.Rfi3010I, M.SsmMI	CTAG	1, 4	m4C	CAGCTCTAGNNNN[NNNN]

AcclII	CAGC TC	M.Cac8I	GCNNGC	2, 5	m5C	GCCAGCTCNNNNNN[NNNN], CAGCTCGCNNNN[NNNN]
AcclII	CAGC TC	M.CagVI	TCCGGA	1, 6	m6A	CAGCTCCGGANN[NNNN]
AcclII	CAGC TC	M.Cdi630IV, M.Ckr177III, M.CmaLM2II, M.CoclI, M.Fnu4HI, M.Fsp4HI	GCNGC	2, 4	m5C	GCAGCTCNNNNNN[NNNN]
AcclII	CAGC TC	M.CfrBI, M.Sen1735IV, M.SenAnaIV	CCWWGG	1, 6	m4C	CAGCTCCWWGGNN[NNNN]
AcclII	CAGC TC	M.Gmu51329I, M.EcoIA03PDcm, M.RorRI1Dcm, M.Sen8391Dcm, M.ZalSM2II	CCWGG	1, 5	Unkno wn	CAGCTCCWGGNN[NNNN]
AcclII	CAGC TC	M.CoclII, M.Fst4947II, M.TdelII, M.Tph12270I	GGNCC	1, 5	m4C	GGNCAGCTCNNNNNN[NNNN]
AcclII	CAGC TC	M.Cph17I, M.Cph4II	CTYCAG	5	m6A	CTYCAGCTCNNNNNN[NNNN]
AcclII	CAGC TC	M.CviQIII, M.CviSIII, M.Hpy299XI, M.Hpy30II, M.Hpy32X, M.Hpy99VII, M.HpyAX, M.HpyFII, M.HpyUM032XVII, M.Tam77409III, M.TaqI, M.TpaRLI, M.TthHB8I	TCCA	1, 4	m6A	CAGCTCGANNNN[NNNN]
AcclII	CAGC TC	M.CviRI, M.Hpy298VIII, M.Hpy299VIII, M.Hpy30VI, M.Hpy32V, M.Hpy57II, M.Hpy66VIII, M.HpyFVI, M.HpyUM032VIII, M.HpyUM037VIII	TGCA	1, 4	m6A	TGCAGCTCNNNNNN[NNNN]
AcclII	CAGC TC	M.Dac11109II, M.Hfo6591I, M.MhuII, M.MjaV, M.MlaZII, M.MspLW5I, M.RsaI	GTAC	1, 4	m4C	GTACAGCTCNNNNNN[NNNN]
AcclII	CAGC	M.Ddel	CTNAG	1, 5	m5C	CTCAGCTCNNNNNN[NNNN]

	TC						CAGCTCAGNNNNN[NNNN], CAGCTCTNAGNNN[NNNN]
AceIII	CAGC TC	M.DfeI, M.Fnu419II, M.Pen113II	CTNNAG	2, 5	m6A		CTNCAGCTCNNNNNN[NNNN], CAGCTCNAGNNN[NNNN]
AceIII	CAGC TC	M.Esp1396I	CCANNN NTGG	3, 9	m6A		CCAGCTCNTGGNNN[NNNN]
AceIII	CAGC TC	M.FatI	CATG	1, 4	m5C		CAGCTCATCNNNN[NNNN]
AceIII	CAGC TC	M.FisAW1I, M.Thal, M.Tph12270II	CGCG	1, 4	m4C		CAGCTCGCGNNN[NNNN]
AceIII	CAGC TC	M.FnoB1II, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VII, M.HpyFV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspIDRI, M.TneBII	CCGG	1, 4	m4C		CAGCTCCGGNNN[NNNN]
AceIII	CAGC TC	M.Gel16401IV	CCAG	2	m5C		CCAGCTCNNNNNN[NNNN]
AceIII	CAGC TC	M.Hka19301IV	TCGCGA	2, 5	m4C		CAGCTCGCGANNN[NNNN]
AceIII	CAGC TC	M.Hme33500II	HGCWGCK	3	m4C		HGCAGCTCNNNNNN[NNNN]
AceIII	CAGC TC	M.Hpy188I, M.Hpy298IX, M.Hpy299V, M.Hpy66V, M.HpyUM032V	TCNGA	1, 5	m6A		CAGCTCNGANNN[NNNN]
AceIII	CAGC TC	M.Hpy30VIII, M.Hpy32VIII, M.Hpy57III, M.Hpy66VI, M.Hpy99XVIII, M.HpyFVII, M.HpyPVIII	TCNNGA	1, 6	m6A		CAGCTCNNGANNN[NNNN]

AceIII	CAGC TC	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.MbuI	CTNAG	1, 5	m4C	CTCAGCTCNNNNNNN[NNNN], CAGCTCAGNNNNN[NNNN], CAGCTCTNAGNNN[NNNN]
AceIII	CAGC TC	M.Hpy57X	CTRYAG	2, 5	m6A	CTRCAGCTCNNNNNNN[NNNN]
AceIII	CAGC TC	M.Hpy99IV, M.HpyFORFX	CCNNGG	1, 6	m4C	CAGCTCCNNGGNN[NNNN]
AceIII	CAGC TC	M.HpyD27I	CCTC	1, 3	m5C, m6A	CAGCTCCTCNNNNN[NNNN]
AceIII	CAGC TC	M.Kpn55II, M.Mha183III	CTTCAG	2, 5	m6A	CTTCAGCTCNNNNNNN[NNNN]
AceIII	CAGC TC	M.Lph145I, M.Lph93I	TCWCC	1	m6A	CAGCTCWCCNNNN[NNNN]
AceIII	CAGC TC	M.Lph145II, M.Lph93II	CYAG	3	m6A	CCAGCTCNNNNNNN[NNNN]
AceIII	CAGC TC	M.Lph15I	GCAG	3	m6A	CCAGCTCNNNNNNN[NNNN]
AceIII	CAGC TC	M.Maf25III, M.Mbo26III, M.Mbo30III, M.MboBCG21I, M.MkaDI, M.Mtu22103III, M.Mtu22115II, M.Mtu26105II, M.Mtu28I, M.Mtu37004II, M.MtuF1I, M.MtuHIII, M.MtuJKIII	CTCCAG	2, 5	m6A	CTCCAGCTCNNNNNNN[NNNN], CAGCTCCAGNNNN[NNNN]
AceIII	CAGC TC	M.MjaII	GGNCC	1, 5	m5C	GGNCCAGCTCNNNNNNN[NNNN]
AceIII	CAGC TC	M.Msp1II	CCTGG	1	m4C	CAGCTCCTGGNNN[NNNN]

AceIII	CAGC TC	M.Msp42II		CCWGG	1, 5	m4C	CAGCTCCWGGNNN[NNNN]
AceIII	CAGC TC	M.NheI		GCTAGC	1, 6	m4C	GCTAGCAGCTCNNNNNNNN[NNNN]
AceIII	CAGC TC	M.Nme18II		TCTGG	2, 4	m5C	CAGCTCTGGNNNN[NNNN]
AceIII	CAGC TC	M.NmeMC58II		CCAGA	1	m5C	CAGCTCCAGANNN[NNNN]
AceIII	CAGC TC	M.Nsp209III, M.Nsp51109II		CTCGAG	1, 6	m5C	CAGCTCGAGNNNN[NNNN], CAGCTCTCGAGNN[NNNN]
AceIII	CAGC TC	M.Pam7686I		CCATGG	1, 6	m5C	CAGCTCCATGGNN[NNNN]
AceIII	CAGC TC	M.Pfa23572II		CCATCG	1, 6	m5C	CAGCTCGATCGNN[NNNN]
AceIII	CAGC TC	M.Pla5I		CTAG	1, 4	Unkno wn	CAGCTCTAGNNNN[NNNN]
AceIII	CAGC TC	M.RpaIV		CTCGAG	3, 4	m5C	CAGCTCGAGNNNN[NNNN]
AceIII	CAGC TC	M.Sam53668II		CCATGG	1, 6	Unkno wn	CAGCTCCATGGNN[NNNN]
AceIII	CAGC TC	M.Sen1781III		CCWWGG	1, 6	Unkno wn	CAGCTCCWGGNN[NNNN]
AceIII	CAGC TC	M.SmoLV1, M.Spn23FII, M.Spn556I, M.Spn7465I, M.SpnRI, M.XbaI		TCTAGA	1, 6	m6A	CAGCTCTAGANNN[NNNN]
AceIII	CAGC TC	M.SonII		GCANNNN GTC	3, 9	m6A	GCAGCTCGTCNNNN[NNNN]

AceIII	CAGC TC	M.TdeI		CTCTTC	1, 2	m4C, m6A	CAGCTCTTCNNNN[NNNN], CAGCTCTCTTCNN[NNNN]
AceIII	CAGC TC	M.TspRI		CASTG	1, 5	m5C	CAGCTCASTGNNNN[NNNN]
AceIII	CAGC TC	M1.AgoG2I, M2.AgoG2I		CCAGT	1, 4	m4C	CAGCTCCAGTNNN[NNNN]
AceIII	CAGC TC	M1.BspACI		CCGC	1	m5C	CAGCTCCGNNNN[NNNN]
AceIII	CAGC TC	M1.HphI		TCAGC	2	m5C	CAGCTCAGCNNNN[NNNN]
AceIII	CAGC TC	M1.Hpy32II		TCYYC	1	m6A	CAGCTCYTCNNNN[NNNN]
AceIII	CAGC TC	M1.HpyAII, M1.MboII, M1.NcuI		TCTTC	1	m6A	CAGCTCTTCNNNN[NNNN]
AceIII	CAGC TC	M1.LplAG30I		GCTTC	1	m5C	CAGCTCCTTCNNNN[NNNN]
AceIII	CAGC TC	M1.Sgr13350III, M2.Sgr13350III		GCTCTTC	2, 3	m4C, m6A	CAGCTCTTCNNNN[NNNN]
AceIII	CAGC TC	M2.ClaSBI, M1.Fli6794IV, M2.HphI, M2.Ngo3502II, M1.Ngo6140II, M1.NgoAXVI, M1.NgoBVIII, M1.NgoFA19I, M2.NgoMVIII, M2.PspHUN102I, M1.Ssp10II		TCAGC	1	m6A	CAGCTCAGCNNNN[NNNN]
AceIII	CAGC TC	M2.Hpy32II, M2.HpyAII, M2.MboII, M2.NcuI		TCTTC	2	m4C	CAGCTCTTCNNNN[NNNN]
AceIII	CAGC TC	M2.HpyAVI, M1.MnlI		CCTC	1	m5C	CAGCTCCTCNNNN[NNNN]

Bbr7I	GAAG AC	M.AatII, M.Lsp6406III	GACGTC	2, 5	m6A	GAAGACGTCNNNN[NNNN]
Bbr7I	GAAG AC	M.AbaBs12I, M.Awo1030I, M.Cdu23823I, M.Csa8155II, M.Cun9529I, M.Eas11Dcm, M.Eba57Dcm, M.Ecl34399Dcm, M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.TsaSP1II, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFDcm, M.Kpn928Dcm, M.KpnNIH1Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I, M.Pmi1I, M.SpaD2III, M.SplRp8II, M.SptADcm, M.Tsu2489III	CCWGG	1, 5	m5C	CCWGGGAAGACNNNNNN[NNNN], GAAGACCGGGNNV[NNNN]
Bbr7I	GAAG AC	M.AbrSp7I, M.Apa101655II, M.Ath21654I, M.BdiNK6I, M.Bel76I, M.Bsp460I, M.BspJG11I, M.BspLA11I, M.BspRAC04II, M.CagI, M.CerMI, M.CerNAI, M.CnaPQ2I, M.Csp16704III, M.CspNS6I, M.CviBI, M.CviQVI, M.Eli8509I, M.Fnu419I, M.Fsp71I, M.HbaI, M.HhaII, M.Hinfi, M.HinHia1I, M.Hpy298IV, M.Hpy299IV, M.Hpy32IV, M.Hpy57IV, M.Hpy66IV, M.Hpy99IX, M.HpyAIV, M.HpyFIV, M.HpyUM032IV, M.HpyUM037IV, M.Mbo45III, M.Msp10I, M.Msp73BI, M.MspAL11I, M.MspLW5II,	GANTC	2, 4	m6A	GAAGACTCNNNN[NNNN]

Bbr7I		M.NhaXII, M.NpeUS61I, M.OspHL35I, M.Pin17395II, M.Pla5III, M.PspLM6I, M.RbaNRL2I, M.Rgi145II, M.Rmo1926I, M.RpaII, M.RsalI, M.RshII, M.Rsp8I, M.SmeC3I, M.SmoLV, M.Sna24252I, M.SscLII, M.SspM41II, M.SstE37II, M.ThalII, M.Xsp91I, M.ZmoCP4I					
Bbr7I	GAAG AC	M.Ace65VI, M.SenAZIII	CTIKMAG	2, 5	m6A	CTGAAGACNNNNNNNN[NNNN]	
Bbr7I	GAAG AC	M.AciI	CCGC	1, 3	m5C	GCGGAAGACNNNNNNNN[NNNN], GAAGACCGCNGNN[NNNN]	
Bbr7I	GAAG AC	M.AflIII	ACRYGT	2, 5	m4C	GAAGACRYGTNNN[NNNN]	
Bbr7I	GAAG AC	M.AgsI	TTSA	1, 5	m6A	TTGAAGACNNNNNNNN[NNNN]	
Bbr7I	GAAG AC	M.Ama55I, M.CnaPQ2III, M.Rmo1926II, M.SenC1765IV, M.SspG1I, M.XorPXII	CGGCGG	1, 6	m5C	GCGCCGAAGACNNNNNNNN[NNNN], GAAGACGGCCGNN[NNNN]	
Bbr7I	GAAG AC	M.Aql	CYCGRG	1, 6	m5C	CYCGRGAAGACNNNNNNNN[NNNN], GAAGACYCGRGNN[NNNN]	
Bbr7I	GAAG AC	M.AvalX, M.Cfr10I, M.Nme18VI, M.VchAI, M.VchO395I	RCCGGY	2, 5	m5C	GAAGACCGGYNNN[NNNN]	
Bbr7I	GAAG AC	M.BcoSlacII, M.MspLW5III, M.Nbr128I, M.Psy18884II, M.Tam774D9II, M.TbrGE1II, M.TpaRLII	CTCGAG	1, 6	m4C	CTCGAGAAGACNNNNNNNN[NNNN], GAAGACTCGAGNN[NNNN]	
Bbr7I	GAAG AC	M.BepI, M.FpsIV, M.Gel16401III, M.Pdl8503I	CGCG	1, 4	m5C	GCGGAAGACNNNNNNNN[NNNN], GAAGACCGCNGNN[NNNN]	

Bbr7I	GAAG AC	M.BsoBI	CYCGRG	1, 6	m4C	CYCGRGAAGACNNNNNN[NNNN], GAAGACYCGRGNN[NNNN]
Bbr7I	GAAG AC	M.BsuFI, M.CpeAIV, M.MspI	CCGG	1, 4	m5C	CCGGAAGACNNNNNN[NNNN], GAAGACCGGNNNN[NNNN]
Bbr7I	GAAG AC	M.CabI3497I, M.Dac111109III, M.Dco13527I, M.Djo10085I, M.Hbo11551I, M.Hka19301III, M.Hme33500I, M.HsaR1I, M.Hsi33800I, M.HspKM1VII, M.HspNII, M.Htul, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.Mjal, M.MthZI, M.Nas12278I, M.Rfi3010I, M.SsmMI	CTAG	1, 4	m4C	CTAGAAGACNNNNNN[NNNN], GAAGACTAGNNNN[NNNN]
Bbr7I	GAAG AC	M.CagVI, M.Lsp6406I, M.Rsp7740I, M.RspPBTs2II, M.VbaLP2AI	TTCGAA	1, 6	m6A	TTCGAAAGACNNNNNN[NNNN]
Bbr7I	GAAG AC	M.CagVI	TCCGGA	1, 6	m6A	TCCGGAAGACNNNNNN[NNNN]
Bbr7I	GAAG AC	M.CagVIII	ACGGGT	2, 5	m4C	GAAGACGGGTNNNN[NNNN]
Bbr7I	GAAG AC	M.CfrBI, M.Sen1735IV, M.SenAnaIV	CCWWGG	1, 6	m4C	CCWWGGAAGACNNNNNN[NNNN], GAAGACCCWWGGNN[NNNN]
Bbr7I	GAAG AC	M.Cmu51329I, M.EcoIA03PDcm, M.RorR1Dcm, M.Sen8391Dcm, M.ZalSM2II	CCWGG	1, 5	Unkno wn	CCWGGGAAGACNNNNNN[NNNN], GAAGACCCWWGGNN[NNNN]
Bbr7I	GAAG AC	M.CviQIII, M.CviSIII, M.Hpy299XI, M.Hpy30II, M.Hpy32X, M.Hpy99VII, M.HpyAX, M.HpyFII, M.HpyUM032XVII, M.Tam77409III, M.TaqI, M.TpaRLI, M.TthHB8I	TCCA	1, 4	m6A	TCGAAGACNNNNNN[NNNN]

Bbr7I	GAAG AC	M.DdeI		CTNAG	1, 5	m5C	CTNAGAAAGACNNNNNN[NNNN], GAAGACTNAGNN[NNNN]
Bbr7I	GAAG AC	M.DfeI, M.Fnu419I, M.Pen113II		CTNNAG	2, 5	m6A	CTGAAGACNNNNNN[NNNN]
Bbr7I	GAAG AC	M.FatI		CATG	1, 4	m5C	CATGAAGACNNNNNN[NNNN], GAAGACATGNNNN[NNNN]
Bbr7I	GAAG AC	M.FisAW1I, M.Thal, M.Tph12270II		CGCG	1, 4	m4C	CGCGAAGACNNNNNN[NNNN], GAAGACGCGNNNN[NNNN]
Bbr7I	GAAG AC	M.FnoB1II, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VIII, M.HpyFV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspJDRI, M.TneBII		CCGG	1, 4	m4C	CCGGAAGACNNNNNN[NNNN], GAAGACCGGNNNN[NNNN]
Bbr7I	GAAG AC	M.FpsIII		CTGAAG	5	m6A	CTGAAGACNNNNNN[NNNN]
Bbr7I	GAAAG AC	M.GsuRedIIIV, M.Osp807II, M.Tth111I		GACNNNG TC	2, 8	m6A	GAAAGACNNNGTCN[NNNN]
Bbr7I	GAAG AC	M.Hka1930IIV		TCGCGA	2, 5	m4C	TCGCGAAGACNNNNNN[NNNN]
Bbr7I	GAAG AC	M.Hpy188I, M.Hpy298IX, M.Hpy299V, M.Hpy66V, M.HpyUM032V		TCNGA	1, 5	m6A	TCNGAAGACNNNNNN[NNNN]
Bbr7I	GAAG AC	M.Hpy298II, M.Hpy299II, M.Hpy66II, M.HpyUM032II, M.HpyUM037II		ACNGT	2, 4	m4C	GAAGACNGTNNNN[NNNN]
Bbr7I	GAAG AC	M.Hpy30VIII, M.Hpy32VIII, M.Hpy57III, M.Hpy66VI, M.Hpy99XVIII, M.HpyFVII, M.HpyPVIII		TCNNGA	1, 6	m6A	TCNNGAAGACNNNNNN[NNNN]

Bbr7I	GAAG AC	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.Mhul	CTNAG	1, 5	m4C	CTNAGAAAGACNNNNNNNN[NNNN], GAAGACTNAGNNN[NNNN]
Bbr7I	GAAG AC	M.Hpy99IV, M.HpyFORFX	CCNNGG	1, 6	m4C	CCNNGGAAGACNNNNNNNN[NNNN], GAAGACCNNGGNN[NNNN]
Bbr7I	GAAG AC	M.Hpy99XI	ACGT	2, 3	m5C	GAAGACGTNNNNN[NNNN]
Bbr7I	GAAG AC	M.HpyD27I	GAGG	2, 4	m6A, m5C	GAAGACCTCNNNN[NNNN], GAGGAAGACNNNNNNNN[NNNN]
Bbr7I	GAAG AC	M.Kpn555II, M.Mha183III	CTGAAG	2, 5	m6A	CTGAAGACNNNNNNNN[NNNN]
Bhr7I	GAAG AC	M.I.ph145I, M.I.ph93I	GGWGA	5	m6A	GGWGAAGACNNNNNNNN[NNNN]
Bbr7I	GAAG AC	M.MlyI, M.Pen113I	GASTC	2, 4	m6A	GAAGACTCNNNNN[NNNN]
Bbr7I	GAAG AC	M.Msp1II	GCAGG	5	m4C	GAAGACCTGCNNNN[NNNN], CCAGGAAGACNNNNNNNN[NNNN]
Bbr7I	GAAG AC	M.Msp42II	CCWGG	1, 5	m4C	CCWGGGAAGACNNNNNNNN[NNNN], GAAGCCWGGNN[NNNN]
Bbr7I	GAAG AC	M.Nme18III	CCAGA	2, 4	m5C	CCAGAAGACNNNNNNNN[NNNN]
Bhr7I	GAAG AC	M.NmeDI	VCCGGY	2, 5	m5C	GAAGACCGGRNNN[NNNN]
Bbr7I	GAAG AC	M.NmeMC58III	CCAGA	1	m5C	TCTGGAAGACNNNNNNNN[NNNN], GAAGACCAGANN[NNNN]
Bbr7I	GAAG AC	M.Nsp209III, M.Nsp51109II	CTCGAG	1, 6	m5C	CTCGAGAAGACNNNNNNNN[NNNN], GAAGACTCGAGNN[NNNN]

Bbr7I	GAAG AC	M.NspI, M.NspHI	RCATGY	2, 5	m5C	GAAGACATGYNNNN[NNNN]
Bbr7I	GAAG AC	M.Pam7686I	CCATGG	1, 6	m5C	CCATGGGAAGACNNNNNNNN[NNNN], GAAGACCATGGNN[NNNN]
Bbr7I	GAAG AC	M.Pfa23572II	CGATCG	1, 6	m5C	CGATCGAAGACNNNNNNNN[NNNN], GAAGACGATCGNN[NNNN]
Bbr7I	GAAG AC	M.Pla5I	CTAG	1, 4	Unkno wn	CTAGAAGACNNNNNNNN[NNNN], GAAGACTAGNNNN[NNNN]
Bbr7I	GAAG AC	M.Sam53668II	CCATGG	1, 6	Unkno wn	CCATGGGAAGACNNNNNNNN[NNNN], GAAGACCATGGNN[NNNN]
Bhr7I	GAAG AC	M.Sen1781III	CCWWGG	1, 6	Unkno wn	CCWWGGAAGACNNNNNNNN[NNNN], GAAGACCCWWGGNN[NNNN]
Bbr7I	GAAG AC	M.SmoLVI, M.Spn23FII, M.Spn556I, M.Spn7465I, M.SpnRI, M.XbaI	TCTAGA	1, 6	m6A	TCTAGAAGACNNNNNNNN[NNNN]
Bbr7I	GAAG AC	M.SonIII	GACNNNN TGC	2, 8	m6A	GAAGACNNNTTGC[NNNN]
Bbr7I	GAAG AC	M.TdelI	CTCTTC	1, 2	m4C, m6A	GAAGAGAAGACNNNNNNNN[NNNN], GAAGACTCTTCNN[NNNN]
Bbr7I	GAAG AC	M.TspRI	CASGT	1, 5	m5C	CASGTGAAGACNNNNNNNN[NNNN], GAAGACASTCGNN[NNNN]
Bhr7I	GAAG AC	M.XmnI	GAANNNN TTC	2, 9	m6A	GAAGACNTTCCNN[NNNN]
Bbr7I	GAAG AC	M1.AgoG2I, M2.AgoG2I	CCAGT	1, 4	m4C	ACTGGAAGACNNNNNNNN[NNNN], GAAGACTGGNNNN[NNNN], GAAGACCATGNNN[NNNN]
Bbr7I	GAAG	M1.BceSIII	ACGGC	2	m4C	GAAGACGGCNNNN[NNNN]

	AC							
Bbr7I	GAAG AC	M1.BspACI		CCGC	1	m5C	GCGGAAGACNNNNNN[NNNN], GAAGACCGCNNNN[NNNN]	
Bbr7I	GAAG AC	M1.Dsh12II, M2.Dsh12II		GACGGC	3, 4	m4C	GAAGACGGCNNNN[NNNN]	
Bbr7I	GAAG AC	M1.HphI		GGTGA	4	m5C	GGTGAAGACNNNNNN[NNNN]	
Bbr7I	GAAG AC	M1.LpIAG30I		GAAAG	5	m5C	GAAGACCTTCNNN[NNNN], GAAGGAAGACNNNNNN[NNNN]	
Bbr7I	GAAG AC	M2.BceSIII, M1.Pru9I, M2.Pru9I		ACGGC	2, 4	m4C	GAAGACGGCNNNN[NNNN]	
Bbr7I	GAAG AC	M2.BfiI		ACTGGG	2	m4C	GAAGACTGGGNNN[NNNN]	
Bbr7I	GAAG AC	M2.BfuAI		ACCTGC	2	m5C	GAAGACCTGCNNN[NNNN]	
Bbr7I	GAAG AC	M2.ClaSBI, M1.Fli6794IV, M2.HphI, M2.Ngo3502II, M1.Ngo6140II, M1.NgoAXVI, M1.NgoBVIII, M1.NgoFA19I, M2.NgoMVIII, M2.PspHUN102I, M1.Ssp10II		GCTGA	5	m6A	GCTGAAGACNNNNNN[NNNN]	
Bbr7I	GAAG AC	M2.HgaI, M1.LlaII		GACGC	3	m5C	GAAGACGCNNNNN[NNNN]	
Bbr7I	GAAG AC	M2.HpyAVI, M1.MnlI		GAGG	4	m5C	GAAGACCTCNNNN[NNNN], GAGGAAGACNNNNNN[NNNN]	
Bbr7I	GAAG AC	M2.Nme579IV, M2.NmeMC58II		GACGC	2	m6A	GAAGACGCNNNNN[NNNN]	
BbsI, BbvII,	GAAG	M.AbaBs12I, M.Awo1030I, M.Cdu23823I,		CCWGG	1, 5	m5C	CCWGGGAAGACNN[NNNN]	

Bbv16II, BpiI, BpuAI, Bsc9I, BspBS3I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	AC	M.Csa8155II, M.Cun9529I, M.Eas11Dcm, M.Eba57Dcm, M.Ecl34399Dcm, M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.EsaSP1II, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFDcm, M.Kpn928Dcm, M.KpnNIH11Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I, M.Pmi1I, M.SpaD2III, M.SplRp8II, M.SptADcm, M.Tsu2489III				
BbsI, BbvII, BpiI, BpuAI, Bsc9I, BspBS3I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.AbrSp7I, M.Apa101655II, M.Ath21654I, M.BdiNK6I, M.Bel76I, M.Bsp460I, M.BspJKG1II, M.BspLA1II, M.BspRAC04II, M.CagI, M.CcrMI, M.CcrNAL, M.CnaPQ2I, M.Csp16704III, M.CspNS6I, M.CviBI, M.CviQVI, M.Eli8509I, M.Fnu419I, M.Fsp71I, M.HbaI, M.HhaII, M.HinI, M.HinHia1I, M.Hpy298IV, M.Hpy299IV, M.Hpy32IV, M.IIpy57IV, M.IIpy66IV, M.Hpy99IX, M.HpyAIV, M.HpyFIV, M.HpyUM032IV, M.HpyUM037IV, M.Mbo45II, M.Msp10I, M.Msp73BI, M.MspAL1I, M.MspLW5II, M.NhaXI, M.NpeUS61I, M.OspHL35I, M.Pin17395II, M.Pla5III, M.PspLM6I, M.RbaNRL2I, M.Rgi145II, M.Rmo1926I,	GANTC	2, 4	m6A	GAAGACTC[NNNN]

BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc9II, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.RpaII, M.RsaI, M.RshII, M.Rsp8I, M.SmeC3I, M.SmoLV, M.Sna24252I, M.SscLI, M.SspM41II, M.SteE37II, M.ThalII, M.Xsp91I, M.ZmoCP4I	CTKMAG	2, 5	m6A	CTGAAGACNN[NNNN]
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc9II, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.AciI	GCGG	2, 4	m5C	GCGGAAGACNN[NNNN]
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc9II, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.AgsI	TTSA	1, 5	m6A	TTGAAGACNN[NNNN]

BstBS32I, BstTS5I, DstV2I	GAAG AC	M.Ama55I, M.CnaPQ2III, M.Rmo1926II, M.SenC1765IV, M.SspG1I, M.XorPXII	CGGCCG	1, 6	m5C	CGGCCGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.AquI	CYCGRG	1, 6	m5C	CYCGRGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.BcoSlaII, M.MspLW5III, M.Nbr128I, M.Psy18884II, M.Tam77409II, M.TbrGE1II, M.TpaRIII	CTCGAG	1, 6	m4C	CTCGAGAAGACNN[NNNN]	

BbsI, Bbv16II, BpuAI, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	BbvII, Bpil, Bsc9II,	GAAG AC	M.BepI, M.Pdi8503I	M.FpsIV, M.GelI6401III,	CGCG	1, 4	m5C	CGCGAAGACNN[NNNN]
BbsI, Bbv16II, BpuAI, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	BbvII, Bpil, Bsc9II,	GAAG AC	M.BsoBI		CYCGRG	1, 6	m4C	CYCGRGAAGACNN[NNNN]
BbsI, Bbv16II, BpuAI, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	BbvII, Bpil, Bsc9II,	GAAG AC	M.BsuFI, M.CpeAIV, M.MspI		CCGG	1, 4	m5C	CCGGAAGACNN[NNNN]
BbsI, Bbv16II, Bpil, Bsc9II,		GAAG AC	M.CabI3497I, M.DacI11109III, M.DcoI3527I, M.Djo10085I, M.HboI1551I,		CTAG	1, 4	m4C	CTAGAAGACNN[NNNN]

BpuAI, Bsc91I, DspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I		M.Hka19301III, M.Hme33500I, M.HsaR1I, M.Isi33800I, M.IispKM1WII, M.IispNII, M.HtuI, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.MjaI, M.MthZI, M.Nas12278I, M.RfI3010I, M.SsmMI				
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.CagV, M.Lsp6406I, M.Rsp7740I, M.RspI7TS2II, M.VbaLP2AI	TTCGAA	1, 6	m6A	TTCGAAGACNN[NNNN]
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.CagVI	TCCGGA	1, 6	m6A	TCCGGAAGACNN[NNNN]
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.CfrBI, M.Sen1735IV, M.SenAnaIV	CCWWGG	1, 6	m4C	CCWWGGAAGACNN[NNNN]

BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I							
BbsI, Bbv16II, BpuAI, DspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Cmu51329I, M.RorR1Dcm, M.Sen8391Dcm, M.ZaiSM2II	CCWGG	1, 5	Unkno wn	CCWGGAGACNN[NNNN]	
BbsI, Bbv16II, BpuAI, DspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.CviQIII, M.Hpy30II, M.HpyAX, M.Tam77409III, M.TthHB8I	TCGA	1, 4	m6A	TCGAAGACNN[NNNN]	
BbsI, Bbv16II, BpuAI, DspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.DdeI	CTNAG	1, 5	m5C	CTNAGAAGACNN[NNNN]	

BstBS32I, BstTS5I, DstV2I							
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc9I, BspBS3I, BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.DfeI, M.Fnu419II, M.Pen113II	CTNAG	2, 5	m6A	CTGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc9I, BspBS3I, BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.FatI	CATG	1, 4	m5C	CATGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc9I, BspBS3I, BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.FisAW1I, M.Thal, M.Tph12270II	CGCG	1, 4	m4C	CGCGAAGACNN[NNNN]	

BbsI, BbvII, BpI, BpuAI, Bsc9II, BspBS3II, BspIS4I, BspTS5I4I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.FnoB1II, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VII, M.HpyFV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspJDR1, M.TmcBII	CCGG	1, 4	m4C	CCGGAAGACNN[NNNN]
BbsI, BbvII, BpI, BpuAI, Bsc9II, BspBS3II, BspIS4I, BspTS5I4I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.FpsJIII	CTGAAG	5	m6A	CTGAAGACNN[NNNN]
BbsI, BbvII, BpI, BpuAI, Bsc9II, BspBS3II, BspIS4I, BspTS5I4I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Hka1930IIV	TCGCGA	2, 5	m4C	TCGGAAGACNN[NNNN]
BbsI, BbvII, BpI, BpuAI, Bsc9II, BspBS3II, BspIS4I, BspTS5I4I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Hpy188I, M.Hpy298IX, M.Hpy299V, M.Hpy66V, M.HpyUM032V	TCNGA	1, 5	m6A	TCNGAAGACNN[NNNN]

BpuAI, Bsc91I, DspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I							
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Hpy30VIII, M.Hpy32VIII, M.Hpy57III, M.Hpy66VI, M.Hpy99XVIII, M.HpyFVII, M.HpyPVIII	TCNNGA	1, 6	m6A	TCNNGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.MhuI	CTNAG	1, 5	m4C	CTNAGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Hpy99IV, M.HpyFORFX	CCNNGG	1, 6	m4C	CCNNGAAGACNN[NNNN]	

BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I							
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, DspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Hpy99XI	ACGT	2, 3	m5C	GAAGACGT[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, DspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.HpyD27I	GAGG	2, 4	m6A, m5C	GAGGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, DspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Kpn555II, M.Mha183III	CTGAAG	2, 5	m6A	CTGAAGACNN[NNNN]	

BstBS32I, BstTS5I, DstV2I							
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Lph145I, M.Lph93I	GGWGA	5	m6A	GGWGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.MlyI, M.Pen113I	GASTC	2, 4	m6A	GAAGACTC[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Msp1II	CCAGG	5	m4C	CCAGGAAGACNN[NNNN]	

BbsI, BbvII, BpI, BpuAI, Bsc9I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Msp42II	CCWGG	1, 5	m4C	CCWGAAGACNN[NNNN]
BbsI, BbvII, BpI, BpuAI, Bsc9I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Nme18III	CCAGA	2, 4	m5C	CCAGAAGACNN[NNNN]
BbsI, BbvII, BpI, BpuAI, Bsc9I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.NmeMC58III	TCTGG	5	m5C	TCTGAAGACNN[NNNN]
BbsI, BbvII, BpI, BpuAI, Bsc9I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Nsp209III, M.Nsp51109II	CTCGAG	1, 6	m5C	CTCGAGAAGACNN[NNNN]

BpuAI, Bsc91I, DspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I							
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Pam7686I	CCATGG	1, 6	m5C	CCATGGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Pfa23572II	CGATCG	1, 6	m5C	CGATCGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Pla5I	CTAG	1, 4	Unkno wn	CTAGAAGACNN[NNNN]	

BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Sam53668II	CCATGG	1, 6	Unkno wn	CCATGGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, Bpil, BpuAI, Bsc9II, DspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.Sen1781III	CCWWGG	1, 6	Unkno wn	CCWWGGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, Bpil, BpuAI, Bsc9II, DspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.SmoLVI, M.Spn23FII, M.Spn7465I, M.SpnRI, M.XbaI	TCTAGA	1, 6	m6A	TCTAGAAGACNN[NNNN]	

BstBS32I, BstTS5I, DstV2I	GAAG AC	M.TdelI		GAAGAG	5, 6	m6A, m4C	GAAGAGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M.TspRI		CATG	1, 5	m5C	CASTGAAGACNN[NNNN]	
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M1.AgoG2I, M2.AgoG2I		ACTGG	2, 5	m4C	ACTGAAGACNN[NNNN]	

BbsI, BbvII, Bpil, BpuAI, Bsc9II, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M1.BspACI	GCGG	4	m5C	GCGGAAGACNN[NNNN]
BbsI, BbvII, Bpil, BpuAI, Bsc9II, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M1.HphI	GGTGA	4	m5C	GGTGAAGACNN[NNNN]
BbsI, BbvII, Bpil, BpuAI, Bsc9II, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M1.LplAG30I	GAAGG	5	m5C	GAAGGAAGACNN[NNNN]
BbsI, BbvII, Bpil, BpuAI, Bsc9II, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M2.ClaSBI, M1.Fli6794IV, M2.HphI, M2.Ngo3502II, M1.Ngo6140II, M1.NgoAXVI,	GGTGA	5	m6A	GGTGAAGACNN[NNNN]

BpuAI, Bsc91I, DspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I		M1.NgoBvIII, M1.NgoFA19I, M2.NgoMvIII, M2.PspIUN102I, M1.Ssp10II				
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M2.HgaI, M1.LlaII	GACGC	3	m5C	GAAGACGC[NNNN]
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M2.HpyAVI, M1.MnII	GAGG	4	m5C	GAGGAAGACNN[NNNN]
BbsI, BbvII, Bbv16II, BpiI, BpuAI, Bsc91I, BspBS31I, BspIS4I, BspTS514I, BstBS32I, BstTS5I, BstV2I	GAAG AC	M2.Nme579IV, M2.NmeMC58II	GACGC	2	m6A	GAAGACGC[NNNN]

BspIS4I, DspTS514I, BstBS32I, BstTS5I, BstV2I	CTCTT C	M.AbrSp7I, M.ApaI01655II, M.Δth21654I, M.BuiNK6I, M.Bel76I, M.Bsp460I, M.BspJKG1I, M.BspLA1I, M.BspRAC04I, M.CagI, M.CerMI, M.CerNAI, M.CnaPQ2I, M.Csp16704III, M.CspNS6I, M.CviBI, M.CviQVI, M.EliB509I, M.Fnu419I, M.Fsp71I, M.HbaI, M.HhaII, M.HinI, M.HinHiaI, M.Hpy298IV, M.Hpy299IV, M.Hpy32IV, M.Hpy57IV, M.Hpy66IV, M.Hpy99IX, M.HpyAIV, M.HpyFIV, M.HpyUM032IV, M.HpyUM037IV, M.Mbo45III, M.Msp10I, M.Msp73BI, M.MspAL1I, M.MspLW5II, M.NhaXII, M.NpeUS6I, M.OspIII35I, M.Pni17395II, M.PlaSIII, M.PspLM6I, M.RbaNRL2I, M.Rgi145II, M.Rmo1926I, M.RpaII, M.RsaI, M.RshII, M.Rsp8I, M.SmeC3I, M.SmoLV, M.Sna24252I, M.SseCL1, M.SspM41I, M.SseE37II, M.ThalII, M.Xsp91I, M.ZmoCP4I	GΔNTC	2, 4	m6A	GACTCTTCN[NNN]
Bco5I, BcoI16I, BcoKI, BseZI, Bst6I, Bsu6I, CatIII, Eam1104I, EarI, Ksp632I	CTCTT C	M.AluBI	AGCT	1, 4	m6A	AGCTCTTCN[NNN]

CatHI, Eam1104I, EarI, Ksp632I							
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.AluI, M.GspCY2I	AGCT	2, 3	m5C	AGCTCTTTCN[NNN]	
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.Alw26I	GTCTC	3, 2	m5C, m6A	GTCTCTTTCN[NNN]	
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.AspDUT2II, M.BstYI, M.CagII, M.MpaI757V, M.Mru1279V, M.RsaII, M.XhoII	RGATCY	2, 5	m4C	RGATCTCTTTCN[NNN]	
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.Asp HL3I, M2.Psp320WII	GAGCTC	3, 4	m4C	GAGCTCTTTCN[NNN]	
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.AvaV, M.SliSI	GATC	1, 4	m4C	GATCTCTTTCN[NNN]	

EarI, Ksp632I	CTCTT	M.BanII, M.EcoT38I	GRGCYC	3, 4	m5C	GRGCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	C					
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.BceSII	GGWCC	1, 5	m5C	GGWCCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.BglII	AGATCT	2, 5	m4C	AGATCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.Ble402III, M.Psp10HI	SAGCTS	3, 4	m4C	SAGCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.Bsp14III, M.Bsp3003I, M.Bsp3004I, M.CaIB3III, M.CpeAIII, M.Esu35585I, M.FpsIV, M.Lba2031I, M.Lbo2004I, M.LiaKR2I, M.Lmo1042I, M.Lmo7956II, M.Lsp10393II, M.Msp17I, M.Sau3AI,	GATC	1, 4	m5C	GATCTCTTCN[NNN]

Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.Sba3I, M.Sin395I, M.Spn23FI, M.Sth368I M.CocII, M.Fsi4947II, M.TdeII, M.Tph12270I	GGNCC	1, 5	m4C	GGNCCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.CviJI	RGCB	2, 3	m5C	VGCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.Dac11109II, M.Hfo6591I, M.MhuII, M.MjaV, M.MlaZII, M.Msp1W5I, M.RsaI	GTAC	1, 4	m4C	GTACTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.Esp3I	GGTCTC	4, 3	m5C, m6A	CGTCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.HpyD27I	CCTC	1, 3	m5C, m6A	CCTCTTCN[NNN]

Bco5I, Bco116I, BcoKI, BseZI, Dst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.MjaII		GGNCC	1, 5	m5C	GGNCCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Dst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.MlyI, M.Pen113I		GATC	2, 4	m6A	GACTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatIII, Eam1104I, EarI, Ksp632I	CTCTT C	M.Mma5219II, M.TarMY3I		AGCT	2, 3	m4C	AGCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.MpeII		RGCY	2, 3	m5C	RGCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.NheI		GCTAGC	1, 6	m4C	GCTAGCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT	M.Nme579III, M.NmeSI, M.ScaI		AGTACT	2, 5	m4C	AGTACTCTTCN[NNN]

BcoKI, BseZI, Bst6I, Bsu6I, CatIII, Eam1104I, EarI, Ksp632I	C	M.Ssp6714II				
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatIII, Eam1104I, EarI, Ksp632I	CTCTT C	M.Pae10332II, M2.Pae50071I	M1.Pae50071I,	CAGCTC	3, 4	m4C CAGCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.SacI, M.Sgr13350I		GAGCTC	3, 4	m5C GAGCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.SuaII		RGATCY	2, 5	m5C RGATCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.TacII		GGCCT	4	m4C GGCCTCTTCN[NNN]
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.Tam77409IV		CCRCTC	4	m4C CCRCTCTTCN[NNN]

Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I							
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.TspPM5I	AGGCCT	2, 5	m4C	AGGCCTCTTCN[NNN]	
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M.XmnI	GAANNN TTC	2, 9	m6A	GAANCTCTTCN[NNN]	
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M1.Sgr13350III, M2.Sgr13350III	GCTCTTC	2, 3	m4C, m6A	GCTCTTCN[NNN]	
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M2.BstSEI	GACTC	4	m6A	GACTCTTCN[NNN]	
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M2.Csu24377II, M2.Hpy298VII, M2.Hpy299VI, M2.Hpy30XII, M1.Hpy32VI, M1.Hpy57VI, M2.Hpy66XIV, M1.Hpy99V,	CCTC	3	m6A	CCCTCTTCN[NNN]	

CatHI, Eam1104I, EarI, Ksp632I		M1.HpyAVI, M1.HpyPVI, M2.HpyUM032VI, M.Mdi27140I, M2.MniI				
Bco5I, Bco116I, BcoKI, BseZI, Bst6I, Bsu6I, CatHI, Eam1104I, EarI, Ksp632I	CTCTT C	M2.Eco31I	GGTCTC	4	m5C	GGTCTCTTCN[NNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, Eco044I	GGTC TC	M.AbaBs12I, M.Awo1030I, M.Cdu23823I, M.Csa8155II, M.Cun9529I, M.Eas11Dcm, M.Eba57Dcm, M.Ecl34399Dcm, M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.EsaSP1II, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFDcm, M.Kpn928Dcm, M.KpnNIH1Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I, M.Pmi1I, M.SpaD2III, M.SplRp8II, M.SptADcm, M.Tsu2489III	CCWGG	1, 5	m5C	CCWGGGTCTCN[NNNN], CCWGGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, Eco044I	GGTC TC	M.AciI	GGG	2, 4	m5C	GGGGGTCTCN[NNNN], GGGGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I,	GGTC TC	M.AhdI	GACNNN NGTC	2, 10	m6A	GACNNNNGGTCTCN[NNNN]

EcoA4I, EcoO44I	GGTC	M.Ama55I, M.CnaPQ2III, M.Rmo1926II, M.SenC1765IV, M.SspG1I, M.XorPXII	CGGCCG	1, 6	m5C	CGGCCGGTCTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.AquI	CYCGRG	1, 6	m5C	CYCGRGGTCTCTCN[NNNN], CYCGGGGTCTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.AspU41II, M.BisIII, M.BsuW23II, M.CnaPQ2II, M.Eco06Dcm, M.Eco12581Dcm, M.Eco1655Dcm, M.Eco29787Dcm, M.Eco3024Dcm, M.Eco3317Dcm, M.Eco3325Dcm, M.Eco3740Dcm, M.Eco3911Dcm, M.Eco4211Dcm, M.Eco4465Dcm, M.Eco600Dcm, M.Eco644Dcm, M.Eco86Dcm, M.Eco9001Dcm, M.Eco9387Dcm, M.EcoE1140Dcm, M.EcoE455Dcm, M.EcoGdcm, M.EcoNU14Dcm, M.EcoRII, M.Kpn223III, M.Kpn35657II, M.Kpn36Dcm, M.Kpn555I, M.Kpn62629II, M.KpnAATV, M.Kqu603Dcm, M.Psp345I, M.Sen158Dcm, M.Sen641Dcm, M.SenA46Dcm, M.SenAboDcm, M.TcoKWC4II	CCWGG	2, 4	m5C	CCWGGTCTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC	M.AvaXI, M.Cfr10I, M.Nme18VI, M.VchAI, M.AvaXI, M.Cfr10I, M.Nme18VI, M.VchAI	RCCGGY	2, 5	m5C	RCCGGTCTCTCN[NNNN]

Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	TC	M.Vch0395I	CTCGAG	1, 6	m4C	CTCGAGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.BcoSlaII, M.MspLW5III, M.Nbr128I, M.Psy18884II, M.Tam77409II, M.TbrGE1II, M.TpaRLII				
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.BepI, M.FpsIV, M.Gel16401III, M.Pdi8503I	CGCG	1, 4	m5C	CGCGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Bhy27164IV, M.BstNI, M.EcoDHB4Dcm, M.EcoK54Dcm, M.Gth3570I, M.Hla49239I, M.Mval, M.Pdi8503II, M.Sso30807Dcm, M.Ypel, M.Ype41I, M.YpeShI	CCWGG	2, 4	m4C	CCWGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.BsoBI	CYCGRG	1, 6	m4C	CYCGRGGTCTCN[NNNN], CYCGGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.BsuFI, M.CpeAIV, M.Mspl	CCGG	1, 4	m5C	CCGGGTCTCN[NNNN], CCGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Cab13497I, M.Dac11109III, M.Dco13527I, M.Djo10085I, M.Hbo11551I, M.Hka19301III, M.Hme33500I, M.HsaRII,	CTAG	1, 4	m4C	CTAGGTCTCN[NNNN]

EcoA4I, EcoO44I		M.Hsi33800I, M.HspKM1WII, M.HspNII, M.Htul, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.Mjal, M.MthZI, M.Nas12278I, M.Rfi3010I, M.SsmMI				
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Cdi630III	CCSSGG	2, 5	m4C	CCSSGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Cfr9I, M.SmaI, M.Sma36365III, M.Xca85III, M.XcyI, M.XmaI, M.XveII	CCCGGG	2, 5	m4C	CCCGGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.CfrBI, M.SenI735IV, M.SenAnaIV	CCWWGG	1, 6	m4C	CCWWGGGTCTCN[NNNN], CCWWGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Cmu51329I, M.EcoIA03PDcm, M.RorR1Dcm, M.Sen8391Dcm, M.ZalSM2II	CCWGG	1, 5	Unkno wn	CCWGGGTCTCN[NNNN], CCWGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Cpa18696III, M.Ype19I, M.YpeDodI, M.YpeEDI, M.YpeI9I, M.YpePGI	CCWGG	2, 4	Unkno wn	CCWGGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI,	GGTC TC	M.Ddel	CTNAG	1, 5	m5C	CTNAGGTCTCN[NNNN]

Bsu537I, Eco31I, EcoA4I, EcoO44I							
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.DsaV, M.Ecl18kl, M.Fps VI, M.NlaX, M.SenPI, M.SsoII	CCNGG	2, 4	m5C	CCNGGTCTCN[NNNN]	
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Eco29kl, M.NgoAIII, M.SacII	CCGCCG	2, 5	m5C	CCGCCGTCTCN[NNNN]	
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.EcoNI	CCTNNNN NAGG	2, 10	m4C	CCTNNNNNAGGTCTCN[NNNN]	
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.EfaRFII, M.Mae7806II	CCGG	2, 3	m4C	CCGGTCTCN[NNNN]	
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.FatI	CATG	1, 4	m5C	CATGGTCTCN[NNNN]	
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.FisAW1I, M.Thal, M.Tph12270II	CGCG	1, 4	m4C	CGCGGTCTCN[NNNN]	

BliI736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.FnoB1II, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VIII, M.HpyFV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspIDRI, M.TmeBII	CCGG	1, 4	m4C	CCGGGTCTCN[NNNN], CCGGTCTCN[NNNN]
BliI736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.FpsII, M.HpaII, M.NmeI8II, M.NmeAI	CCGG	2, 3	m5C	CCGGTCTCN[NNNN]
BliI736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Gel16401IV	CTGG	3	m5C	CTGGTCTCN[NNNN]
BliI736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.GsuRed1IV, M.Osp807II, M.Tth111I	GACNNG TC	2, 8	m6A	GACNNGGTCTCN[NNNN]
BliI736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Hpy298II, M.Hpy299II, M.Hpy66II, M.HpyUM032II, M.HpyUM037II	ACNGT	2, 4	m4C	ACGGTCTCN[NNNN]
BliI736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.MhuI	CTNAG	1, 5	m4C	CTNAGGTCTCN[NNNN]
BliI736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC	M.Hpy991V, M.HpyFORFX	CCNNGG	1, 6	m4C	CCNNGGTCTCN[NNNN]

Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	TC	M.HpyAV, M.Nme18IV	GAAGG	3, 4	m6A, m5C	CCNNGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, TC	GGTC	M.HpyD27I	GAGG	2, 4	m6A, m5C	GAGGTCTCTCN[NNNN], GAGGTCTCTCN[NNNN]
Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC	M.Lph145I, M.Lph93I	TCWCC	1	m6A	GGTCTCC[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, TC	GGTC	M.Msp1II	CCAGG	5	m4C	CCAGGTCTCTCN[NNNN], CCAGGTCTCTCN[NNNN]
Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC	M.Msp42II	CCWGG	1, 5	m4C	CCWGGGTCTCTCN[NNNN], CCWGGGTCTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, TC	GGTC	M.Nme18III	TCTGG	2, 4	m5C	TCTGGTCTCTCN[NNNN]

EcoA4I, EcoO44I	GGTC	M.NmeDI		VCCGGY		2, 5	m5C	VCCGGTCTCN[NNNN], RCCGGGTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC							
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.NmeMC58II		TCTGG		5	m5C	TCTGGGTCTCN[NNNN], TCTGGTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Nsp209III, M.Nsp51109II		CTCGAG		1, 6	m5C	CTCGAGGTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Pam7686I		CCATGG		1, 6	m5C	CCATGGGTCTCN[NNNN], CCATGGTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Pfa23572II		CGATCG		1, 6	m5C	CGATCGGTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.Pla5I		CTAG		1, 4	Unkno wn	CTAGGTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC	M.Sam53668II		CCATGG		1, 6	Unkno	CCATGGGTCTCN[NNNN],

Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	TC	M.Sen1781III	CCWWGG	1, 6	Unkno wn	CCATGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC					CCWWGGGTCTCN[NNNN], CCWWGGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.SonIII	GCANNN GTC	3, 9	m6A	GCANNNGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.TdelI	GAAGAG	5, 6	m6A, m4C	GAAGAGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M.TspRI	CASTG	1, 5	m5C	CASTGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M1.AgoG2I, M2.AgoG2I	ACTGG	2, 5	m4C	ACTGGGTCTCN[NNNN], ACTGGGTCTCN[NNNN]
Bli736I, Bsal, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M1.Bag18758II, M2.Bag18758II	CACGGG	3, 5	Unkno wn	CACGGGTCTCN[NNNN]

EcoA4I, EcoO44I	GGTC	M1.BcnI, M2.BcnI, M.NciI	CCSGG	2, 4	m4C	CCSGGTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC					
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M1.BfiI	ACTGGG	5	m4C	ACTGGGTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M1.BspACI	GCGG	4	m5C	GCGGGTCTCN[NNNN], GCGGTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M1.Hpy32II	TCYYC	1	m6A	GGTCTCC[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M1.LplAG30I	GAAGG	5	m5C	GAAGGTCTCN[NNNN], GAAGGTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC TC	M2.BfuAI	GCAGGT	5	m5C	GCAGGTCTCN[NNNN]
Bli736I, BsaI, Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	GGTC	M2.HpyAVI, M1.MnlI	GAGG	4	m5C	GAGGTCTCN[NNNN]

Bso31I, BspTNI, Bsu537I, Eco31I, EcoA4I, EcoO44I	TC					GAGGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.AatII, M.Lsp6406III	GACGTC	2, 5	m6A	GACGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.AflIII	ACRYGT	2, 5	m4C	ACRCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.AhdI	GACNNNN NGTC	2, 10	m6A	GACNNNNCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.Ama55I, M.CnaPQ2III, M.Rmo1926II, M.SenC1765IV, M.SspG1I, M.XorPXII	CGGCCG	1, 6	m5C	CGGCCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.AvaV, M.SliSI	GATC	1, 4	m4C	GATCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.BceSII	GGWCC	1, 5	m5C	GGWCCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.BepI, M.FpsIV, M.GelI6401III, M.Pdi8503I	CGCG	1, 4	m5C	CGCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.Bgl2196III	CGTACG	2, 5	m4C	CGTACGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.BspI4III, M.Bsp3003I, M.Bsp3004I, M.CalB3III, M.CpeAIII, M.Esu35585I, M.FpsIV, M.Lba2031I, M.Lbo2004I, M.LlaKR2I, M.Lmo1042I, M.Lmo7956II, M.Lsp10393II, M.MspI7I, M.Snu3AI, M.Sba3I, M.Sin395I, M.Spn23FI, M.Sth368I	GATC	1, 4	m5C	GATCGTCTCN[NNNN]	
BsmBI, BstGZ53I,	CGTCT	M.CagVII	ACGCGT	2, 5	m4C	ACGCGTCTCN[NNNN]	

Esp3I	C							
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.CocII, M.Fst4947II, M.Tph12270I	M.TdeII,	GGNCC	1, 5	m4C	GGNCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.CviII		RGCB	2, 3	m5C	RGGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.DacI1109II, M.Hfo6591I, M.MjaV, M.MlaZII, M.MspLW5I, M.RsaI	M.MhuII,	GTAC	1, 4	m4C	GTACGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.FisAW1I, M.Thal, M.Tph12270II		CGCG	1, 4	m4C	CGCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.GsuRedIIV, M.Osp807II, M.Tth111I		GACNNNG TC	2, 8	m6A	GACNNCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.Hpy298II, M.Hpy299II, M.HpyUM032II, M.HpyUM037II	M.Hpy66II,	ACNGT	2, 4	m4C	ACCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.Hpy66I, M.Hpy99I		CGWCG	2, 4	m4C	CGWCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.Hpy99XI		ACGT	2, 3	m5C	ACGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.Lph145I, M.Lph93I		TCWCC	1	m6A	CGTCTCC[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.MjaII		GGNCC	1, 5	m5C	GGNCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.NheI		GCTAGC	1, 6	m4C	GCTAGCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.Pfa23572II		CGATCG	1, 6	m5C	CGATCGTCTCN[NNNN]	
BsmBI, BstGZ53I, Esp3I	CGTCT C	M.SonIII		GCANNNN	3, 9	m6A	GCANNNCGTCTCN[NNNN]	

Esp3I	C		GTC			
BsmBI, BstGZ53I, Esp3I	CGTCT C	M1.BceSIII	GCGGT	4	m4C	GCGGTCTCN[NNNN]
BsmBI, BstGZ53I, Esp3I	CGTCT C	M1.Dsh12II, M2.Dsh12II	GCGGTC	3, 4	m4C	GCGGTCTCN[NNNN]
BsmBI, BstGZ53I, Esp3I	CGTCT C	M1.HgaI, M2.LlaI	GCGTC	2	m5C	GCGTCTCN[NNNN]
BsmBI, BstGZ53I, Esp3I	CGTCT C	M1.Hpy32II	TCYYC	1	m6A	GCTCTCC[NNNN]
BsmBI, BstGZ53I, Esp3I	CGTCT C	M2.BceSIII, M1.Pru9I, M2.Pru9I	GCGGT	2, 4	m4C	GCGGTCTCN[NNNN]
BsmBI, BstGZ53I, Esp3I	CGTCT C	M2.HgaI, M1.LlaI	GCGTC	3	m5C	GCGTCTCN[NNNN]
BsmBI, BstGZ53I, Esp3I	CGTCT C	M2.Nme579IV, M2.NmeMC58II	GCGTC	4	m6A	GCGTCTCN[NNNN]
BtgZI	GCGA TG	M.AbaBs12I, M.Awo1030I, M.Cdu23823I, M.Csa8155II, M.Cun9529I, M.EasL1Dcm, M.EbaS7Dcm, M.Ecl34399Dcm, M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.EsaSP1II, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFDcm, M.Kpn928Dcm, M.KpnNIH1Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I, M.PmlI, M.SpaD2III, M.SplRp8II,	CCWGG	1, 5	m5C	CCWGGGATGNNNNNNNNNN[NNNN]

BtgZI	GGGA TG	M.SptADcm, M.Tsu2489III M.AciI	CCGC	1, 3	m5C	GGCGCGATGNNNNNNNNNN[NNNN], CCGCGATGNNNNNNNNNN[NNNN]
BtgZI	GGGA TG	M.Ama55I, M.CnaPQ2III, M.Rmo1926II, M.SenC1765IV, M.SspG1I, M.XorPXII	CGGCCG	1, 6	m5C	CGGCCGCGATGNNNNNNNNNN[NNNN]
BtgZI	GGGA TG	M.AquI	CYCGRG	1, 6	m5C	CYCGRGCGATGNNNNNNNNNN[NNNN]
BtgZI	GGGA TG	M.AsiSI	GCGATCGC	2, 7	m5C	GCGATCGCGATGNNNNNNNNNN[NNNN]
BtgZI	GGGA TG	M.AvaIX, M.Chr10I, M.Nme18VI, M.VchAI, M.VchO395I	RCCGGY	2, 5	m5C	RCCGGCGATGNNNNNNNNNN[NNNN]
BtgZI	GGGA TG	M.AvaV, M.SliSI	GATC	1, 4	m4C	GCGATGATCNNNNNNNN[NNNN]
BtgZI	GGGA TG	M.BamWS8I, M.CthV, M.EfaRFI, M.Saf8902II, M.TacII, M.TcoKWC4III	GCWGC	2, 4	m5C	GCWGGCGATGNNNNNNNNNN[NNNN]
BtgZI	GGGA TG	M.BbvI, M.BceSIV	GCAGC	2, 4	m5C	GCTGCGATGNNNNNNNNNN[NNNN], GCAGCGATGNNNNNNNNNN[NNNN]
BtgZI	GGGA TG	M.BceSII	GGWCC	1, 5	m5C	GCGATGGWCCNNNNNNNN[NNNN]
BtgZI	GGGA TG	M.EcoSlaII, M.MspLW5III, M.Nbr128I, M.Psy18884II, M.Tam77409II, M.TbrGE1II, M.TpaRLII	CTCGAG	1, 6	m4C	CTCGAGCGATCNNNNNNNNNN[NNNN]
BtgZI	GGGA TG	M.BepI, M.FpsIV, M.Gel16401III, M.Pdi8503I	CCCG	1, 4	m5C	CGCGCGATGNNNNNNNNNN[NNNN], CGCGATGNNNNNNNNNN[NNNN]
BtgZI	GGGA TG	M.BfaSIII, M.Ccl10114I, M.CfrB38I, M.CfrH1I, M.EcoGVI, M.EcoKII,	ATGCAT	2, 5	m6A	GCGATGCATNNNNNNNN[NNNN]

		M.Fnu10562II, M.SbaUIV, M.Sbo12419III, M.Sbo268II, M.Scen10384II, M.Sen10708II, M.Sen1080II, M.Sen1175III, M.Sen13IV, M.Sen13311IV, M.Sen1387II, M.Sen1427III, M.Sen14882III, M.Sen15791III, M.Sen158III, M.Sen16I, M.Sen1655III, M.Sen1676III, M.Sen1677III, M.Sen1728II, M.Sen1735III, M.Sen1736IV, M.Sen1764III, M.Sen1766III, M.Sen1781IV, M.Sen1783III, M.Sen18569II, M.Sen1878III, M.Sen1880II, M.Sen1896III, M.Sen1898III, M.Sen1899III, M.Sen1903III, M.Sen1906III, M.Sen1908III, M.Sen1910III, M.Sen1921III, M.Sen1927II, M.Sen195II, M.Sen2050II, M.Sen2064II, M.Sen2069III, M.Sen22462I, M.Sen255II, M.Sen318III, M.Sen3387III, M.Sen3890II, M.Sen403III, M.Sen624III, M.Sen641III, M.Sen7378II, M.Sen8391III, M.Sen8692II, M.SenA46III, M.SenAbatl, M.SenAboII, M.SenAnaIII, M.SenC1736III, M.SenC1765III, M.SenC1808IV, M.SenC1810III, M.SenJIII, M.SenL1II, M.ScenL15III, M.ScenN1660II, M.SenSARA26II, M.SenSPBIII, M.SenTFII				
BtgZI	GCGA TG	M.Bgil, M.Cagl, M.DdeVCDI, M.GsuRedII	GCCNNNN NGGC	2, 10	m4C	GCCNNNNGGCGATGNNNNNNNNNN[NNN N]

BtgZI	GCGA TG	M.BsoBI	CYCGRG	1, 6	m4C	CYCGRGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.BspI4III, M.Bsp3003I, M.Bsp3004I, M.CalB3III, M.CpeAIII, M.Esu35585I, M.FpsIV, M.Lba2031I, M.Lbo2004I, M.LlaKR2I, M.Lmo1042I, M.Lmo7956II, M.Lsp10393II, M.MspI7I, M.Sau3AI, M.Sba3I, M.Sin395I, M.Spn23FI, M.Sth368I	GATC	1, 4	m5C	GCGATGATCNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.BstXI	CCANNNN NNTGG	3, 10	m6A	CCANNGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.BsuFI, M.CpeAIV, M.MspI	CCGG	1, 4	m5C	CCGGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.Cab13497I, M.Dac11109III, M.Dco13527I, M.Djo10085I, M.Hbo11551I, M.Hka19301III, M.Hme33500I, M.HsaR1I, M.Hsi33800I, M.HspKM1VII, M.HspNII, M.Htul, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.Mjal, M.MthZI, M.Nas12278I, M.Rfi3010I, M.SsmMI	CTAG	1, 4	m4C	CTAGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.Cac8I	GCNNGC	2, 5	m5C	GCNNGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.CerNAV	YGCCGGGR	3, 6	m5C	YGCCGGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.Cdi630IV, M.Ckr177III, M.CnaLM2II, M.CoclI, M.Fnu4HI, M.Fsp4HI	GCNGC	2, 4	m5C	GCNGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA	M.CfrBI, M.Sen1735IV, M.SenAnaIV	CCWWGG	1, 6	m4C	CCWWGGCGATGNNNNNNNNNN[NNNN]

BtgZI	TG	M.Cmu51329I, M.EcoIA03PDcm, M.RorK1Dcm, M.Sen8391Dcm, M.ZalSM2II	CCWGG	1, 5	Unkno wn	CCWGGCGATGNNNNNNNNNN[NNNN]
BtgZI	TG	M.CocII, M.Fsi4947II, M.TdeII, M.Tph12270I	GGNCC	1, 5	m4C	GCGATGGNCCNNNNNNNN[NNNN]
BtgZI	TG	M.CsiFSMII, M.MfoCT6IV, M.NgoAIV, M.NgoMIV, M1.Sgr13350II	GCCGGC	2, 5	m5C	GCCGGCGATGNNNNNNNNNN[NNNN]
BtgZI	TG	M.CviJI	RGCB	2, 3	m5C	RGCATGNNNNNNNNNN[NNNN]
BtgZI	TG	M.CviRI, M.Hpy298VIII, M.Hpy299VIII, M.Hpy30VI, M.Hpy32V, M.Hpy57II, M.Hpy66VIII, M.HpyFVI, M.HpyUM032VIII, M.HpyUM037VIII	TGCA	1, 4	m6A	GCGATGCANNNNNNNNN[NNNN]
BtgZI	TG	M.DacI1109II, M.Hfo6591I, M.MhuII, M.MjaV, M.MlaZII, M.MspLW5I, M.RsaI	GTAC	1, 4	m4C	GCGATGTACNNNNNNNN[NNNN]
BtgZI	TG	M.DdeI	CTNAG	1, 5	m5C	CTNAGCGATGNNNNNNNNNN[NNNN]
BtgZI	TG	M.Esp1396I	CCANNN NTGC	3, 9	m6A	CCANGCGATGNNNNNNNNNN[NNNN]
BtgZI	TG	M.FatI	CATG	1, 4	m5C	CATGCGATGNNNNNNNNNN[NNNN]
BtgZI	TG	M.FisAW1I, M.Thal, M.Tph12270II	CGCG	1, 4	m4C	GCGCGCGATGNNNNNNNNNN[NNNN]. GCGCGATGNNNNNNNNNN[NNNN]
BtgZI	TG	M.FnoB1II, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VII	CCGG	1, 4	m4C	CCGGCGATGNNNNNNNNNN[NNNN]

			M.HpyFV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspJDRI, M.TmcBII					
BtgZI	GCGA TG		M.HhaI, M.HinPII, M.Hpy99III, M.HpyAVIII, M.HpyUM032XV	GCGC	2, 3	m5C	GCGCGATGNNNNNNNNNNNNNNNNNN	
BtgZI	GCGA TG		M.HkaI9301IV	TCGCGA	2, 5	m4C	TCGCGATGNNNNNNNNNNNNNNNNNN	
BtgZI	GCGA TG		M.Hmc33500II	MGCWGGCD	5	m4C	MGCWGGCGATGNNNNNNNNNNNNNNNNNN	
BtgZI	GCGA TG		M.Hpy30VIII, M.Hpy32VIII, M.Hpy57III, M.Hpy66VI, M.Hpy99XVIII, M.HpyFVII, M.HpyPVIII	TCNNGA	1, 6	m6A	TCGCGATGNNNNNNNNNNNNNNNNNN	
BtgZI	GCGA TG		M.IIpy30X, M.Hpy57VIII, M.HpyFXI, M.Mhul	CTNAG	1, 5	m4C	CTNAGCGATGNNNNNNNNNNNNNNNNNN	
BtgZI	GCGA TG		M.Hpy99IV, M.HpyFOREX	CCNNGG	1, 6	m4C	CCNNGGCGATGNNNNNNNNNNNNNNNNNN	
BtgZI	GCGA TG		M.HpyD27I	GAGG	2, 4	m6A, m5C	GAGGCGATGNNNNNNNNNNNNNNNNNN	
BtgZI	GCGA TG		M.MjaII	GGNCC	1, 5	m5C	GCGATGGNCCNNNNNNNNNNNNNNNNNN	
BtgZI	GCGA TG		M.Msp1II	CCAGG	5	m4C	CCAGGCGATGNNNNNNNNNNNNNNNNNN	
BtgZI	GCGA TG		M.Msp42II	CCWGG	1, 5	m4C	CCWGGCGATGNNNNNNNNNNNNNNNNNN	
BtgZI	GCGA TG		M.NheI	GCTAGC	1, 6	m4C	GCTAGCGATGNNNNNNNNNNNNNNNNNN], GCGATGCTAGCNNNNNNNNNNNNNNNNNN	
BtgZI	GCGA		M.Nme18V	GCGCGC	2, 5	m5C	GCGCGCGATGNNNNNNNNNNNNNNNNNN	

	TG						
BtgZI	GCGA TG	M.NmcDI		VCCGGY	2, 5	m5C	VCCGGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.NmeMC58III		TCTGG	5	m5C	TCTGGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.Nsp209III, M.Nsp51109II		CTCGAG	1, 6	m5C	CTCGAGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.Nspl, M.NspHI		RCATGY	2, 5	m5C	RCATGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.Pam7686I		CCATGG	1, 6	m5C	CCATGGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.Pfa23572II		CGATCG	1, 6	m5C	CGATCGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.Pla5I		CTAG	1, 4	Unkno wn	CTAGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.Sam53668II		CCATGG	1, 6	Unkno wn	CCATGGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.Sen1781III		CCWWGG	1, 6	Unkno wn	CCWWGGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.SonIII		GACNNNN TGC	2, 8	m6A	GACGGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.TdelI		GAAGAG	5, 6	m6A, m4C	GAAGAGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA TG	M.TspRI		CASTG	1, 5	m5C	CASTGCGATGNNNNNNNNNN[NNNN]
BtgZI	GCGA	M.XcmI		CCANNNN	3, 13	m6A	CCANNNNGCGATGGNNNNNNNN[NNNN]

	TG		NNNNNTG			J
BtgZl	GCGA TG	M.Yen2516l	GCGCGC	2, 5	Unkno wn	GCGCGCGATGNNNNNNNNNN[NNNN]
BtgZl	GCGA TG	M1.AgoGZl, M2.AgoGZl	ACTGG	2, 5	m4C	ACTGCGGATGNNNNNNNNNN[NNNN]
BtgZl	GCGA TG	M1.Bccl, M1.Hpy30XI, M1.Jsp2502I, M2.MmyCVI	GATGG	2	m6A	GCGATGNNNNNNNNNN[NNNN]
BtgZl	GCGA TG	M1.Bhal, M1.Bst19I, M.Fnu10562III, M1.Mbo45IV, M2.Mbo45IV	GATGC	3	m6A	GCGATGCNNNNNNNNNN[NNNN]
BtgZl	GCGA TG	M1.BspACI	GCGG	4	m5C	GCGGCGATGNNNNNNNNNN[NNNN]
BtgZl	GCGA TG	M1.Csp16704IV, M2.Csp16704IV, M.Lsp10393I, M.Mha183II, M.Mha185II, M.Mha807II, M.Mva1312II	GATGC	2, 3	m6A	GCGATGCNNNNNNNNNN[NNNN]
BtgZl	GCGA TG	M1.Hgal, M2.LlaII	GAGGC	4	m5C	GACGCCATGNNNNNNNNNN[NNNN]
BtgZl	GCGA TG	M1.Hpy299X, M2.Hpy32XIII, M1.HpyFXIII, M2.HpyFXIII, M2.HpyUM037VI, M1.McaCI, M2.McaCI, M1.PspHUN102II, M2.PspHUN102II	GATGG	2, 3	m6A	GCGATGNNNNNNNNNN[NNNN]
BtgZl	GCGA TG	M1.LplAG30I	GAAGG	5	m5C	GAAGGCGATGNNNNNNNNNN[NNNN]
BtgZl	GCGA TG	M1.Sgr13350III, M2.Sgr13350III	GAAGAGC	5, 6	m6A, m4C	GAAGAGCGATGNNNNNNNNNN[NNNN]
BtgZl	GCGA TG	M2.Bccl, M2.Hpy299X, M1.Hpy66X,	GATGG	3	m6A	GCGATGNNNNNNNNNN[NNNN]

	TG	M1.HpyC1I, M2.HpyC1I, M1.HpyUM032X, M2.Jsp2502I, M1.MmyCVI					
BtgZI	GCGA TG	M2.BceSIII, M1.Pru9I, M2.Pru9I	ACGGC	2, 4	m4C	ACGGCGATGNNNNNNNNNN[NNNN]	
BtgZI	GCGA TG	M2.BhaI	GATGC	2	m6A	GCGATGCNNNNNNNNNN[NNNN]	
BtgZI	GCGA TG	M2.HpyAVI, M1.MnII	GAGG	4	m5C	GAGGCGATCNNNNNNNNNN[NNNN]	
RleAI	CCCAC A	M.AflIII	ACRYGT	2, 5	m4C	CCCACAYGTNNNNNN[NNN]	
RleAI	CCCAC A	M.AluBI	AGCT	1, 4	m6A	CCCACAGCTNNNNNN[NNN]	
RleAI	CCCAC A	M.ApaI	GGGCCC	3, 4	m5C	GGGCCCACANNNNNNNN[NNN]	
RleAI	CCCAC A	M.AspDUT2II, M.BstVI, M.CagII, M.CpaI, M.Mpa1757V, M.Mru1279V, M.RsaII, M.XhoII	RGATCY	2, 5	m4C	RGATCCCACANNNNNNNN[NNN]	
RleAI	CCCAC A	M.AspWA102I, M.AvaVII, M.BceIII, M.CbeI, M.Ckr177II, M.CmaLM2IV, M.CpeAVII, M.Csp323II, M.PhoI, M.Ssp6714III, M.SuaI	GACC	2, 3	m4C	GGCCACANNNNNNNNNN[NNN]	
RleAI	CCCAC A	M.Asu14238I, M.Bhy27164II, M.CagIX, M.Cje11351II, M.CviAII, M.CviQVII, M.Fli6794III, M.Gli15749I, M.HauII, M.HpyI, M.Hpy166IV, M.Hpy188II, M.Hpy298I, M.Hpy299I, M.Hpy30I, M.Hpy32I, M.Hpy57I, M.Hpy66XI,	CATG	2, 3	m6A	CCCACATGNNNNNNNN[NNN]	

RleAI	CCCAC A	M.Hpy99X, M.HpyAI, M.HpyFI, M.HpyPI, M.HpyUM032I, M.HpyUM037I, M.Lsp48II, M.LspLP1II, M.NlaIII, M.NsaI, M.Sba3II, M.SdeAIII, M.SspT436I, M.ThalIV, M.TvoI	GATC	1, 4	m4C	GATCCCACACANNNNNNNNN[NNN]	
RleAI	CCCAC A	M.BamHI	GGATCC	2, 5	m4C	GGATCCCACACANNNNNNNNN[NNN]	
RleAI	CCCAC A	M.BanII, M.EcoT38I	GRGCTC	3, 4	m5C	GRGCTCCCACACANNNNNNNNN[NNN]	
RleAI	CCCAC A	M.BanLII, M.CpeAII, M.SinI	GGWCC	2, 4	m5C	GGWCCCACACANNNNNNNNN[NNN]	
RleAI	CCCAC A	M.BceSII	GGWCC	1, 5	m5C	GGWCCCCCACACANNNNNNNNN[NNN], GGWCCCACACANNNNNNNNN[NNN]	
RleAI	CCCAC A	M.BglII, M.CagI, M.DdeVCDI, M.GsuRed1I	GCCNNNN NGGC	2, 10	m4C	GCCCACANGGCCNNNNNN[NNN]	
RleAI	CCCAC A	M.BhaII, M.BspRI, M.BsuRI, M.CthVI, M.DthLII, M.HaeIII, M.MthTI, M.NgoAI, M.NgoPII, M.Pod1539I	GGCC	2, 3	m5C	GGCCCACACANNNNNNNNN[NNN]	
RleAI	CCCAC A	M.Bsp14III, M.Bsp3003I, M.Bsp3004I, M.CalB3III, M.CpeAIII, M.Esu35585I, M.FpsIV, M.Lba2031I, M.Lbo2004I, M.LlaKR2I, M.Lmo1042I, M.Lmo7956II, M.Lsp10393II, M.MspI7I, M.Sau3AI, M.Sba3I, M.Sin395I, M.Spn23FI, M.Sth368I	GATC	1, 4	m5C	GATCCCACACANNNNNNNNN[NNN]	
RleAI	CCCAC	M.BstXI	CCANNNN	3, 10	m6A	CCCACANNNNTGGNN[NNN]	

RleAI	A	CCCAC	M.CocII, M.Fst4947II, M.TdeII, M.Tph12270I	GNCC	1, 5	m4C	GGNCCCACANNNNNNNNNN[NNN], GGNCCCACANNNNNNNNNN[NNN], GGCCACANNNNNNNNNN[NNN]
RleAI	A	CCCAC	M.CviII	RGC	2, 3	m5C	VGCCACANNNNNNNNNN[NNN]
RleAI	A	CCCAC	M.DacI1109II, M.Hfo6591I, M.MhuII, M.MjaV, M.MlaZII, M.MspLW5I, M.RsaI	GTAC	1, 4	m4C	GTACCCACANNNNNNNNNN[NNN]
RleAI	A	CCCAC	M.Djo10085II, M.Ssp6803II	GCC	2, 3	Unkno wn	GGCCACANNNNNNNNNN[NNN]
RleAI	A	CCCAC	M.Eco0109I	RGNCCY	3, 5	m5C	RGNNCCCACANNNNNNNNNN[NNN]
RleAI	A	CCCAC	M.EcoVIII, M.HindIII, M.Lad60222I, M.LlaCI	AAGCTT	1, 6	m6A	CCCACAAAGCTTNNNN[NNN]
RleAI	A	CCCAC	M.Esp1396I	GCANNNN NTGG	3, 9	m6A	CCCACANNNTGGNNNN[NNN]
RleAI	A	CCCAC	M.FatI	CATG	1, 4	m5C	CCCACATGNNNNNNNN[NNN]
RleAI	A	CCCAC	M.HgiCI	GGYRCC	2, 5	m5C	GGYRCCCACANNNNNNNNNN[NNN]
RleAI	A	CCCAC	M.Hpy298II, M.Hpy299II, M.Hpy66II, M.HpyUM032II, M.HpyUM037II	ACNGT	2, 4	m4C	CCCACAGTNNNNNNNN[NNN]
RleAI	A	CCCAC	M.HtuII, M.Nas12278II	CATTC	2	m6A	CCCACATTCNNNNNNNN[NNN]
RleAI	A	CCCAC	M.Mae7806III, M.PspPI, M.Sau96I	GNCC	2, 4	m5C	GGNCCCACANNNNNNNNNN[NNN], GGCCACANNNNNNNNNN[NNN]

RleAI	CCCAC A	M.MjaII		GGNCC	1, 5	m5C	GGNCCCCACANNNNNNNNNNNN[NNN], GGNCCACANNNNNNNNNNNN[NNN], GGCCACANNNNNNNNNNNN[NNN]
RleAI	CCCAC A	M.MpeII		RGCY	2, 3	m5C	RGCCACANNNNNNNNNNNN[NNN]
RleAI	CCCAC A	M.NgoAXV, M.NgoMV		GGNNCC	2, 5	m5C	GGNNCCACANNNNNNNNNNNN[NNN], GGNCCACANNNNNNNNNNNN[NNN]
RleAI	CCCAC A	M.NheI		GCTAGC	1, 6	m4C	GCTAGCCACANNNNNNNNNNNN[NNN]
RleAI	CCCAC A	M.NspI, M.NspHI		RCATGY	2, 5	m5C	CCCACATGYNNNNNNNN[NNN]
RleAI	CCCAC A	M.Psp737I		CCCCCD	4	m4C	CCCCACANNNNNNNNNNNN[NNN]
RleAI	CCCAC A	M.SbaUII		CAGCTG	2, 5	m6A	CCCACAGCTGNNNNNN[NNN]
RleAI	CCCAC A	M.SuaII		RGATCY	2, 5	m5C	RGATCCGACANNNNNNNNNN[NNN]
RleAI	CCCAC A	M.TspRI		CASTG	1, 5	m5C	CCCACASTGNNNNNNNN[NNN]
RleAI	CCCAC A	M1.CspNS6II		TCCGCC	5	m5C	TCCGCCACANNNNNNNNNN[NNN]
RleAI	CCCAC A	M2.BstF5I		CATCC	2	m6A	CCCACATCCNNNNNN[NNN]
RleAI	CCCAC A	M4.BstF5I, M1.Cin755I, M2.Cin755I, M1.CmeSR3I, M2.CmeSR3I, M1.Csp12AI, M1.Csp323I, M2.Csp323I, M1.Csp410I,		CATCC	2, 3	m6A	CCCACATCCNNNNNNNN[NNN]

AarI		M2.Csp410I, M.FokI, M1.Msp315I, M2.Msp315I, M.Sau13435I, M.StsI, M1.Tsu2489IV, M2.Tsu2489IV, M3.Tsu2489IV					
	CACCT GC	M.AbaBs12I, M.Awo1030I, M.Cdu23823I, M.Csa8155II, M.Cun9529I, M.EasL1Dcm, M.Eba57Dcm, M.Ecl34399Dcm, M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.EsaSP1II, M.Kpn34618Dcm, M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFDcm, M.Kpn928Dcm, M.KpnNIH1Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I, M.PmiI, M.SpaD2III, M.SplRp8II, M.SptADcm, M.Tsu2489III	CCWGG	1, 5	m5C	CACCTGCCWGG[NNNN]	
	CACCT GC	M.Acc65VI, M.SenAZIII	CTKMAG	2, 5	m6A	CACCTGCACNN[NNNN]	
	CACCT GC	M.AciI	CCGC	1, 3	m5C	CACCTGCGGNN[NNNN], CACCTGCGGCN[NNNN]	
AarI	CACCT GC	M.Ade56712I, M.Asp588III, M.AveCB51II, M.BbrUIII, M.Blo30698I, M.BsuBI, M.CpaYL7III, M.Dba7044I, M.Dfa45958II, M.Eco12581I, M.Eco2012I, M.EcoGIII, M.EcoG089I, M.Eho1669II, M.Hal5920II, M.Lsp6406II, M.MspNR4III, M.Pae490I,	CTCCAG	2, 5	m6A	CACCTGCACNN[NNNN]	

AarI	CACCT	M.Pox23570I, M.PstI, M.RpaII, M.Sca3403II, M.Tsp61I, M.XmnII, M.YenWI, M.YinY228I	CTGCAG	3, 4	m5C	CACCTGCAGNN[NNNN]
	GC					
AarI	CACCT	M.AvaIX, M.Cfr10I, M.Nme18VI, M.VchAI, M.VchO39SI	RCCGGY	2, 5	m5C	CACCTGCCGGY[NNNN]
	GC					
AarI	CACCT	M.AvaV, M.SliSI	GATC	1, 4	m4C	GATCACCTGCNNNN[NNNN]
	GC					
AarI	CACCT	M.BamWS8I, M.CthV, M.EfaRfI, M.Saf8902II, M.TacII, M.TcoKWC4III	GCWGC	2, 4	m5C	CACCTGCWGCN[NNNN]
	GC					
AarI	CACCT	M.RbvI, M.BceSIV	GCAGC	2, 4	m5C	CACCTGCTGCN[NNNN]. CACCTGCAGCN[NNNN]
	GC					
AarI	CACCT	M.BceSII	GGWCC	1, 5	m5C	GGWCCACCTGCNNNN[NNNN]
	GC					
AarI	CACCT	M.BepI, M.FpsIV, M.GelI640I, III, M.Pdi8503I	GC CG	1, 4	m5C	CACCTGCCGCN[NNNN]
	GC					
AarI	CACCT	M.Bsp14III, M.Bsp3003I, M.Bsp3004I, M.CalB3III, M.CpeAIII, M.Esu35585I, M.FpsIV, M.Lba2031I, M.Lbo2004I, M.LlaKR2I, M.Lmo1042I, M.Lmo7956II, M.Lsp10393II, M.MspI7I, M.Sau3AI, M.Sba3I, M.Sin395I, M.Spn23FI, M.Sth368I	GATC	1, 4	m5C	GATCACCTGCNNNN[NNNN]
	GC					
AarI	CACCT	M.BspRAC04I, M.EcaI, M.Eco3936IV	GCTNACC	3, 5	m6A	GGTCACCTGCNNNN[NNNN]
	GC					
AarI	CACCT	M.BstXI	CCANNNN NNTGG	3, 10	m6A	CCACCTGCNTGG[NNNN]
	GC					

AarI	CACCT GC	M.BsuFI, M.CpeAIV, M.MspI	CCGG	1, 4	m5C	CACCTGCCGGN[NNNN]	
AarI	CACCT GC	M.CabI3497I, M.Dac11109III, M.Dco13527I, M.Djo10085I, M.Hbo11551I, M.Hka19301III, M.Hme33500I, M.HsaR1I, M.Hsi33800I, M.HspKM1WII, M.HspNII, M.HtuI, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.MjaI, M.MthZI, M.Nas12278I, M.RfI3010I, M.SsmMI	CTAG	1, 4	m4C	CACCTGCTAGN[NNNN]	
AarI	CACCT GC	M.Cac8I	GCNNGC	2, 5	m5C	CACCTGCNNGC[NNNN]	
AarI	CACCT GC	M.Cdi630IV, M.Ckr177III, M.CmaLM2II, M.CocII, M.Fnu4HI, M.Fsp4HI	GCNGC	2, 4	m5C	CACCTGCNGCN[NNNN]	
AarI	CACCT GC	M.Cmu51329I, M.EcoIA03PDcm, M.RorR1Dcm, M.Sen8391Dcm, M.ZalISM2II	CCWGG	1, 5	Unkno wn	CACCTGCCWGG[NNNN]	
AarI	CACCT GC	M.CocII, M.Fst4947II, M.TdeII, M.Tph12270I	GGNCC	1, 5	m4C	GGNCCACCTGCNNNN[NNNN]	
AarI	CACCT GC	M.CsiFSMII, M.MfoCT6IV, M.NgoAIV, M.NgoMIV, M1.Sgr13350II	GCCGGC	2, 5	m5C	CACCTGCCGGC[NNNN]	
AarI	CACCT GC	M.CviRI, M.Hpy298VIII, M.Hpy299VIII, M.Hpy30VI, M.Hpy32V, M.Hpy57II, M.Hpy66VIII, M.HpyFVI, M.HpyUM032VIII, M.HpyUM037VIII	TGCA	1, 4	m6A	TGCACCTGCNNNN[NNNN], CACCTGCANN[NNNN]	
AarI	CACCT GC	M.Dac11109II, M.Hfo6591I, M.MhuII, M.MjaV, M.MlaZII, M.MspLW5I, M.RsaI	GTAC	1, 4	m4C	GTACACCTGCNNNN[NNNN]	
AarI	CACCT	M.DdeI	CTNAG	1, 5	m5C	CACCTGCTNAG[NNNN]	

AarI	GC	M.DfeI, M.Fnu419II, M.Pen113II	CTNNAG	2, 5	m6A	CACCTGCAGNN[NNNN]
AarI	GC	M.Esp1396I	CCANNNN NTGG	3, 9	m6A	CCACCTGCTGGN[NNNN]
AarI	CACCT GC	M.FatI	CATG	1, 4	m5C	CACCTGCATGN[NNNN]
AarI	CACCT GC	M.FisAW1I, M.Thal, M.Tph12270II	GCGG	1, 4	m4C	CACCTGGCGGN[NNNN]
AarI	CACCT GC	M.FnoB1II, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VII, M.HpyFV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspIDRI, M.TmeBII	CGGG	1, 4	m4C	CACCTGCCGGN[NNNN]
AarI	CACCT GC	M.HgiCI	GGYRCC	2, 5	m5C	GGCACCTGCNNNN[NNNN]
AarI	CACCT GC	M.HhaI, M.HinP1I, M.Hpy99III, M.HpyAVIII, M.HpyUM032XV	GCGC	2, 3	m5C	CACCTGCCGNN[NNNN]
AarI	CACCT GC	M.Hme33500II	HGCWGC	3	m4C	CACCTGCWGC[NNNN]
AarI	CACCT GC	M.Hpy30IX, M.Hpy32IX, M.Hpy57IX, M.Hpy66XIII, M.Hpy8I, M.HpyFIX, M.HpyUM032XI, M.HpyUM037V, M.MjaIV, M.MspRAC08II	GTNNAC	2, 5	m6A	GTNCACCTGCNNNN[NNNN]
AarI	CACCT GC	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.Mhul	CTNAG	1, 5	m4C	CACCTGCTNAG[NNNN]

AarI	CACCT GC	M.Hpy57X		CTRYAG	2, 5	m6A	CACCTGCAGNN[NNNN]
AarI	CACCT GC	M.Hpy99II, M.Tsp45I, M.Tsu2489II		GTSAC	2, 4	m6A	GTCACCTGCNNNN[NNNN]
AarI	CACCT GC	M.HpyD27I		CCTC	1, 3	m5C, m6A	CACCTGCCTCN[NNNN]
AarI	CACCT GC	M.MjaII		GGNCC	1, 5	m5C	GGNCCACCTGCNNNN[NNNN]
AarI	CACCT GC	M.Msp1II		CCTGG	1	m4C	CACCTGCCTGG[NNNN]
AarI	CACCT GC	M.Msp42II		CCWGG	1, 5	m4C	CACCTGCCWGG[NNNN]
AarI	CACCT GC	M.NgoAXV, M.NgoMV		GGNNCC	2, 5	m5C	GGCACCTGCNNNN[NNNN]
AarI	CACCT GC	M.NheI		GCTAGC	1, 6	m4C	GCTAGCACCTGCNNNN[NNNN], CACCTGCTAGC[NNNN]
AarI	CACCT GC	M.Nme18V		GCCCGC	2, 5	m5C	CACCTGCCCGC[NNNN]
AarI	CACCT GC	M.NmeDI		VCCGGY	2, 5	m5C	CACCTGCCGGB[NNNN]
AarI	CACCT GC	M.NmeMC58III		CCAGA	1	m5C	CACCTGCCAGA[NNNN]
AarI	CACCT GC	M.NspI, M.NspHI		RCATGY	2, 5	m5C	CACCTGCATGY[NNNN]
AarI	CACCT GC	M.Pla5I		CTAG	1, 4	Unkno wn	CACCTGCTAGN[NNNN]

AarI	CACCT GC	M.SonIII		GACNNNN TGC	2, 8	m6A	GACCACCTGCNNNN[NNNN]
AarI	CACCT GC	M.TspRI		CASTG	1, 5	m5C	CACCTGCASTG[NNNN]
AarI	CACCT GC	M.Yen2516I		GCGCGC	2, 5	Unkno wn	CACCTGGCGCG[NNNN]
AarI	CACCT GC	M1.AgoG2I, M2.AgoG2I		CCAGT	1, 4	m4C	CACCTGCCAGT[NNNN]
AarI	CACCT GC	M1.BspACI		CCGC	1	m5C	CACCTGGCGCN[NNNN]
AarI	CACCT GC	M1.HgaI, M2.LlaII		GCGTC	2	m5C	CACCTGCGTCN[NNNN]
AarI	CACCT GC	M1.HphI		TCACC	2	m5C	TCACCTGCNNNN[NNNN]
AarI	CACCT GC	M1.LpIAG30I		CCTTC	1	m5C	CACCTGCCTTC[NNNN]
AarI	CACCT GC	M2.BceSIII, M1.Pru9I, M2.Pru9I		GCCGT	2, 4	m4C	CACCTGCCGTN[NNNN]
AarI	CACCT GC	M2.HpyAVI, M1.MnII		CCTC	1	m5C	CACCTGCCTCN[NNNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.AbaBs12I, M.Awo1030I, M.Cdu23823I, M.Csa8155II, M.Cun9529I, M.EasL1Dcm, M.Eba57Dcm, M.Ecl34399Dcm, M.Eco12761Dcm, M.Eco3609Dcm, M.Eco4255Dcm, M.EcoVR50Dcm, M.EsaSP1II, M.Kpn34618Dcm,		CCWGG	1, 5	m5C	CCWGGCTCTTCN[NNN]

		M.Kpn38547Dcm, M.Kpn39I, M.Kpn43816I, M.Kpn6770RFDcm, M.Kpn928Dcm, M.KpnNIH1Dcm, M.KpnNIH10Dcm, M.KpnNIH30Dcm, M.Kva120Dcm, M.Lsp48I, M.Pmi1I, M.SpaD2III, M.SplRp8II, M.SptADcm, M.Tsu2489III	CGGC	1, 3	m5C	GGGGCTCTTCN[NNN], CCGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, Vpak32I	GCTCT TC	M.AciI				
BspQI, LguI, PciSI, SapI, Vpak32I	GCTCT TC	M.AluBI	AGCT	1, 4	m6A	AGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, Vpak32I	GCTCT TC	M.AluI, M.CspCY2I	AGCT	2, 3	m5C	AGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, Vpak32I	GCTCT TC	M.Ama55I, M.CnaPQ2III, M.Rmo1926II, M.SenC1765IV, M.SspGII, M.XorPXII	CGGCGG	1, 6	m5C	CGGGCGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, Vpak32I	GCTCT TC	M.AquI	CYCGRG	1, 6	m5C	CYCGRGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, Vpak32I	GCTCT TC	M.AsiSI	GGGATCGC	2, 7	m5C	GGGATCGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, Vpak32I	GCTCT TC	M.Asp HL3I, M2.Psp320WII	GAGCTC	3, 4	m4C	GAGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, Vpak32I	GCTCT TC	M.AvaIX, M.Cfr10I, M.Nme18VI, M.VchAI, M.VchO395I	RCCGGY	2, 5	m5C	RCCGGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, Vpak32I	GCTCT TC	M.BamWS8I, M.CthV, M.EfaRFL, M.Saf8902II, M.TacII, M.TcoKWC4III	GCWGC	2, 4	m5C	GCWGCCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, Vpak32I	GCTCT TC	M.BanII, M.EcoT38I	GRGCYC	3, 4	m5C	GRGCTCTTCN[NNN]

BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.BbvI, M.BceSIV	GCAGC	2, 4	m5C	GCTGCTCTTCN[NNN], GCAGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.BcoSlaCI, M.MspLW5III, M.Nbr128I, M.Psy18884II, M.Tam77409II, M.TbrGE1II, M.TpaRIII	CTCGAG	1, 6	m4C	CTCGAGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.BepI, M.PspIV, M.Gel16401III, M.Pdi8503I	CCCG	1, 4	m5C	CCGGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.BglII, M.CagI, M.DdeVCDI, M.GsuRed1I	GCCNNNN NGGC	2, 10	m4C	GCCNNNNGGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.Ble402III, M.Psp10HI	SAGCTS	3, 4	m4C	SAGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.DsoBI	CYCGRG	1, 6	m4C	CYCGRGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.BsuFI, M.CpeAIV, M.MspI	CCGG	1, 4	m5C	CCGGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.Cab13497I, M.Dac111109III, M.Dco13527I, M.Djo10085I, M.Hbo11551I, M.Hka19301III, M.Hmc33500I, M.HsaR1I, M.Hsi33800I, M.HspKM1VII, M.HspNII, M.HtuI, M.HvoDSI, M.Lba2029II, M.Mbo45II, M.MjaI, M.MthZI, M.Nas12278I, M.Rfi3010I, M.SsmMI	CTAG	1, 4	m4C	CTAGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.Cac8I	GCNNGC	2, 5	m5C	GCNNGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.Cdi630IV, M.Ckr177III, M.CmaLM2II, M.CocII, M.Fnu4HI, M.Fsp4HI	GCNCG	2, 4	m5C	GCNCGCTCTTCN[NNN]

BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.CfrRI, M.Sen1735IV, M.SenAnaIV	CCWWGG	1, 6	m4C	CCWWGGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.Cmu51329I, M.EcoIA03PDcm, M.RorR1Dcm, M.Sen8391Dcm, M.ZalSM2II	CCWGG	1, 5	Unkno wn	CCWGGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.CsiFSMII, M.MfoCT6IV, M.NgoAIV, M.NgoMIV, M1.Sgr13350II	GCCGGC	2, 5	m5C	GCCGGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.CviJI	RGCB	2, 3	m5C	VGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.Ddel	CTNAG	1, 5	m5C	CTNAGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.FatI	CATG	1, 4	m5C	CATGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.FisAW1I, M.Thal, M.Tph12270II	CGCG	1, 4	m4C	CGCGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.FnoB1II, M.Hal24586I, M.Hpy298V, M.Hpy299IX, M.Hpy30V, M.Hpy32XI, M.Hpy57V, M.Hpy66IX, M.Hpy99VII, M.HpyFV, M.HpyUM032IX, M.HpyUM037IX, M.MjaVI, M.PspIDRI, M.TmeBII	CGGG	1, 4	m4C	CGGGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.HhaI, M.HinP1I, M.Hpy99III, M.HpyAVIII, M.HpyUM032XV	GCGC	2, 3	m5C	GCGCTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.Hme33500II	MGCWGD	5	m4C	MGCWGTCTTCN[NNN]
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M.Hpy30X, M.Hpy57VIII, M.HpyFXI, M.MhuI	CTNAG	1, 5	m4C	CTNAGCTCTTCN[NNN]
BspQI, LguI, PciSI,	GCTCT	M.Hpy99IV, M.HpyFORFX	GCNNGG	1, 6	m4C	GCNNGGCTCTTCN[NNN]

SapI, VpaK32I	TC							
BspQI, LguI, PciSI,	GCTCT	M.HpyD27I		GAGG	2, 4	m6A, m5C	GAGGCTCTTCN[NNN]	
SapI, VpaK32I	TC							
BspQI, LguI, PciSI,	GCTCT	M.Mma5219II, M.TarMY3I		AGCT	2, 3	m4C	AGCTCTTCN[NNN]	
SapI, VpaK32I	TC							
BspQI, LguI, PciSI,	GCTCT	M.MpeII		RGCT	2, 3	m5C	RGCTCTTCN[NNN]	
SapI, VpaK32I	TC							
BspQI, LguI, PciSI,	GCTCT	M.Msp1II		CCAGG	5	m4C	CCAGGCTCTTCN[NNN]	
SapI, VpaK32I	TC							
BspQI, LguI, PciSI,	GCTCT	M.Msp42II		CCWGG	1, 5	m4C	CCWGGCTCTTCN[NNN]	
SapI, VpaK32I	TC							
BspQI, LguI, PciSI,	GCTCT	M.NgoAI		RGCGCT	3, 4	m5C	RGCGCTCTTCN[NNN]	
SapI, VpaK32I	TC							
BspQI, LguI, PciSI,	GCTCT	M.NheI		GCTAGC	1, 6	m4C	GCTAGCTCTTCN[NNN]	
SapI, VpaK32I	TC							
BspQI, LguI, PciSI,	GCTCT	M.Nme18V		GCGGCGC	2, 5	m5C	GCGGCGCTCTTCN[NNN]	
SapI, VpaK32I	TC							
BspQI, LguI, PciSI,	GCTCT	M.NmeDI		VCCGGY	2, 5	m5C	VCCGGCTCTTCN[NNN]	
SapI, VpaK32I	TC							
BspQI, LguI, PciSI,	GCTCT	M.NmeMC58III		TCTGG	5	m5C	TCTGGCTCTTCN[NNN]	
SapI, VpaK32I	TC							
BspQI, LguI, PciSI,	GCTCT	M.Nsp209III, M.Nsp51109II		CTCGAG	1, 6	m5C	CTCGAGCTCTTCN[NNN]	
SapI, VpaK32I	TC							
BspQI, LguI, PciSI,	GCTCT	M.Nspl, M.NsphI		RCATGY	2, 5	m5C	RCATGCTCTTCN[NNN]	
SapI, VpaK32I	TC							
BspQI, LguI, PciSI,	GCTCT	M.Pae10332II, M1.Pae50071I,		CAGCTC	3, 4	m4C	CAGCTCTTCN[NNN]	

SapI, Vpak32I	TC	M2.Pae50071I					
BspQI, LguI, PciSI,	GCTCT	M.Pam7666I		CCATGG	1, 6	m5C	CCATGGCTCTTCN[NNN]
SapI, Vpak32I	TC						
BspQI, LguI, PciSI,	GCTCT	M.Pfa23572II		CGATCG	1, 6	m5C	CGATCGCTCTTCN[NNN]
SapI, Vpak32I	TC						
BspQI, LguI, PciSI,	GCTCT	M.Pla5I		CTAG	1, 4	Unkno wn	CTAGCTCTTCN[NNN]
SapI, Vpak32I	TC						
BspQI, LguI, PciSI,	GCTCT	M.SacI, M.Sgr13350I		GAGCTC	3, 4	m5C	GAGCTCTTCN[NNN]
SapI, Vpak32I	TC						
BspQI, LguI, PciSI,	GCTCT	M.Sam53668II		CCATGG	1, 6	Unkno wn	CCATGGCTCTTCN[NNN]
SapI, Vpak32I	TC						
BspQI, LguI, PciSI,	GCTCT	M.Sen1781III		CCWWGG	1, 6	Unkno wn	CCWWGGCTCTTCN[NNN]
SapI, Vpak32I	TC						
BspQI, LguI, PciSI,	GCTCT	M.Tam77409IV		CCRCTC	4	m4C	CCGCTCTTCN[NNN]
SapI, Vpak32I	TC						
BspQI, LguI, PciSI,	GCTCT	M.TspRI		CATG	1, 5	m5C	CATGCTCTTCN[NNN]
SapI, Vpak32I	TC						
BspQI, LguI, PciSI,	GCTCT	M.XmnI		GAANNNN TTC	2, 9	m6A	GAAGCTCTTCN[NNN]
SapI, Vpak32I	TC						
BspQI, LguI, PciSI,	GCTCT	M.Yen2516I		GCGCGC	2, 5	Unkno wn	GCGCGCTCTTCN[NNN]
SapI, Vpak32I	TC						
BspQI, LguI, PciSI,	GCTCT	M1.AgoG2I, M2.AgoG2I		ACTGG	2, 5	m4C	ACTGGCTCTTCN[NNN]
SapI, Vpak32I	TC						
BspQI, LguI, PciSI,	GCTCT	M1.BspACI		GCGG	4	m5C	GCGGCTCTTCN[NNN]
SapI, Vpak32I	TC						
BspQI, LguI, PciSI,	GCTCT	M1.HgaI, M2.LlaII		GACGC	4	m5C	GACGCTCTTCN[NNN]
SapI, Vpak32I	TC						

SapI, VpaK32I	TC						
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M1.Lp AG30I	GAAGG	5	m5C	GAAGGCTCTTCN[NNN]	
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M2.BceSIII, M1.Pru9I, M2.Pru9I	ACGGC	2, 4	m4C	ACGGCTCTTCN[NNN]	
BspQI, LguI, PciSI, SapI, VpaK32I	GCTCT TC	M2.HpyAVI, M1.MnII	GAGG	4	m5C	GAGGCTCTTCN[NNN]	

Uso de otros tipos de metiltransferasas.

A pesar de utilizar enzimas metilasa del ADN de tipo II en los ejemplos, también se pueden considerar enzimas de modificación de restricción de tipo I. Las enzimas de modificación de restricción de tipo I son enzimas de subunidades múltiples que tienen actividad tanto de metiltransferasa como de endonucleasa de restricción. Por lo general, constan de tres tipos de subunidades codificadas por tres genes diferentes: subunidad de metilación, subunidad de restricción y subunidad de reconocimiento de secuencia de ADN (subunidad de especificidad). Tanto la subunidad de metilación como la subunidad de especificidad son necesarias para la actividad de la metiltransferasa.

Por ejemplo, se sabe que M.EcoCI5I reconoce **GACNNNNNRtAAY** (base metilada/base opuesta en negrita, RTT complementario inverso **AYNNNNNNNGtC**). Se podría proporcionar un sitio de restricción protegible por metilación según la descripción en el presente documento usando esta secuencia para proteger un sitio BsaI solapado: **RTTATNNNNNGGTCTC** (base metilada/base opuesta en negrita, secuencia de reconocimiento de BsaI subrayada, secuencia de reconocimiento de M.EcoCI5I en cursiva). Un *E. coli* La cepa que expresa tanto la subunidad de metilación M.EcoCI5I como la subunidad de especificidad S.EcoCI5I puede ser capaz de metilar las bases correspondientes en el elemento de control de corte, lo que bloquearía el sitio BsaI protegido por metilación.

La Tabla 4 proporciona algunos ejemplos adicionales de combinaciones de enzimas de restricción y enzimas de restricción-modificación de tipo I, o secuencias de reconocimiento de las mismas, que pueden proporcionarse según la descripción en el presente documento. El experto entenderá que estos son ejemplos y no una lista exhaustiva de combinaciones potenciales que estarían abarcadas por las reivindicaciones.

Tabla 4

Enzima de restricción	Lugar de restricción	Enzima de modificación de restricción tipo I	Secuencia de metilación	Posición de la base metilada	Tipo de metilación	Sitio de metilación/restricción superpuesto
AleXI, BbvI, BseKI, BseXI, Bsp42 3I, Bst12I, Bst71I, BstV1I, CcoCI, Esp110 9I	GCAGC	M.AspU 418I	GCANNNN NTGA	4, 11	m6A	GGCAGCNNNTGANNNN I
AleXI, BbvI, BseKI, BseXI, Bsp42 3I, Bst12I, Bst71I, BstV1I, CcoCI, Esp110 9I	GCAGC	M.Sna3 61I	GCANNNN NTCG	3, 11	m6A	GCAGCNNNTCG(NNNN)
Stal	GGATC	M.Bne D34I	GAGNNNN NTGG	2, 11	m6A	GAGNNNGGATGGNNNN NNNN(NNNN)
BpmI, CsaI	CTGGA G	M.CbaI 3I	GAGNNNN TGT	2, 10	m6A	CTGGAGNNNNRTGTNNNN N[NN]
BamBI, BspQ25 3I, Esp3I	CGTCT C	M.SmaF 1I	CAYNNNN TCA	2, 10	m6A	CAYNNCGTCTCANNNN
PaeC8 II	ACCCA G	M.La21 6I	CCANNNN NNNTGC	4, 13	m6A	GCACCCAGNNNTGGNNNN NNNNNN[NN], CCACCCAGNNNTGCNNNN NNNNNN[NN]

Ejemplo de secuencias vectoriales

Las secuencias de plásmidos para Metclo sin cicatrices se enumeran a continuación en formato fasta. El secuencia subrayada es la "secuencia de descarte" que contiene el fragmento de selección LacZ y un origen de replicación para la preparación de ADN del vector con alto número de copias. La secuencia encuadrada es el "elemento compuesto recesivo/truncante" que contiene un "elemento protegible contra la metilación" y un "elemento no protegible" dispuestos cabeza con cabeza a una distancia adecuada para la conmutación del adaptador. El resto de la secuencia es la columna vertebral del vector que contiene un origen de replicación de pocas copias (basado en pl5A) y un marcador de selección (ya sea kanamicina o espectinomicina).

Para Metclo sin cicatrices basado en BsaI/M.Osp807II, los plásmidos son:

pMXLK_VL, pMXLK_VM, pMXLK_VR (kanamicina)
pMXLS_VL, pMXLS_VM, pMXLS_VR (espectinomicina)

Para Metclo sin cicatrices basado en LglI/M.XmnI, los plásmidos son:

pMZLK_VL, pMZLK_VM, pMZLK_VR (kanamicina)
pMZLS_VL, pMZLS_VM, pMZLS_VR (espectinomicina)

Las secuencias de plásmidos para el ensamblaje estándar de 4 fragmentos de ADN de 1 kb y el ensamblaje de 2 etapas de aproximadamente 200 kb de ADN a partir de 28 fragmentos también se enumeran en formato fasta. La secuencia subrayada para vectores de ensamblaje de 1 kb contiene el marcador de selección LacZ. La secuencia subrayada para vectores de ensamblaje de 200 kb contiene el marcador de selección LacZ y un origen de replicación para la preparación de ADN del vector con alto número de copias. El resto de las secuencias son cadenas principales de vectores.

Para un ensamblaje de 1 kb, los plásmidos son:

pMXP_A4E, pMYP_A5E, pMZP_P6T

Para un ensamblaje de 200 kb, los plásmidos son:

pMXBG_A7I, pMXBG_I7G, pMXBG_G7F, pMXBG_F7E, pMXBK_A7E

Dichas secuencias de vectores son aplicables para los vectores de los ejemplos en el presente documento, intercambiándose los elementos compuestos de truncamiento, mantenimiento o inserción específicos según corresponda. Un experto familiarizado con la clonación molecular puede reconstruir los vectores aquí mediante PCR a partir del kit Moclo disponible comercialmente.

>pMXLK_VL (SEQ ID NO: 28)

```
CACTCTGTGACTTGGTCTCA CTCC TGAGACC
GTCAGACCAAGTTTACTCAT
ATATACTTTAGATTGATTTAAACTTTCATTTTTAATTTAAAAGGATCTAG
GTGAAGATCCTTTTTTGATAATCTCATGACCAAATCCCTTAACGTGAGTT
TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
GAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAAACAAAAAACCA
CCGCTACCAGCGGTGGTTTGTGTTGCCGATCAAGAGCTACCAACTCTTTT
TCCGAAGGTAAGTGGCTTCAAGCAGAGCGCAGATACCAAATACTGTCTTCT
TAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAAGTCTGTAGCACCGCCT
ACATACCTCGCTCTGCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGA
TAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGG
CGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGAGG
CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA
GGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGG
TATCTTTATAGTCCTGTGCGGTTTTGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACG
```

CGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTT
 TTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGA
 GTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAG
 TGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACCGCCTCTCCCCGCG
 CGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
 AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
 CACCCAGGCTTTACACTTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT
 GTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGC
 CAAGCTTGCAATGCCTGCAGGTGCACTCTAGAGGATCCCCGGGTACCGAGC
 TCGAATTCACCTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGG
 CGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTCGCCAGCTGGC
 GTAATAGCGAAGAGGGCCGACCGATCGCCCTTCCCAACAGTTGCGCAGC
 CTGAATGGCGAATGGCGCCTGATGCGGTATTTCTCCTTACGCATCTGTG
 CGGTATTTTACACCGCATATGGTGCACCTCTCAGTACAATCTGCTCTGATG
 CCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC
 TGACGGGCTTGCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC
 GGTCTCA CGAG ACCAAGTCACGAAGTGTGCGGCCGCCTTCGTGCGCGGAAC
 CCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGA
 GACAATAACCCGTATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATG
 GCTAAAATGAGAATATCACCGGAATTGAAAAACTGATCGAAAAATACCG
 CTGCGTAAAAGATACGGAAGGAATGTCTCCTGCTAAGGTATATAAGCTGG
 TGGGAGAAAATGAAAACCTATATTTAAAAATGACGGACAGCCGGTATAAA
 GGGACCACCTATGATGTGGAACGGGAAAAGGACATGATGCTATGGCTGGA
 AGGAAAGCTGCCTGTTCCAAAGGTCCCTGCACTTTGAACGGCATGATGGCT
 GGAGCAATCTGCTCATGAGTGAGGCCGATGGCGTCCTTTGCTCGGAAGAG
 TATGAAGATGAACAAAGCCCTGAAAAGATTATCGAGCTGTATGCGGAGTG
 CATCAGGCTCTTTCACTCCATCGACATATCGGATTGTCCCTATACGAATA
 GCTTAGACAGCCGCTTAGCCGAATTGGATTACTTACTGAATAACGATCTG
 GCCGATGTGGATTGCGAAAACCTGGGAAGAGGACACTCCATTTAAAGATCC
 GCGCGAGCTGTATGATTTTTTTAAAGACGGAAAAGCCCGAAGAGGAACTTG
 TCTTTTCCCACGGCGACCTGGGAGACAGCAACATCTTTGTGAAAGATGGC
 AAAGTAAGTGGCTTTATTGATCTTGGGAGAAGCGGCAGGGCGGACAAGTG
 GTATGACATTGCCTTCTGCGTCCGGTCGCTCAGGGAGGATATCGGGGAAG
 AACAGTATGTGAGCTATTTTTTTGACTTACTGGGGATCAAGCCTGATTGG
 GAGAAAATAAAATATTATATTTTACTGGATGAATTGTTTTAGGAGCCTGT
 TTAATAAGATGATCTTCTTGAGATCGTTTTTGGTCTGCGCGTAATCTCTTG
 CTCTGAAAACGAAAAAACCGCCTTGCAGGGCGGTTTTTCGAAGGTTCTCT
 GAGCTACCAACTCTTTGAACCGAGGTAACCTGGCTTGGAGGAGCGCAGTCA
 CAAAACCTTGTCCTTTTCAGTTTAGCCTTAACCGGCGCATGACTTCAAGAC
 TAACTCCTCTAAATCAATTACCAGTGGCTGCTGCCAGTGGTGCTTTTGCA
 TGTCTTTCCGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCG
 GTCGGACTGAACGGGGGGTTCGTGCATACAGTCCAGCTTGGAGCGAACTG
 CCTACCCGGAACAGTGTGTCAGGCGTGGAATGAGACAAACGCGGCCATAA
 CAGCGGAATGACACCGGTAAACCGAAAGGCAGGAACAGGAGAGCGCAGCA
 GGGAGCCGCCAGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGTTT
 CGCCACCACTGATTTGAGCGTCAGATTTCTGATGCTTGTGAGGGGGGCG
 GAGCCTATGGAAAAACGGCTTTGCCGCGGCCCTCTCACTTCCCTGTAAAG
 TATCTTCTGCGATCTTCCAGGAAATCTCCGCCCCGTTCGTAAGCCATTT
 CCGCTCGCCGAGTCGAACGACCGAGCGTAGCGAGTCAGTGAGCGAGGAA
 GCGGAATATATCCTGTATCACATATTCTGCTGACGCACCGGTGCAGCCTT
 TTTTCTCCTGCCACATGAAGCACTTCACTGACACCCTCATCAGTGCCAAC
 ATAGTAAGCCAGTATACACTCCGCTAGCGCGCAAGCGGCCGCA

>pMXLK_VM (SEQ ID NO: 29)

CACTCTGTGACTTGGT CTCC TGAGACC
GTCAGACCAAGTTTACTCAT
ATATACTTTAGATTGATTTAAACTTCATTTTAAATTTAAAAGGATCTAG
GTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACTGAGTT
TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
GAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCA
CCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTT
TCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTC
TAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACC GCCT
ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGA
TAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGG
CGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAG
CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA
GGGTGCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGG
TATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACG
CGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTT
TTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGA
GTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAG
TGAGCGAGGAAGCGGATGAGCGCCCAATACGCAACCGCCTCTCCCCGCG
CGTTGGCCGATTTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
CACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT
GTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGC
CAAGCTTGCTATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGC
TCGAATTCAGTGGCCGTCGTTTTTACAACGTCGTGACTGGGAAAACCTTGG
CGTTACCCAACTTAATCGCCTTGCGAGCACATCCCCCTTTTCGCCAGCTGGC
GTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC
CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG
CGGTATTTACACCGCATATGGTGCCTCTCAGTACAATCTGCTCTGATG
CCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC
TGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC
GGTCTCA CGAG ACCAAGTCACGAAGTGTGCGGCCGCCTTCGTGCGCGGAAC
CCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGA
GACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATG
GCTAAAATGAGAATATCACCGGAATTGAAAAAACTGATCGAAAAATACCG
CTGCGTAAAAGATACGGAAGGAATGTCTCCTGCTAAGGTATATAAGCTGG
TGGGAGAAAATGAAAACCTATATTTAAAAATGACGGACAGCCGGTATAAA
GGGACCACCTATGATGTGGAACGGGAAAAGGACATGATGCTATGGCTGGA
AGGAAAGCTGCCTGTTCCAAAGGTCCTGCACTTTGAACGGCATGATGGCT
GGAGCAATCTGCTCATGAGTGAGGCCGATGGCGTCCTTTGCTCGGAAGAG
TATGAAGATGAACAAAGCCCTGAAAAGATTATCGAGCTGTATGCGGAGTG
CATCAGGCTCTTTCACTCCATCGACATATCGGATTGTCCCTATACGAATA
GCTTAGACAGCCGCTTAGCCGAATTGGATTACTTACTGAATAACGATCTG
GCCGATGTGGATTGCGAAAACCTGGGAAGAGGACACTCCATTTAAAGATCC
GCGCGAGCTGTATGATTTTTTAAAGACGGAAAAGCCCGAAGAGGAACTTG
TCTTTTCCCACGGCGACCTGGGAGACAGCAACATCTTTGTGAAAGATGGC
AAAGTAAGTGGCTTTATGATCTTGGGAGAAGCGGCAGGGCGGACAAGTG
GTATGACATTGCCTTCTGCGTCCGGTCGCTCAGGGAGGATATCGGGGAAG
AACAGTATGTCGAGCTATTTTTTACTTACTGGGGATCAAGCCTGATTGG

GAGAAAATAAAATATTATATTTTACTGGATGAATTGTTTTAGGAGCCTGT
 TTAATAAGATGATCTTCTTGAGATCGTTTTGGTCTGCGCGTAATCTCTTG
 CTCTGAAAACGAAAAAACCGCCTTGACAGGGCGGTTTTTCGAAGGTTCTCT
 GAGCTACCAACTCTTTGAACCGAGGTAAGTGGCTTGAGGAGCGCAGTCA
 CCAAACTTGTCTTTTTCAGTTTAGCCTTAACCGGCGCATGACTTCAAGAC
 TAACTCCTCTAAATCAATTACCAGTGGCTGCTGCCAGTGGTGTCTTTTGCA
 TGTCTTTCCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCG
 GTCGGACTGAACGGGGGGTTCGTGCATACAGTCCAGCTTGAGAGCAACTG
 CCTACCCGGAAGTGAAGTGTGAGGCGTGGAATGAGACAAACGCGGCCATAA
 CAGCGGAATGACACCGGTAAACCGAAAGGCAGGAACAGGAGAGCGCACGA
 GGGAGCCGCCAGGGGGAAACGCCTGGTATCTTTATAGTCTGTGCGGGTTT
 CGCCACCACTGATTTGAGCGTCAGATTTCTGTGATGCTTGTGAGGGGGGCG
 GAGCCTATGGAAAAACGGCTTTGCCGCGGCCCTCTCACTTCCCTGTTAAG
 TATCTTCTGTCATCTTCCAGGAAATCTCCGCCCCGTTCTGTAAGCCATTT
 CCGCTCGCCGCGAGTGAACGACCGAGCGTAGCGAGTCAGTGAGCGAGGAA
 GCGGAATATATCCTGTATCACATATTCTGCTGACGACCGGTGACGCCTT
 TTTTCTCCTGCCACATGAAGCACTTCACTGACACCCTCATCAGTGCCAAC
 ATAGTAAGCCAGTATACACTCCGCTAGCGCGCAAGCGGCCGCA

>pMXLK_VR (SEQ ID NO: 30)

CACTCTGTGACTTGGT CTCC TGAGACC
 GTCAGACCAAGTTTACTCAT
 ATATACTTTAGATTGATTTAAACTTTCATTTTTAATTTAAAAGGATCTAG
 GTGAAGATCCTTTTTGATAATCTCATGACCAAATCCCTTAACGTGAGTT
 TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
 GAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCA
 CCGCTACCAGCGGTGGTTTGTGTGCCGGATCAAGAGCTACCAACTCTTTT
 TCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCTTCT
 TAGTGTTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCT
 ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGA
 TAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGG
 CGCAGCGGTTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGAG
 CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
 CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA
 GGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGG
 TATCTTTATAGTCTGTGCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
 TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACG
 CGGCCTTTTTTACGGTTCTTGGCCTTTTGCTGGCCTTTTGCTCACATGTTC
 TTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGA
 GTGAGCTGATACCGCTCGCCGCGAGCCGAACGACCGAGCGCAGCGAGTCAG
 TGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACCGCCTCTCCCCGCG
 CGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
 AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
 CACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT
 GTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGC
 CAAGCTTGCATGCCTGCAGGTGCACTCTAGAGGATCCCCGGGTACCGAGC
 TCGAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCCTGG
 CGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTGCCAGCTGGC
 GTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC
 CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG
 CGGTATTTACACCGCATATGGTGCACCTCTCAGTACAATCTGCTCTGATG
 CCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC

TGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC
GGTCTCA CGAG TGAGACCAAGTCACGAAGTGTGCGGCCGCCTTCGTGCGCGGAAC
 CCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGA
 GACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATG
 GCTAAAATGAGAATATCACCGGAATTGAAAAAAGCTGATCGAAAAATACCG
 CTGCGTAAAAGATACGGAAGGAATGTCTCCTGCTAAGGTATATAAGCTGG
 TGGGAGAAAATGAAAACCTATATTTAAAAATGACGGACAGCCGGTATAAA
 GGGACCACCTATGATGTGGAACGGGAAAAGGACATGATGCTATGGCTGGA
 AGGAAAAGCTGCCTGTTCCAAAGGTCTGCACTTTGAACGGCATGATGGCT
 GGAGCAATCTGCTCATGAGTGAGGCCGATGGCGTCCTTTGCTCGGAAGAG
 TATGAAGATGAACAAAGCCCTGAAAAGATTATCGAGCTGTATGCGGAGTG
 CATCAGGCTCTTTCACTCCATCGACATATCGGATTGTCCCTATACGAATA
 GCTTAGACAGCCGCTTAGCCGAATTGGATTACTTACTGAATAACGATCTG
 GCCGATGTGGATTGCGAAAACTGGGAAGAGGACACTCCATTTAAAGATCC
 GCGCGAGCTGTATGATTTTTTAAAGACGGAAGCCCGAAGAGGAAGCTTG
 TCTTTTCCACGGCGACCTGGGAGACAGCAACATCTTTGTGAAAGATGGC
 AAAGTAAGTGGCTTTTATTGATCTTGGGAGAAGCGGCAGGGCGGACAAGTG
 GTATGACATTGCCTTCTGCGTCCGGTCGCTCAGGGAGGATATCGGGGAAG
 AACAGTATGTGAGCTATTTTTTGACTTACTGGGGATCAAGCCTGATTGG
 GAGAAAATAAAATATTATATTTTACTGGATGAATTGTTTTAGGAGCCTGT
 TTAATAAGATGATCTTCTTGAGATCGTTTTTGGTCTGCGCGTAATCTCTTG
 CTCTGAAAACGAAAAAACCGCCTTGCAAGGGCGGTTTTTCGAAGGTTCTCT
 GAGCTACCAACTCTTTGAACCGAGGTAAGTGGCTTGGAGGAGCGCAGTCA
 CCAAACTTGTCTTTTCACTTTAGCCTTAACCGGCGCATGACTTCAAGAC
 TAACTCCTCTAAATCAATTACCAGTGGCTGCTGCCAGTGGTGTCTTTTGCA
 TGTCTTTCCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCG
 GTCGGACTGAACGGGGGGTTTCGTGCATACAGTCCAGCTTGGAGCGAACTG
 CCTACCCGGAAGTGAAGTGTGAGGCGTGGAATGAGACAAACGCGGCCATAA
 CAGCGGAATGACACCGGTAAACCGAAAGGCAGGAACAGGAGAGCGCACGA
 GGGAGCCGCCAGGGGAAACGCCTGGTATCTTTATAGTCCTGTCTGGGTTT
 CGCCACCACTGATTTGAGCGTCAGATTTCTGTGATGCTTGTGAGGGGGCG
 GAGCCTATGGAAAAACGGCTTTGCCGCGGCCCTCTCACTTCCCTGTAAAG
 TATCTTCTGTCATCTTCCAGGAAATCTCCGCCCCGTTCTGTAAGCCATTT
 CCGCTCGCCGAGTCGAACGACCGAGCGTAGCGAGTCAGTGAGCGAGGAA
 GCGGAATATATCCTGTATCACATATTCTGCTGACGCACCGGTGCAGCCTT
 TTTTCTCCTGCCACATGAAGCACTTCACTGACACCTCATCAGTGCCAAC
 ATAGTAAGCCAGTATACACTCCGCTAGCGCGCAAGCGGCCGCA

>pMZLK_VL (SEQ ID NO: 31)

CACTCTGT GAAGCTCTTCA TCC TGAAGAGC
GGTCAGACCAAGTTTACTCAT
ATATACTTTTAGATTGATTTAAACTTCATTTTAAATTTAAAAGGATCTAG
GTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTT
TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
GAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCA
CCGCTACCAGCGGTGGTTTGTTGCCGGATCAAGAGCTACCAACTCTTTT
TCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTC
TAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACC GCCT
ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGA
TAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGG
CGCAGCGGTCTGGGCTGAACGGGGGGTTGCTGCACACAGCCCAGCTTGGAG
CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA

GGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGG
TATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACG
CGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTT
TTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGA
GTGAGCTGATACCGCTCGCCGCGAGCCGAACGACCGAGCGCAGCGAGTCAG
TGAGCGAGGAAGCGGATGAGCGCCCAATACGAAACCGCCTCTCCCCGCG
CGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
CACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT
GTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGC
CAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGC
TCGAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGG
CGTTACCCAACCTTAATCGCCTTGACGACATCCCCCTTTCCGCCAGCTGGC
GTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC
CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG
CGGTATTTACACCCGCATATGGTGCACCTCTCAGTACAATCTGCTCTGATG
CCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC
TGACGGGCTTGCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACC
GCTCTTCA CGA AGAGCTTC ACGAAGTGTGCGGCCCGCCTTCGTGCGCGGAAC
CCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGA
GACAATAACCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATG
GCTAAAATGAGAATATCACCGGAATTGAAAAAACTGATCGAAAAATACCG
CTGCGTAAAAGATACGGAAGGAATGTCTCCTGCTAAGGTATATAAGCTGG
TGGGAGAAAAATGAAAACCTATATTTAAAAATGACGGACAGCCGGTATAAA
GGGACCACCTATGATGTGGAACGGGAAAAGGACATGATGCTATGGCTGGA
AGGAAAGCTGCCTGTTCCAAAGGTCCTGCACTTTGAACGGCATGATGGCT
GGAGCAATCTGCTCATGAGTGAGGCCGATGGCGTCCTTTGCTCGGAAGAG
TATGAAGATGAACAAAGCCCTGAAAAGATTATCGAGCTGTATGCGGAGTG
CATCAGGCTCTTTCACTCCATCGACATATCGGATTGTCCCTATACGAATA
GCTTAGACAGCCGCTTAGCCGAATTGGATTACTTACTGAATAACGATCTG
GCCGATGTGGATTGCGAAAACCTGGGAAGAGGACACTCCATTTAAAGATCC
GCGCGAGCTGTATGATTTTTTAAAGACGGAAAAGCCCGAAGAGGAACTTG
TCTTTTCCCACGGCGACCTGGGAGACAGCAACATCTTTGTGAAAGATGGC
AAAGTAAGTGGCTTTATTTGATCTTGGGAGAAGCGGCAGGGCGGACAAGTG
GTATGACATTGCCTTCTGCGTCCGGTCGCTCAGGGAGGATATCGGGGAAG
AACAGTATGTGAGCTATTTTTTTGACTTACTGGGGATCAAGCCTGATTGG
GAGAAAATAAAATATTATATTTTACTGGATGAATTGTTTTAGGAGCCTGT
TTAATAAGATGATCTTCTTGAGATCGTTTTTGGTCTGCGCGTAATCTCTTG
CTCTGAAAACGAAAAAACCGCCTTGACAGGGCGGTTTTTCGAAGGTTCTCT
GAGCTACCAACTCTTTGAACCGAGGTAACCTGGCTTGGAGGAGCGCAGTCA
CCAAAACCTTGTCCTTTTCACTTTAGCCTTAACCGGCGCATGACTTCAAGAC
TAACTCCTCTAAATCAATTACAGTGCGTGTGCCAGTGGTGCTTTTTGCA
TGCTTTTCCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCG
GTCGGACTGAACGGGGGTTTCGTGCATACAGTCCAGCTTGGAGCGAACTG
CCTACCCGGAACCTGAGTGTGAGCGTGGAATGAGACAAACGCGGCCATAA
CAGCGGAATGACACCGGTAAACCGAAAGGCAGGAACAGGAGAGCGCACGA
GGGAGCCGCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTT
CGCCACCACTGATTTGAGCGTCAGATTTCTGTGATGCTTGTCAGGGGGGCG
GAGCCTATGGAAAAACGGCTTTGCCGCGGCCCTCTCACTTCCCTGTAAAG
TATCTTCTGTCATCTTCCAGGAAATCTCCGCCCCGTTTCGTAAGCCATTT
CCGCTCGCCGAGTCGAACGACCGAGCGTAGCGAGTCAGTGAGCGAGGAA
GCGGAATATATCCTGTATCACATATTCTGCTGACGCACCGGTGCAGCCTT

ES 2 979 391 T3

TTTTCTCCTGCCACATGAAGCACTTCACTGACACCCTCATCAGTGCCAAAC
ATAGTAAGCCAGTATACACTCCGCTAGCGCGCAAGCGGCCGCA

>pMZLK_VM (SEQ ID NO: 32)

CACTCTGT GAAGCTCT TCC TGAAGAGC
GGTCAGACCAAGTTTACTCAT
ATATACTTTAGATTGATTTAAACTTCATTTTTTAATTTAAAAGGATCTAG
GTGAAGATCCTTTTTGATAATCTCATGACCAAATCCCTTAACGTGAGTT
TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
GAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCA
CCGCTACCAGCGGTGGTTTGTGTGCCGGATCAAGAGCTACCAACTCTTTT
TCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCTTC
TAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCT
ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGA
TAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGG
CGCAGCGGTCCGGCTGAACGGGGGGTTTCGTGCACACAGCCCAGCTTGGAG
CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA
GGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGG
TATCTTTATAGTCCTGTCCGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACG
CGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTT
TTTCTGCGTTATCCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGA
GTGAGCTGATACCGCTCGCCCGCAGCCGAACGACCGAGCGCAGCGAGTCAG
TGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACCGCCTCTCCCCGCG
CGTTGGCCGATTTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
CACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT
GTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGC
CAAGCTTGCAATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGC
TCGAATTCACCTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCCTGG
CGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTTCGCCAGCTGGC
GTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC
CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG
CGGTATTTACACCCGCATATGGTGCACCTCTCAGTACAATCTGCTCTGATG
CCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC
TGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACC
GCTCTTCA CGA AGAGCTTC ACGAAGTGTGCGGCCGCTTTCGTGCGCGGAAC
CCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGA
GACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATG
GCTAAAATGAGAATATCACCGGAATTGAAAAAACTGATCGAAAAATACCG
CTGCGTAAAAGATACGGAAGGAATGTCTCCTGCTAAGGTATATAAGCTGG
TGGGAGAAAAATGAAAACCTATATTTAAAAATGACGGACAGCCGGTATAAA
GGGACCACCTATGATGTGGAACGGGAAAAGGACATGATGCTATGGCTGGA
AGGAAAGCTGCCTGTTCCAAAGGTCCTGCACTTTGAACGGCATGATGGCT
GGAGCAATCTGCTCATGAGTGAGGCCGATGGCGTCCTTTGCTCGGAAGAG
TATGAAGATGAACAAAGCCCTGAAAAGATTATCGAGCTGTATGCGGAGTG
CATCAGGCTCTTTCACTCCATCGACATATCGGATTGTCCCTATACGAATA
GCTTAGACAGCCGCTTAGCCGAATTGGATTACTTACTGAATAACGATCTG
GCCGATGTGGATTGCGAAAACCTGGGAAGAGGACACTCCATTTAAAGATCC
GCGCGAGCTGTATGATTTTTTAAAGACGGAAAAGCCCGAAGAGGAACTTG
TCTTTTCCACGGCGACCTGGGAGACAGCAACATCTTTGTGAAAGATGGC
AAAGTAAGTGGCTTTATTGATCTTGGGAGAAGCGGCAGGGCGGACAAGTG

ES 2 979 391 T3

GTATGACATTGCCTTCTGCGTCCGGTCCGCTCAGGGAGGATATCGGGGAAG
AACAGTATGTCGAGCTATTTTTTGAAGTACTGGGGATCAAGCCTGATTGG
GAGAAAATAAAATATTATATTTTACTGGATGAATTGTTTTAGGAGCCTGT
TTAATAAGATGATCTTCTTGAGATCGTTTTTGGTCTGCGCGTAATCTCTTG
CTCTGAAAACGAAAAAACCGCCTTGACAGGGCGGTTTTTCGAAGGTTCTCT
GAGCTACCAACTCTTTGAACCGAGGTAAGTGGCTTGGAGGAGCGCAGTCA
CCAAAACCTTGTCCTTTTCAAGTTTAGCCTTAACCGGCGCATGACTTCAAGAC
TAACTCCTCTAAATCAATTACCAGTGGCTGCTGCCAGTGGTGCTTTTTGCA
TGCTTTTCCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCG
GTCGGACTGAACGGGGGGTTCGTGCATACAGTCCAGCTTGGAGCGAACTG
CCTACCCGGAAGTGAAGTGTGAGGCGTGGAATGAGACAAACGCGGCCATAA
CAGCGGAATGACACCGGTAAACCGAAAGGCAGGAACAGGAGAGCGCACGA
GGGAGCCGCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCCGGGTTT
CGCCACCAGTGAATTTGAGCGTCAGATTTTCGTGATGCTTGTGAGGGGGCG
GAGCCTATGGAAAAACGGCTTTGCCGCGGCCCTCTCACTTCCCTGTTAAG
TATCTTCCCTGGCATCTTCCAGGAAATCTCCGCCCCGTTGTAAGCCATTT
CCGCTCGCCGCAGTCGAACGACCGAGCGTAGCGAGTCAGTGAGCGAGGAA
GCGGAATATATCCTGTATCACATATTCTGCTGACGCACCGGTGCAGCCTT
TTTTCTCCTGCCACATGAAGCACTTCACTGACACCCATCATCAGTGCCAAC
ATAGTAAGCCAGTATACACTCCGCTAGCGCGCAAGCGGCCGCA

>pMZLK_VR (SEQ ID NO: 33)

CACTCTGT GAAGCTCT TCC TGAAGAGC
GGTCAGACCAAGTTTACTCAT
ATATACTTTAGATTGATTTAAAAC TTCATTTTAAATTTAAAAGGATCTAG
GTGAAGATCCTTTTTTGATAATCTCATGACC AAAATCCCTTAACGTGAGTT
TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
GAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACC
CCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTT
TCCGAAGGTAAC TGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTC
TAGTG TAGCCGTAGTTAGGCCACCACTTCAAGA ACTCTGTAGCACCGCCT
ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGA
TAAGTCGTGTCTTACCGGGTTGGA CTCAAGACGATAGTTACCGGATAAGG
CGCAGCGGTCTGGGCTGAACGGGGGGTTTCGTGCACACAGCCCAGCTTGGAG
CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCG GTAAGCGGCA
GGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGG
TATCTTTATAGTCCTGTCTGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACG
CGGCCTTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTT
TTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTTGA
GTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAG
TGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACGCCTCTCCCCGCG
CGTTGGCCGATT CATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
CACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT
GTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGC
CAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGC
TCGAATTC ACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGG
CGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTGCCAGCTGGC
GTAATAGCGAAGAGGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC
CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG
CGGTATTTACACCGCATATGGTGC ACTCTCAGTACAATCTGCTCTGATG

CCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC
TGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACC
GCTCTTCA CGA TGAAGAGCTTC ACGAAGTGTGCGGCCGCCTTCGTGCGCGGAAC
CCCTATTTGTTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGA
GACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATG
GCTAAAATGAGAATATCACCGGAATTGAAAAACTGATCGAAAAATACCG
CTGCGTAAAAGATACGGAAGGAATGTCTCCTGCTAAGGTATATAAGCTGG
TGGGAGAAAATGAAAACCTATATTTAAAAATGACGGACAGCCGGTATAAA
GGGACCACCTATGATGTGGAACGGGAAAAGGACATGATGCTATGGCTGGA
AGGAAAGCTGCCTGTTCCAAAGGTCTGCACTTTGAACGGCATGATGGCT
GGAGCAATCTGCTCATGAGTGAGGCCGATGGCGTCCTTTGCTCGGAAGAG
TATGAAGATGAACAAAGCCCTGAAAAGATTATCGAGCTGTATGCGGAGTG
CATCAGGCTCTTTCACTCCATCGACATATCGGATTGTCCCTATACGAATA
GCTTAGACAGCCGCTTAGCCGAATTGGATTACTTACTGAATAACGATCTG
GCCGATGTGGATTGCGAAAACCTGGGAAGAGGACACTCCATTTAAAGATCC
GCGCGAGCTGTATGATTTTTTAAAGACGGAAAAGCCCGAAGAGGAACTTG
TCTTTTCCCACGGCGACCTGGGAGACAGCAACATCTTTGTGAAAGATGGC
AAAGTAAGTGGCTTTATGATCTTGGGAGAAGCGGCAGGGCGGACAAGTG
GTATGACATTGCCTTCTGCGTCCGGTCGCTCAGGGAGGATATCGGGGAAG
AACAGTATGTCGAGCTATTTTTTGACTTACTGGGGATCAAGCCTGATTGG
GAGAAAATAAAATATTATATTTTACTGGATGAATTGTTTTAGGAGCCTGT
TTAATAAGATGATCTTCTTGAGATCGTTTTTGGTCTGCGCGTAATCTCTTG
CTGTGAAAACGAAAAACCGCCTTGACAGGGCGGTTTTTCGAAGGTTCTCT
GAGCTACCAACTCTTTGAACCGAGGTAACCTGGCTTGAGAGAGCGCAGTCA
CCAAAACCTTGTCCTTTCAGTTTTAGCCTTAACCGGCGCATGACTTCAAGAC
TAACTCCTCTAAATCAATTACCAGTGGCTGCTGCCAGTGGTGCTTTTGCA
TGTCTTTCCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCG
GTCGGACTGAACGGGGGGTTTCGTGCATACAGTCCAGCTTGGAGCGAACTG
CCTACCCGGAACCTGAGTGTGAGCGTGGAATGAGACAAACGCGGCCATAA
CAGCGGAATGACACCGGTAAACCGAAAGGCAGGAACAGGAGAGCGCACGA
GGGAGCCGCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGTTT
CGCCACCACTGATTTGAGCGTCAGATTTCTGTGATGCTTGTGAGGGGGGCG
GAGCCTATGGAAAAACGGCTTTGCCGCGGCCCTCTCACTTCCCTGTTAAG
TATCTTCTGGCATCTTCCAGGAAATCTCCGCCCCGTTCTGTAAGCCATTT
CCGCTCGCCGCAGTCGAACGACCGAGCGTAGCGAGTCAGTGAGCGAGGAA
GCGGAATATATCTGTATCACATATTCTGCTGACGCACCGGTGCAGCCTT
TTTTCTCTGCCACATGAAGCACTTCACTGACACCCCTCATCAGTGCCAAC
ATAGTAAGCCAGTATACACTCCGCTAGCGCGCAAGCGGCCGCA

>pMXLS_VL (SEQ ID NO: 34)

ES 2 979 391 T3

CACTCTGTGACTTGGTCTCA CTCC TGAGACC

GTCAGACCAAGTTTACTCAT

ATATACTTTAGATTGATTTAAAACCTTCATTTTTAATTTAAAAGGATCTAG
GTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTT
TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
GAGATCCTTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCA
CCGCTACCAGCGGTGGTTTGTTTGCCGATCAAGAGCTACCAACTCTTTT
TCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTC
TAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCT
ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGA
TAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGG
CGCAGCGGTCGGGCTGAACGGGGGGTTTCGTGCACACAGCCCAGCTTGGAG

CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA
GGGTGCGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCCTGG
TATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACG
CGGCCTTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTTC
TTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGA
GTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAG
TGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACCGCCTCTCCCCGCG
CGTTGGCCGATTTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
CACCCAGGCTTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT
GTGAGCGGATAACAATTTTACACAGGAAACAGCTATGACCATGATTACGC
CAAGCTTGCATGCCTGCAGGTGCACTCTAGAGGATCCCCGGGTACCGAGC
TCGAATTCACCTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGG
CGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTTCGCCAGCTGGC
GTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC
CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG
CGGTATTTTACACCGCATATGGTGCCTCTCAGTACAATCTGCTCTGATG
CCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC
TGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC

GGTCTCA CGAG ACCAAGTCACGAAGTG TGCGGGCCGCCTTCG
TGCGCGGAACCCCTATTTGTTTTATTTTTCTAAATACATTCAAATATGTAT
CCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAG
GAAGAGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCG
GTGGTAACGGCGCAGTGGCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTTT
GGGTACAGTCTATGCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGT
GGGTGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCA
GTCGCCCTAAAACAAAGTTAAACATCATGAGGGAAGCGGTGATCGCCGAA
GTATCGACTCAACTATCAGAGGTAGTTGGCGTCATCGAGCGCCATCTCGA
ACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGGCC
TGAAGCCACACAGCGATATTGATTTGCTGGTTACGGTGACCGTAAGGCTT
GATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAACTTCGGC
TTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTCACCATTGTTG
TGCACGACGACATCATTCCGTGGCGTTATCCAGCTAAGCGCGAACTGCAA
TTTGGAGAATGGCAGCGCAATGACATTCTTGAGGTATCTTCGAGCCAGC
CACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACATA
GCGTTGCCTTGGTAGGTCCAGCGCGGAGGAACTCTTTGATCCGTTTCCT
GAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAACGCTATGAACTC
GCCGCCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTACGTTGTCCC
GCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTCGCT
GCCGACTGGGCAATGGAGCGCCTGCCGGCCAGTATCAGCCCGTCATACT
TGAAGCTAGACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGC
GCGCAGATCAGTTGGAAGAATTTGTCCATTACGTAAAAGGCGAGATCACC
AAGGTAGTCGGCAAATAAGAGCCTGTTTAATAAGATGATCTTCTTGAGAT
CGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAAACGCCTT
GCAGGGCGGTTTTTTCGAAGTTCTCTGAGCTACCAACTCTTTGAACCGAG
GTAAGTGGCTTGGAGGAGCGCAGTCACCAAACTTGTCTTTTTCAGTTTAG
CCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAG
TGGCTGCTGCCAGTGGTGCTTTTGCATGTCTTTCCGGGTGGACTCAAGA
CGATAGTTACCGGATAAGGCGCAGCGGTCGGAAGTGAACGGGGGGTTCGTG
CATAAGTCCAGCTTGGAGCGAACTGCCTACCCGGAAGTGAAGTGTGAGC

GTGGAATGAGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCG
AAAGGCAGGAACAGGAGAGCGCACGAGGGAGCCGCCAGGGGGAAACGCCT
GGTATCTTTATAGTCCTGTCGGGTTTCGCCACCACTGATTGAGCGTCAG ATTCGTGATGCTTGTAG-
GGGGGCGGAGCCTATGGAAAAACGGCTTTGC CGCGGCCCTCTCACTTCCCTGTTAAGTATCTTCCT-
GGCATCTTCCAGGAA ATCTCCGCCCCGTTGTAAGCCATTTCCGCTCGCCCGAGTCGAACGACCG
AGCGTAGCGAGTCAGTGAGCGAGGAAGCGGAATATATCCTGTATCACATA
TTCTGCTGACGCACCGGTGCAGCCTTTTTTCTCCTGCCACATGAAGCACT TCACTGACACCCTCATCAGT-
GCCAACATAGTAAGCCAGTATACACTCCGC TAGCGCGCAAGCGGCCGCA

>pMXLS VM (SEQ ID NO: 35)

CACTCTGTGACTTGGT CTCC TGAGACC
GTCAGACCAAGTTTACTCAT
ATATACTTTAGATTGATTTAAACTTTCATTTTTTAATTTAAAAGGATCTAG
GTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTT
TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
GAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCA
CCGCTACCAGCGGTGGTTTGTGTGCGGATCAAGAGCTACCAACTCTTTT
TCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTC
TAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCT
ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGA
TAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGG
CGCAGCGGTCTGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAG
CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA
GGGTCCGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGG
TATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACG
CGGCCTTTTTTACGGTTCTTGGCCTTTTGCTGGCCTTTTGCTCACATGTTT
TTTCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGA
GTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAG
TGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACCGCCTCTCCCCGCG
CGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
CACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT
GTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGC
CAAGCTTGCATGCCTGCAGGTGCACTCTAGAGGATCCCCGGGTACCGAGC
TCGAATTCCTGCGCGTCTGTTTTACAACGTCGTGACTGGGAAAACCTGG
CGTTACCCAACTTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGC
GTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC
CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG
CGGTATTTACACCGCATATGGTGCCTCTCAGTACAATCTGCTCTGATG
CCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC
TGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC
GGTCTCA CGAG ACCAAGTCACGAAGTG TGCGGCCGCCTTCG
TGCGCGGAACCCCTATTTGTTTTATTTTCTAAATACATTCAAATATGTAT
CCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAG
GAAGAGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCG
GTGGTAACGGCGCAGTGGCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTTT
GGGGTACAGTCTATGCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGT
GGGTGATGTTTGTATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCA
GTCGCCCTAAAACAAAGTTAAACATCATGAGGGAAGCGGTGATCGCCGAA

GTATCGACTCAACTATCAGAGGTAGTTGGCGTCATCGAGCGCCATCTCGA
ACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGGCC
TGAAGCCACACAGCGATATTGATTTGCTGGTTACGGTGACCGTAAGGCTT
GATGAAACAACGCGCGGAGCTTTGATCAACGACCTTTTGAAACTTCGGC
TTCCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAAGTCACCATTTGTTG
TGCACGACGACATCATTCCGTGGCGTTATCCAGCTAAGCGCGAACTGCAA
TTTGGAGAATGGCAGCGCAATGACATTCTTGCAAGGTATCTTCGAGCCAGC
CACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACATA
GCGTTGCCTTGGTAGGTCCAGCGGCGGAGGAAGTCTTTGATCCGGTTCCCT
GAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAACGCTATGGAACCTC
GCCGCCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTACGTTGTCCC
GCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTGCGT
GCCGACTGGGCAATGGAGCGCCTGCCGGCCAGTATCAGCCCGTCATACT
TGAAGCTAGACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGC
GCGCAGATCAGTTGGAAGAATTTGTCCATTACGTAAAAGGCGAGATCACC
AAGGTAGTCGGCAAATAAGAGCCTGTTTAATAAGATGATCTTCTTGAGAT
CGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAAACCGCCTT
GCAGGGCGGTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAACCGAG
GTAACCTGGCTTGGAGGAGCGCAGTCACCAAAAGTGTCTTTTTCAGTTTAG
CCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAG
TGGCTGCTGCCAGTGGTGCTTTTGCATGTCTTTCGGGTTGGACTCAAGA
CGATAGTTACCGGATAAAGGCGCAGCGGTCGGACTGAACGGGGGGTTTCGTG
CATACAGTCCAGCTTGGAGCGAACTGCCTACCCGGAAGTGAAGTGTAGGC
GTGGAATGAGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCG
AAAGGCAGGAACAGGAGAGCGCACGAGGGAGCCGCCAGGGGGAAACGCCT
GGTATCTTTATAGTCCTGTGCGGTTTCGCCACCACTGATTTGAGCGTCAG
ATTTTCGTGATGCTTGTACAGGGGGCGGAGCCTATGGAAAAACGGCTTTGC
CGCGGCCCTCTCACTTCCCTGTTAAGTATCTTCTTGGCATCTTCCAGGAA
ATCTCCGCCCCGTTCTGTAAGCCATTTCCGCTCGCCGCAGTCGAACGACCG
AGCGTAGCGAGTCAGTGAGCGAGGAAGCGGAATATATCCTGTATCACATA
TTCTGCTGACGCACCGGTGCAGCCTTTTTTCTCCTGCCACATGAAGCACT
TCACTGACACCCTCATCAGTGCCAACATAGTAAGCCAGTATACACTCCGC
TAGCGCGCAAGCGGCCGCA

>pMXLS VR (SEQ ID NO: 36)

CACTCTGT⁻GACTTGGT CTCC TGAGACC
 GTCAGACCAAGTTTACTCAT
 ATATACTTTAGATTGATTTAAACTTCATTTTTTAATTTAAAAGGATCTAG
 GTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTT
 TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
 GAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCA
 CCGCTACCAGCGGTGGTTTGTGGCCGATCAAGAGCTACCAACTCTTTT
 TCCGAAGGTAACGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTC
 TAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCT
 ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGA
 TAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGG
 CGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAG
 CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
 CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA
 GGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAACGCCTGG
 TATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATT
 TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAACGCGCAGCAACG
 CGGCCTTTTTACGGTTCCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTT
 TTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGA
 GTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAG
 TGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACCGCCTCTCCCCGCG
 CGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
 AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
 CACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT
 GTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGC
 CAAGCTTGCGATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGC
 TCGAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGG
 CGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGC
 GTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC
 CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG
 CGGTATTTACACACCGCATATGGTGCCTCTCAGTACAATCTGCTCTGATG
 CCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC
 TGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC
 GGTCTCA CGAG TGAGACCAAGTCAACGAAGTG TGCGGCCGCTTCG
 TGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTAT
 CCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAG
 GAAGAGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCG
 GTGGTAACGGCGCAGTGCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTTT
 GGGGTACAGTCTATGCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGT
 GGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCA
 GTCGCCCTAAAACAAAGTTAAACATCATGAGGGAAGCGGTGATCGCCGAA
 GTATCGACTCAACTATCAGAGGTAGTTGGCGTCATCGAGCGCCATCTCGA
 ACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGGCC
 TGAAGCCACACAGCGATATTGATTTGCTGGTTACGGTGACCGTAAGGCTT
 GATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAACCTTCGGC
 TTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTACCATTGTTG
 TGCACGACGACATCATTCGCTGGCGTTATCCAGCTAAGCGCGAACTGCAA
 TTTGGAGAATGGCAGCGCAATGACATTCTTGCGAGGTATCTTCGAGCCAGC
 CACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACATA
 GCGTTGCCCTGGTAGGTCCAGCGGCGGAGGAACTCTTTGATCCGGTTCCT
 GAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAACGCTATGGAACCTC
 GCCGCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTACGTTGTCC

GCATTTGGTACAGCGCAGTAACCGGCCAAAATCGCGCCGAAGGATGTCGCT
 GCCGACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACT
 TGAAGCTAGACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGC
 GCGCAGATCAGTTGGAAGAATTTGTCCATTACGTAAAAGGCGAGATCACC
 AAGGTAGTCGGCAAATAAGAGCCTGTTTAATAAGATGATCTTCTTGAGAT
 CGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAAACCGCCTT
 GCAGGGCGGTTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAACCGAG
 GTAACCTGGCTTGGAGGAGCGCAGTCACCAAACTTGTCTTTTTCAGTTTAG
 CCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAG
 TGGCTGCTGCCAGTGGTGCTTTTGCATGTCTTTCCGGGTTGGAATCAAGA
 CGATAGTTACCGGATAAGGCGCAGCGGTCGGACTGAACGGGGGGTTCGTG
 CATAAGTCCAGCTTGGAGCGAACTGCCTACCCGGAACAGAGTGTACAGG
 GTGGAATGAGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCG
 AAAGGCAGGAACAGGAGAGCGCACGAGGGAGCCGCCAGGGGGAAACGCCT
 GGTATCTTTATAGTCCTGTCTGGGTTTCGCCACCACTGATTTGAGCGTCAG
 ATTTTCGTGATGCTTGTGAGGGGGGCGGAGCCTATGGAAAAACGGCTTTGC
 CGCGGCCCTCTCACTTCCCTGTTAAGTATCTTCCTGGCATCTTCCAGGAA
 ATCTCCGCCCCGTTCTGAAGCCATTTCCGCTCGCCGCAGTCGAACGACCG
 AGCGTAGCGAGTCAGTGAGCGAGGAAGCGGAATATATCCTGTATCACATA
 TTCTGCTGACGCACCGGTGCAGCCTTTTTTCTCCTGCCACATGAAGCACT
 TCACTGACACCCTCATCAGTGCCAACATAGTAAGCCAGTATACACTCCGC
 TAGCGCGCAAGCGGCCGCA

>pMZLS_VL (SEQ ID NO: 37)

CACTCTGT |GAAGCTCTTCA TCC TGAAGAGC|
 GGTCAGACCAAGTTTACTCAT
 ATATACTTTAGATTGATTTAAAACCTTCATTTTTTAATTTAAAAGGATCTAG
 GTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTT
 TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
 GAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCA
 CCGCTACCAGCGGTGGTTTGTGTGCCGGATCAAGAGCTACCAACTCTTTT
 TCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTC
 TAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCT
 ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGA
 TAAGTCGTGTCTTACCGGGTTGGAATCAAGACGATAGTTACCGGATAAGG
 CGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAG
 CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
 CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA
 GGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGG
 TATCTTTATAGTCCTGTCTGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
 TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACG
 CGGCCTTTTTACGGTTCCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTT
 TTTCTGCGTTATCCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGA
 GTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAG
 TGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACCGCCTCTCCCCGCG
 CGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
 AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
 CACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT
 GTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGC
 CAAGCTTGCATGCCTGCAGGTGCACTCTAGAGGATCCCCGGGTACCGAGC
 TCGAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGG
 CGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGC

GTAATAGCGAAGAGGCCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC
 CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG
 CCGTATTTTACACCGCATATGGTGCACCTCTCAGTACAATCTGCTCTGATG
 CCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC
 TGACGGGCTTGCTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACC
 GCTCTTCA CGA AGAGCTT C ACGAAGTG TGCGGCCGCTTCG
 TGCGCGGAACCCCTATTTGTTTTATTTTCTAAATACATTCAAATATGTAT
 CCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAG
 GAAGAGTATGCGCTCACGCAACTGGTCCAGAACCCTTGACCGAACGCAGCG
 GTGGTAACGGCGCAGTGGCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTTT
 GGGGTACAGTCTATGCCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGT
 GGGTGCATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCA
 GTCGCCCTAAAACAAAGTTAAACATCATGAGGGAAGCGGTGATCGCCGAA
 GTATCGACTCAACTATCAGAGGTAGTTGGCGTCATCGAGCGCCATCTCGA
 ACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGGCC
 TGAAGCCACACAGCGATATTGATTTGCTGGTTACGGTGACCGTAAGGCTT
 GATGAAACAACCGCGCGAGCTTTGATCAACGACCTTTTGGAACTTCGGC
 TTCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTCACCATTTGTTG
 TGCACGACGACATCATTTCCGTGGCGTTATCCAGCTAAGCGCGAACTGCAA
 TTTGGAGAATGGCAGCGCAATGACATTTCTTGCAGGTATCTTCGAGCCAGC
 CACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACATA
 GCGTTGCCTTGGTAGGTCCAGCGGCGGAGGAACTCTTTGATCCGGTTCCCT
 GAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAACGCTATGGAACCTC
 GCCGCCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTACGTTGTCCC
 GCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTGCGCT
 GCCGACTGGGCAATGGAGCGCCTGCCGCGCCAGTATCAGCCCGTCATACT
 TGAAGCTAGACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGC
 GCGCAGATCAGTTGGAAGAATTTGTCCATTACGTAAAAGGCGAGATCACC
 AAGGTAGTCGGCAAATAAGAGCCTGTTTAATAAGATGATCTTCTTGAGAT
 CGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAAACCGCCTT
 GCAGGGCGGTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAACCGAG
 GTAACCTGGCTTGGAGGAGCGCAGTCACCAAACTTGTCTTTTCAGTTTAG
 CCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAG
 TGGCTGCTGCCAGTGGTGCTTTTGCATGTCTTTCCGGGTGGACTCAAGA
 CGATAGTTACCGGATAAGGCGCAGCGGTGCGACTGAACGGGGGGTTCGTG
 CATAACAGTCCAGCTTGGAGCGAACTGCCTACCCGGAAGTGAAGTGTGAGG
 GTGGAATGAGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCG
 AAAGGCAGGAACAGGAGAGCGCACGAGGGAGCCGCCAGGGGGAAACGCCT
 GGTATCTTTATAGTCCTGTGCGGTTTTCGCCACCACTGATTTGAGCGTCAG
 ATTTTCGTGATGCTTGTCAGGGGGGCGGAGCCTATGGAAAAACGGCTTTCG
 CGCGGCCCTCTCACTTCCCTGTTAAGTATCTTCTTGGCATCTTCCAGGAA
 ATCTCCGCCCCGTTGTAAGCCATTTCCGCTCGCCGCGAGTCGAACGACCG
 AGCGTAGCGAGTCAGTGAGCGAGGAAGCGGAATATATCCTGTATCACATA
 TTCTGCTGACGCACCGGTGCAGCCTTTTTTCTCCTGCCACATGAAGCACT
 TCACTGACACCCTCATCAGTGCCAACATAGTAAGCCAGTATACACTCCGC
 TAGCGCGCAAGCGGCCGCA

>pMZLS_VM (SEQ ID NO: 38)

CACTCTGT GAAGCTCT TCC TGAAGAGC
GGTCAGACCAAGTTTACTCAT
ATATACTTTAGATTGATTTAAACTTCATTTTTTAATTTAAAAGGATCTAG
GTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTT
TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
GAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCA
CCGCTACCAGCGGTGGTTTGTTCGCCGGATCAAGAGCTACCAACTCTTTT
TCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCTTC
TAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAAGTCTGTAGCACCCGCT
ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGA
TAAGTCGTGTCCTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGG
CGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAG
CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
CGCCACGCTTCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA
GGGTCCGAACAGGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGG
TATCTTTATAGTCCTGTTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACG
CGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTT
TTTCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGA
GTGAGCTGATACCGCTCGCCGCGAGCCGAACGACCGAGCGCAGCGAGTCAG
TGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACCGCCTCTCCCCGCG
CGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
CACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT
GTGAGCGGATAACAATTTTCACACAGGAACAGCTATGACCATGATTACGC
CAAGCTTGCGTGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGC
TCGAATTCAGTGCCGTCGTTTTACAACGTCGTGACTGGGAAAAACCTTG
CGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTTCGCCAGCTGGC
GTAATAGCGAAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC
CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG
CGGTATTTACACCCGCATATGGTGCCTCTCAGTACAATCTGCTCTGATG
CCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGGCC
TGACGGGCTTGCTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACC
GCTCTTCA CGA AGAGCTC ACGAAGTG TCGGGCCGCTTCG
TGCGCGGAACCCCTATTTGTTTTATTTTTCTAAATACATTCAAATATGTAT
CCGCTCATGAGACAATAACCCGTGATAAATGCTTCAATAATATTGAAAAAG
GAAGAGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCG
GTGGTAACGGCGCAGTGGCGTTTTTCATGGCTTGTTATGACTGTTTTTTTTT
GGGGTACAGTCTATGCCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGT
GGGTGCGATGTTTATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCA
GTCGCCCTAAAAACAAAGTTAAACATCATGAGGGAAGCGGTGATCGCCGAA
GTATCGACTCAACTATCAGAGGTAGTTGGCGTCATCGAGCGCCATCTCGA
ACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGGCC
TGAAGCCACACAGCGATATTGATTTGCTGGTTACGGTGACCGTAAGGCTT
GATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAACTTCGGC
TTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTACCATTGTTG
TTCACGACGACATCATTCGCTGGCGTTATCCAGCTAAGCGCGAACTGCAA
TTTTGGAGAATGGCAGCGCAATGACATTCTTGCAAGGTATCTTCGAGCCAGC
CACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACATA
GCGTTGCCTTGGTAGGTCCAGCGGCGGAGGAACTCTTTGATCCGGTTCTT
GAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAACGCTATGGAACCTC
GCCGCCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTACGTTGTCCC

GCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTCGCT
 GCCGACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACT
 TGAAGCTAGACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGC
 GCGCAGATCAGTTGGAAGAATTTGTCCATTACGTAAAAGGCGAGATCACC
 AAGGTAGTTCGGCAAATAAGAGCCTGTTTAATAAGATGATCTTCTTGAGAT
 CGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAAACC GCCTT
 GCAGGGCGGTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAACCGAG
 GTAAC TGGCTTGGAGGAGCGCAGTCACCAAACTTGTCTTTTCAGTTTAG
 CCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAG
 TGGCTGCTGCCAGTGGTGCTTTTGCATGTCTTCCGGGTTGGACTCAAGA
 CGATAGTTACCGGATAAGGCGCAGCGGTCGGACTGAACGGGGGGTTTCGTG
 CATAAGTCCAGCTTGGAGCGAACTGCCTACCCGGAAC T GAGTGT CAGGC
 GTGGAATGAGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCG
 AAAGGCAGGAACAGGAGAGCGCACGAGGGAGCCGCCAGGGGGAAACGCCT
 GGTATCTTTATAGTCCTGTCTGGGTTTTCGCCACCACTGATTTGAGCGTCAG
 ATTTTCGTGATGCTTGT CAGGGGGGCGGAGCCTATGGAAAAACGGCTTTGC
 CGCGGCCCTCTCACTTCCCTGTTAAGTATCTTCCTGGCATCTTCCAGGAA
 ATCTCCGCCCCGTTTCGTAAGCCATTTCCGCTCGCCGCAGTCGAACGACCG
 AGCGTAGCGAGTCAGTGAGCGAGGAAGCGGAATATATCCTGTATCACATA
 TTCTGCTGACGCACCGGTGCAGCCTTTTTTCTCCTGCCACATGAAGCACT
 TCACTGACACCCTCATCAGTGCCAACATAGTAAGCCAGTATACACTCCGC
 TAGCGCGCAAGCGGCCGCA

>pMZLS_VR (SEQ ID NO: 39)

CACTCTGT GAAGCTCT TCC TGAAGAGC
 GGTGAGACCAAGTTTACTCAT
 ATATACTTTTAGATTGATTTAAAACCTTCATTTTTAATTTAAAAGGATCTAG
 GTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTT
 TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTT
 GAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCA
 CCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTT
 TCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTC
 TAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCT
 ACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGA
 TAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGG
 CGCAGCGGTGCGGCTGAACGGGGGGTTGCTGCACACAGCCCAGCTTGGAG
 CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
 CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCA
 GGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCCTGG
 TATCTTTATAGTCCTGTGCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
 TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACG
 CGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTT
 TTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGA
 GTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAG
 TGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACCGCCTCTCCCCGCG
 CGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
 AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
 CACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATT
 GTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGC
 CAAGCTTGATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGC
 TCGAATTCCTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGG
 CGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGC
 GTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC

CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG
 CCGTATTTTACACCCGCATATGGTGCCTCTCAGTACAATCTGCTCTGATG
 CCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC
 TGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACC
 GCTCTTCA CGA TGAAGAGCTTC ACGAAGTG TGCGGCCCGCCTTCG
 TGCGCGGAACCCCTATTTGTTTTATTTTCTAAATACATTCAAATATGTAT
 CCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAG
 GAAGAGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCG
 GTGGTAACGGCGCAGTGGCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTTT
 GGGGTACAGTCTATGCCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGT
 GGGTTCGATGTTTATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCA
 GTCGCCCTAAAAACAAAGTTAAACATCATGAGGGAAGCGGTGATCGCCGAA
 GTATCGACTCAACTATCAGAGGTAGTTGGCGTCATCGAGCGCCATCTCGA
 ACCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGGCC
 TGAAGCCACACAGCGATATTGATTTGCTGGTTACGGTGACCGTAAGGCTT
 GATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAACTTCGGC
 TTCCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTCACCATTTGTG
 TGCACGACGACATCATTTCCGTGGCGTTATCCAGCTAAGCGCGAACTGCAA
 TTTGGAGAATGGCAGCGCAATGACATTCTTGCAAGGTATCTTCGAGCCAGC
 CACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACATA
 GCGTTGCCTTGGTAGGTCCAGCGGCGGAGGAACTCTTTGATCCGGTTCCCT
 GAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAACGCTATGGAATC
 GCCGCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTACGTTGTCCC
 GCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTGCGCT
 GCCGACTGGGCAATGGAGCGCCTGCCGCGCCAGTATCAGCCCGTCATACT
 TGAAGCTAGACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGC
 GCGCAGATCAGTTGGAAGAATTTGTCCATTACGTAAAAGGCGAGATCACC
 AAGGTAGTCGGCAAATAAGAGCCTGTTTAATAAGATGATCTTCTTGAGAT
 CGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAAACCGCCTT
 GCAGGGCGGTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAACCGAG
 GTAACCTGGCTTGGAGGAGCGCAGTCACCAAACTTGTCTTTTCAGTTTAG
 CCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAG
 TGGCTGCTGCCAGTGGTGCTTTTGCATGTCTTTCCGGGTGGACTCAAGA
 CGATAGTTACCGGATAAGGCGCAGCGGTGCGACTGAACGGGGGGTTCGTG
 CATAAGTCCAGCTTGGAGCGAACTGCCTACCCGGAAGTGAAGTGTACGGC
 GTGGAATGAGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCG
 AAAGGCAGGAACAGGAGAGCGCACGAGGGAGCCGCCAGGGGGAAACGCCT
 GGTATCTTTATAGTCCCTGTGCGGTTTTCGCCACCACTGATTTGAGCGTCAG
 ATTTTCGTGATGCTTGTGAGGGGGGCGGAGCCTATGGAAAAACGGCTTTCG
 CGCGGCCCTCTCACTTCCCTGTTAAGTATCTTCTGCGCATCTTCAGGAA
 ATCTCCGCCCCGTTCTGTAAGCCATTTCCGCTCGCCGAGTCGAACGACCG
 AGCGTAGCGAGTCAGTGAGCGAGGAAGCGGAATATATCCTGTATCACATA
 TTCTGCTGACGCACCGGTGCAGCCTTTTTTCTCCTGCCACATGAAGCACT
 TCACTGACACCCCTCATCAGTGCCCAACATAGTAAGCCAGTATACACTCCGC
 TAGCGCGCAAGCGGCCGA

>pMXP A4E (SEQ ID NO: 40)

CACTCTGT GACTTGGTCTCA TGCC TGAGACC GCAGCTGGCAGACAGG
 TTCCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTA
 GCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTA
 TGTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTAT
 GACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATC
 CCCGGGTACCGAGCTCGAATTCAGTGGCCGTCGTTTTACAACGTCGTGAC
 TGGGAAAACCCCTGGCGTTACCCAACTTAATCGCCTTGACAGCACATCCCCC
 TTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCCGACCGATCGCCCTTCCC
 AACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTC
 CTTACGATCTGTGCGGTATTTACACACCGCATATGGTGCACTCTCAGTAC
 AATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACAC
 CCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGA
 CAAGCTGTGAC GGTCTCA CAGA TGAGACCAAGTC ACGAAGTGTTTAA
 ACTATCAGTGTTTGACAGGATATATTGGCGGGTAAACCTAAGAGAAAAGA
 GCGTTTATTAGAATAATCGGATATTTAAAAGGGCGTGAAAAGGTTTATCC
 GTTCGTCCATTTGTATGTGCATGCCAACACAGGGTTCCCCAGATCAGGC
 GCTGGCTGCTGAACCCCCAGCCGGAAGTACCCCCACAAGGCCCTAGCGTT
 TGCAATGCACCAGGTCATCATTGACCCAGGCGTGTTCCACCAGGCCGCTG
 CCTCGCAACTCTTCGAGGCTTCGCCGACCTGCTCGCGCCACTTCTTCAC
 GCGGGTGGAATCCGATCCGCACATGAGGCGGAAGGTTTCCAGCTTGAGCG
 GGTACGGCTCCCGGTGCGAGCTGAAATAGTCGAACATCCGTGCGGCCGTC
 GGCGACAGCTTGCGGTACTTCTCCCATATGAATTTCTGTAGTGCGGCC
 AGCAAACAGCACGACGATTTCTCTCGTCGATCAGGACCTGGCAACGGGACG
 TTTTCTTGCCACGGTCCAGGACGCGGAAGCGGTGACGACGACACCGAT
 TCCAGGTGCCAACGCGGTGCGACGTGAAGCCCATCGCCGTGCGCTGTAG
 GCGCGACAGGCATTCCTCGGCCTTCGTGTAATACCGGCCATTGATCGACC
 AGCCCAGGTCTTGCCAAAGCTCGTAGAACGTGAAGGTGATCGGCTCGCCG
 ATAGGGGTGCGCTTCGCGTACTCCAACACCTGCTGCCACACCGATTGCTG
 ATCGTCCGCGCCGAGCTCGACGCCGTTGTAGGTGATCTTCACGTTCCTTG
 TGAGCTGGAAAATGACCTTGTTTTGCAGCGCCTCGCGCGGATTTTCTTG
 TTGCGCGTGTTGAACAGGGCAGAGCGGGCCGTGTCGTTTGGCATCGCTCG
 CATCGTGTCCGGCCACGGCGCAATATCGAACAAGGAAAGCTGCATTTCTC
 TGATCTGCTGCTTCGTGTGTTTTCAGCAACGCGGCCCTGCTTGGCCTCGCTG
 ACCTGTTTTTGCCAGGTCCTCGCCGGCGGTTTTTTCGCTTCTTGGTCTCAT
 AGTTCCTCGCGTGTGATGGTCATCGACTTCGCCAAACCTGCCGCCCTCCT
 GTTCAAGACGACGCGAACGCTCCACGGCGGCCGATGGCGCGGGCAGGGCA
 GGGGGAGCCAGTTGCACGCTGTGCGCTCGATCTTGGCCGTAGCTTGCTG
 GACCATCGAGCCGACGACTGGAAGGTTTCGCGGGGCGCACGCATGACGG
 TGCGGCTTGCGATGGTTTCGGCATCCTCGGCGGAAAACCCCGCTCGATC
 AGTTCCTTGCCGTGTATGCCTTCCGGTCAAACGTCCGATTCAATCACCCCTC
 TTGCGGGATTGCCCCGACTCACGCCGGGGCAATGTGCCCTTATTCCTGAT
 TTGACCCGCCCTGGTGCCTTGGTGTCCAGATAATCCACCTTATCGGCAATG
 AAGTCGGTCCCGTAGACCGTCTGGCCGTCCTTCTCGTACTTGGTATTCGG
 AATCTTGCCCTGCACGAATACCAGCGACCCCTTGCCCCAAATACTTGCCGT
 GGGCCTCGGCCTGAGAGCCAAAACACTTGATGCGGAAGAACTCGGTGCGC
 TCCTGCTTGTGCGCCGGCATCGTTGCGCCACATCTAGGATCTGCCAGGAAC
 CGTAAAAAGGCCGCTTGTGGCCTTTTCCATAGGCTCCGCCCCCTGA
 CGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAACCCGACAG
 GACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT
 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTC
 GGGAAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG
 TGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTACG

CCGGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGT
 AAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA
 GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC
 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCC
 AGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA
 CCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGA
 AAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGC
 TCAGTGGAAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAA
 AAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA
 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAAT
 CAGTGAGGCACCTATCTCAGCGATCTGTCTATTTGTTTCATCCATAGTTG
 CCTGACTCCCCGTGCTGTAGATAACTACGATACGGGAGGGCTTACCATCT
 GGCCCCAGTGCTGCAATGATACCGCGAGAACCACGCTCACCGGCTCCAGA
 TTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTC
 CTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT
 AGAGTAAAGTAGTTCCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGC
 TACAGGCATCGTGGTGTCAAGCTCGTCTGTTTGGTATGGCTTCATTCCAGT
 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAA
 AAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGC
 CGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTG
 TCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG
 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTC
 AATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA
 TTGGAACCGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTG
 AGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATC
 TTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATC
 CCGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC
 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAG
 CGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC
 GCACGAATTGGCCAGCGCTGCCATTTTGGGGTGAGGCCGTTTCGCGGCCG
 AGGGGCGCAGCCCTGGGGGGATGGGAGGCCCGCGTTAGCGGGCCGGGAG
 GGTTCGAGAAGGGGGGGCACCCCCCTTCGGCGTGCGCGGTACGCGCACA
 GGGCGCAGCCCTGGTTAAAAACAAGGTTTATAAATATTGGTTTAAAAGCA
 GGTTAAAAGACAGGTTAGCGGTGGCCGAAAAACGGGCGGAAACCCTTGCA
 AATGCTGGATTTTCTGCCTGTGGACAGCCCCCTCAAATGTCAATAGGTGCG
 CCCCTCATCTGTCAAGCACTCTGCCCCCTCAAGTGTCAAGGATCGCGCCCCCT
 CATCTGTCAGTAGTCGCGCCCCCTCAAGTGTCAATACCGCAGGGCACTTAT
 CCCCAGGCTTGTCCACATCATCTGTGGGAAACTCGCGTAAAATCAGGCGT
 TTTGCGCGATTTGCGAGGCTGGCCAGCTCCACGTGCGCGGCCGAAATCGA
 GCCTGCCCCCTCATCTGTCAACGCCGCGCCGGGTGAGTGGCCCCCTCAAGT
 GTCAACGTCCGCCCCCTCATCTGTCAAGTGAAGGCCAAGTTTTCCGCGAGGT
 ATCCACAACGCCGGCGGCCGCGGTGTCTCGCACACGGCTTCGACGGCGTT
 TCTGGCGCGTTTGCAGGGCCATAGACGGCCGCCAGCCCAGCGGCGAGGGC
 AACCAGCCCCGTGAGCGTCGCAAAGGAGATCCTGATCTGACTGATGGGCT
 GCCTGTATCGAGTGGTGATTTTGTGCCGAGCTGCCGGTCGGGGAGCTGTT
 GGCTGGCTGGTGGCAGGATATATTGGTGTAAACAAATTGACGCTTAGA
 CAACTTAATAACACATTGCGGACGTTTTTAATGTACTGGGGTGGATGCAG
 TGGGCCC

>pMYP_A5E (SEQ ID NO: 41)

CACTCTGT GAAGACGC TGCC TTGTCTTC GCAGCTGGCAGCAGAGGTTT
 CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCT
 CACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGT
 TGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGAC
 CATGATTACGCCAAGCTTGCATGCCTGCAGGTGCACTCTAGAGGATCCCC
 GGGTACCGAGCTCGAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGG
 GAAAACCCCTGGCGTTACCCAACTTAATCGCCTTGACAGCACATCCCCCTTT
 CGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAAC
 AGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTCTCCTT
 ACGCATCTGTGCGGTATTTACACCGCATATGGTGCACTCTCAGTACAAT
 CTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCG
 CTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAA
 GCTGTGAC GAAGACAA CAGA GCGTCTTC ACGAAGTGGTTTTAACTATC
 AGTGTTTTGACAGGATATATTGGCGGGTAAACCTAAGAGAAAAGAGCGTTT
 ATTAGAATAATCGGATATTTAAAAGGGCGTGAAAAGGTTTATCCGTTTCGT
 CCATTTGTATGTGCATGCCAACACAGGGTTCCCCAGATCAGGCGCTGGC
 TGCTGAACCCCCAGCCGGAAGTACCCCAACAGGCCCTAGCGTTTGCAT
 GCACCAGGTCATCATTGACCCAGGCGTGTTCCACCAGGCCGCTGCCTCGC
 AACTCTTCGAGGCTTCGCCGACCTGCTCGCGCCACTTCTTCACGCGGGT
 GGAATCCGATCCGCACATGAGGCGGAAGGTTTTCCAGCTTGAGCGGGTACG
 GCTCCCGGTGCGAGCTGAAATAGTCGAACATCCGTCGGGCGGTCGGCGAC
 AGCTTGCGGTACTTCTCCCATATGAATTTTCGTGTAGTGGTCGCCAGCAAA
 CAGCACGACGATTTCTCGTCGATCAGGACCTGGCAACGGGACGTTTTCT
 TGCCACGGTCCAGGACGCGGAAGCGGTGCAGCAGCGACACCGATTCCAGG
 TGCCCAACGCGGTTCGGACGTGAAGCCCATCGCCGTCGCTGTAGGCGCGA
 CAGGCATTCTCGGCCTTCGTGTAATACCGGCCATTGATCGACCAGCCCA
 GGTCTTGCAAAGCTCGTAGAACGTGAAGGTGATCGGCTCGCCGATAGGG
 GTGCGCTTCGCGTACTCCAACACCTGCTGCCACACAGTTTCGTATCGTC
 GGCCCGCAGCTCGACGCGCGGTGTAGGTGATCTTCACGTCTTGTGTGACGT
 GGAAAATGACCTTGTTTTGACGCGCCTCGCGCGGGATTTTCTTGTGCGC
 GTGGTGAACAGGGCAGAGCGGGCGGTGTGTTTTGGCATCGCTCGCATCGT
 GTCCGGCCACGGCGCAATATCGAACAAGGAAAGCTGCATTTCTTGATCT
 GCTGCTTCGTGTGTTTTAGCAACGCGGCCTGCTTGGCCTCGCTGACCTGT
 TTTGCCAGGTCCTCGCCGGCGGTTTTTCGCTTCTTGGTTCGTATAGTTCC
 TCGCGTGTGATGGTCATCGACTTCGCCAAACCTGCCGCCTCCTGTTCAA
 GACGACGCGAACGCTCCACGGCGGCCGATGGCGCGGGCAGGGCAGGGGGA
 GCCAGTTGCACGCTGTGCGCTCGATCTTGGCCGTAGCTTGCTGGACCAT
 CGAGCCGACGGACTGGAAGGTTTTCGCGGGGCGCACGCATGACGGTGCAGG
 TTGCGATGGTTTTCGGCATCCTCGGCGGAAAACCCCGCGTCGATCAGTTCT
 TGCCTGTATGCCTTCCGGTCAAACGTCCGATTCAATCACCCCTCCTTGCGG
 GATTGCCCCGACTCACGCCGGGGCAATGTGCCCTTATTCCTGATTTGACC
 CGCCTGGTGCCTTGGTGTCCAGATAATCCACCTTATCGGCAATGAAGTCG
 GTCCCGTAGACCGTCTGGCCGTCTTCTCGTACTTGGTATTCGGAATCTT
 GCCCTGCACGAATACCAGCGACCCCTTGCCCCAAATACTTGCCGTGGGCCT
 CGGCCTGAGAGCCAAAACACTTGATGCGGAAGAAGTCGGTGCAGCTCCTGC
 TTGTGCGCCGGCATCGTTGCGCCACATCTAGGATCTGCCAGGAACCGTAAA
 AAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCA
 TCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAACCCGACAGGACTAT
 AAAGATAACAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTT
 CCGACCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAG
 CGTGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGG
 TCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGAC

CGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACA
 CGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA
 GGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGC
 TACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC
 CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG
 GTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAA
 GGATCTCAAGAAGATCCTTTTGATCTTTTCTACGGGGTCTGACGCTCAGTG
 GAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGA
 TCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTAAATCAATCTAA
 AGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA
 GGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGAC
 TCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCC
 AGTGCTGCAATGATACCGCGAGAACCACGCTCACC GGCTCCAGATTTATC
 AGCAATAAACCCAGCCAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCCTGCAA
 CTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTA
 AGTAGTTCCGCAGTTAATAGTTTGC GCAACGTTGTTGCCATTGCTACAGG
 CATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTTCAGCTCCGGTT
 CCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCG
 GTTAGCTCCTTCGGTCCCTCCGATCGTTGTGAGAAAGTGGCCGAGT
 GTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCTATGC
 CATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTC
 TGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACG
 GGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTTGGAA
 AACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCC
 AGTTCGATGTAACCCACTCGTGACCCAACTGATCTTCAGCATCTTTTAC
 TTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAA
 AAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTT
 TTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATA
 CATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCGCGCACGA
 ATTGGCCAGCGCTGCCATTTTGGGGTGAGGCCGTTTCGCGGCCGAGGGGC
 GCAGCCCCGTTGGGGGATGGGAGGCCCGCGTTAGCGGGCCGGGAGGGTTTCG
 AGAAGGGGGGGCACCCCCCTTCGGCGTGCGCGGTCACGCGCACAGGGCGC
 AGCCCTGGTTAAAAACAAGGTTTATAAATATTGGTTTAAAAGCAGGTTAA
 AAGACAGGTTAGCGGTGGCCGAAAAACGGGCGGAAACCCCTTGCAAATGCT
 GGATTTTCTGCCTGTGGACAGCCCCCTCAAATGTCAATAGGTGCGCCCCCTC
 ATCTGTGACGACTCTGCCCCCTCAAGTGTC AAGGATCGCGCCCCCTCATCTG
 TCAGTAGTCGCGCCCCCTCAAGTGTC AATACCGCAGGGCACTTATCCCCAG
 GCTTGTCCACATCATCTGTGGGAAACTCGCGTAAAAATCAGGCGTTTTCGC
 CGATTTGCGAGGCTGGCCAGCTCCACGTCGCGGCCGAAATCGAGCCTGC
 CCCTCATCTGTCAACGCCGCGCCGGGTGAGTCGGCCCCCTCAAGTGTC AAC
 GTCCGCCCCCTCATCTGTGAGTGAGGGCCAAGTTTTCGCGAGGTATCCAC
 AACGCCGGCGGCCGCGGTGTCTCGCACACGGCTTCGACGGCGTTTCTGGC
 GCGTTTGCAGGGCCATAGACGGCCGCCAGCCAGCGGCGAGGGCAACCGAG
 CCCGGTGAGCGTCGCAAAGGAGATCCTGATCTGACTGATGGGCTGCCTGT
 ATCGAGTGGTGATTTTGTGCCGAGCTGCCGGTCGGGGAGCTGTTGGCTGG
 CTGGTGGCAGGATATATTGTGGTGTAACAAAATTGACGCTTAGACAACCTT
 AATAACACATTGCGGACGTTTTTAATGTACTGGGGTGGATGCAGTGGGCC
 C

>pMZP_P6T (SEQ ID NO: 42)

CACTCTGT GAAGCTCTTCA GCG TGAAGAGC GCAGCTGGCAGCAGGTT
 TTCCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAG
 CTCACTCATTAGGCCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTAT
 GTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATG
 ACCATGATTACGCCAAGCTTGCATGCCTGCAGGTGCGACTCTAGAGGATCC
 CCGGGTACCGAGCTCGAATTCAGTGGCCGTGCTTTTACAACGTCGTGACT
 GGGAAAACCCCTGGCGTTACCCAACTTAATCGCCTTGACAGCACATCCCCCT
 TTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCA
 ACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCC
 TTACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACCTCTCAGTACA
 ATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACC
 CGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGAC
 AAGCTGTGAC GCTCTTCA TAG TGAAGAGCTTC ACGAAGTGTTTAAAC
 TATCAGTGTTTGACAGGATATATTGGCGGGTAAACCTAAGAGAAAAGAGC
 GTTTATTAGAATAATCGGATATTTAAAAGGGCGTGAAAAGGTTTATCCGT
 TCGTCCATTTGTATGTGCATGCCAACACAGGGTTCCCCAGATCAGGCGC
 TGGCTGCTGAACCCCCAGCCGGAAGTGAACCCACAAGGCCCTAGCGTTTG
 CAATGCACCAGGTCATCATTGACCCAGGCGTGTTCCACCAGGCCGCTGCC
 TCGCAACTCTTCGCAGGCTTCGCCGACCTGCTCGCGCCACTTCTTCACGC
 GGGTGGAAATCCGATCCGCACATGAGGCGGAAGGTTTCCAGCTTGAGCGGG
 TACGGCTCCCGGTGCGAGCTGAAATAGTGAACATCCGTCCGGGCCGTGCG
 CGACAGCTTGCGGTACTTCTCCCATATGAATTTCTGTAGTGGTCCGCCAG
 CAAACAGCACGACGATTTCTCTCGTCGATCAGGACCTGGCAACGGGACGTT
 TTCTTGCCACGGTCCAGGACGCGGAAGCGGTGCAGCAGCGACACCGATTG
 CAGGTGCCAACGCGGTGCGACGTGAAGCCCATCGCCGTGCGCTGTAGGC
 GCGACAGGCATTCCTCGGCCTTCGTGTAATACCGGCCATTGATCGACCAG
 CCCAGGTCCCTGGCAAAGCTCGTAGAACGTGAAGGTGATCGGCTCGCCGAT
 AGGGGTGCGCTTCGCGTACTCCAACACCTGCTGCCACACAGTTCTGTGAT
 CGTCGGCCCGCAGCTCGACGCGCGGTGTAGGTGATCTTCACGCTCTGTG
 ACGTGGAATGACCTTGTTTTGACGCGCCTCGCGCGGGATTTTCTTGTT
 GCGCGTGGTGAACAGGGCAGAGCGGGCCGTGTCGTTTGGCATCGCTCGCA
 TCGTGTCCGGCCACGGCGCAATATCGAACAAGGAAAGCTGCATTTCTTG
 ATCTGCTGCTTCGTGTGTTTCAGCAACGCGGCCTGCTTGGCCTCGCTGAC
 CTGTTTTGCCAGGTCCTCGCCGCGGTTTTTTCGCTTCTTGGTCTGTCATAG
 TTCTTCGCGTGTGATGGTTCATCGACTTCGCCAAACCTGCCGCCCTCCTGT
 TCAAGACGACGCGAACGCTCCACGGCGGCCGATGGCGCGGGCAGGGCAGG
 GGGAGCCAGTTGCACGCTGTGCGCTCGATCTTGGCCGTAGCTTGCTGGA
 CCATCGAGCCGACGGACTGGAAGGTTTCGCGGGCGCACGCATGACGGTG
 CGGCTTGCGATGGTTTTCGGCATCCTCGGCGGAAAACCCCGCTCGATCAG
 TTCTTGCCGTGTATGCCTTCCGGTCAAACGTCCGATTCAATCACCCCTCCTT
 GCGGGATTGCCCCGACTCACGCCGGGGCAATGTGCCCTTATTCCTGATTT
 GACCCGCCTGGTGCCTTGGTGTCCAGATAATCCACCTTATCGGCAATGAA
 GTCGGTCCCGTAGACCGTCTGGCCGTCCTTCTCGTACTTGGTATTCGGAA
 TCTTGCCCTGCACGAATACCAGCGACCCCTTGCCCCAAATACTTGCCGTGG
 GCCTCGGCCCTGAGAGCCAAAACACTTGATGCGGAAGAAGTCGGTGCGCTC
 CTGCTTGTGCGCCGGCATCGTTGCGCCACATCTAGGATCTGCCAGGAACCG
 TAAAAAGGCCCGGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACG
 AGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAACCCGACAGGA
 CTATAAAGATACCAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC
 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGG
 GAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTG
 TAGGTCTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCC

CGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAA
GACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA
GCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA
CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAG
TTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
GCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAA
AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTC
AGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAA
AGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT
CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCA
GTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCC
TGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGG
CCCCAGTGCTGCAATGATACCGCGAGAACCACGCTCACC GGCTCCAGATT
TATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCT
GCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG
AGTAAGTAGTTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTA
CAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTACAGCTCC
GGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAA
AGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCG
CAGTGTTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC
ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC
ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAA
TACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATT
GGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAG
ATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTT
TTACTTTTACCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATGCC
GCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT
CCTTTTTCAATATTATTGAAGCATTTTATCAGGGTTATTGTCTCATGAGCG
GATACATATTTGAATGTATTTAGAAAAATAAAACAAATAGGGGTTCCGCGC
ACGAATTGGCCAGCGCTGCCATTTTTTGGGGTGAGGCCGTTTCGCGGCCGAG
GGGCGCAGCCCCTGGGGGGATGGGAGGCCCGCGTTAGCGGGCCGGGAGGG
TTCGAGAAGGGGGGGCACCCCCCTTCGGCGTGC GCGGTACGCGCACAGG
GCGCAGCCCTGGTTAAAAACAAGGTTTATAAATATTGGTTTAAAAAGCAGG
TAAAAGACAGGTTAGCGGTGGCCGAAAAACGGGCGGAAACCCTTGCAAA
TGCTGGATTTTCTGCCTGTGGACAGCCCCTCAAATGTCAATAGGTGCGCC
CCTCATCTGT CAGCACTCTGCCCTCAAGTGTCAAGGATCGCGCCCCTCA
TCTGT CAGTAGTCGCGCCCCCTCAAGTGTCAATACCGCAGGGCACTTATCC
CCAGGCTTGTCCACATCATCTGTGGGAAACTCGCGTAAATCAGGCGTTT
TCGCCGATTTGCGAGGCTGGCCAGCTCCACGTCGCCGGCCGAAATCGAGC
CTGCCCTCATCTGTCAACGCCGCGCGGGGTGAGTCGGCCCCCTCAAGTGT
CAACGTCCGCCCTCATCTGT CAGTGAGGGCCAAGTTTTCCGCGAGGTAT
CCACAACGCCGGCGGCCGCGGTGTCTCGCACACGGCTTCGACGGCGTTTC
TGGCGCGTTTGCAGGGCCATAGACGGCCGCCAGCCAGCGGCGAGGGCAA
CCAGCCCGGTGAGCGTCGCAAAGGAGATCCTGATCTGACTGATGGGCTGC
CTGTATCGAGTGGTGATTTTGTGCCGAGCTGCCGGTCGGGGAGCTGTTGG
CTGGCTGGTGGCAGGATATATTGTGGTGTAACAAATTGACGCTTAGACA
ACTTAATAACACATTGCGGACGTTTTTAAATGTACTGGGGTGGATGCAGTG
GGCCC

>pMXBG_A71 (SEQ ID NO: 43)

CACTCTGT GACTTGGTCTCA TGCC TGAGACC GTCAGACCAAGTTTACT
 CATATATACTTTAGATTGATTTAAACTTCATTTTTTAATTTAAAAGGATC
 TAGGTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGA
 GTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTT
 CTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAA
 CCACCGCTACCAGCGGTGGTTTGGTTGCCGGATCAAGAGCTACCAACTCT
 TTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCC
 TTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCG
 CCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGG
 CGATAAGTCGTGTCTTACC GGTTGGACTCAAGACGATAGTTACCGGATA
 AGGCGCAGCGGTGCGGCTGAACGGGGGGTTTCGTGCACACAGCCAGCTTG
 GAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGA
 AAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCG
 GCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCC
 TGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCG
 ATTTTGTGTATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCA
 ACGCGGCCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATG
 TTCTTTCCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTT
 TGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGT
 CAGTGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACCGCCTCTCCCC
 GCGCGTTGGCCGATTCTTAATGCAGCTGGCAGCAGAGGTTCCCGACTG
 GAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATT
 AGGCACCCCAAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGA
 ATGTGTAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTA
 CGCCAAAGCTTGCAATGCCTGCAGGTGCACTCTAGAGGATCCCCGGGTACCG
 AGCTCGAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCC
 TGGCGTTACCCAACCTTAATCGCCTTGACGACATCCCCCTTTCGCCAGCT
 GGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGC
 AGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCT
 GTGCGGTATTTACACCGCATATGGTGCCTCTCAGTACAATCTGCTCTG
 ATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCG
 CCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC
GGTCTCA CTTC TGAGACCAAGTC ACGAAGTGTGCGGCCGCCTTCAGCG
 GCGGCACGAATCCATGTGGGAGTTTATTTCTTGACACAGATATTTATGATA
 TAATAACTGAGTAAGCTTAACATAAGGAGGAAAAACATATGTTACGCAGC
 AGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGGTGG
 CTCAAGTATGGGCATCATTTCGCACATGTAGGCTCGGCCCTGACCAAGTCA
 AATCCATGCGGGCTGCTCTTGATCTTTTCGGTCGTGAGTTCGGAGACGTA
 GCCACCTACTCCCAACATCAGCCGGAAGTCCGATTACCTCGGGAACTTGCT
 CCGTAGTAAGACATTCATCGCGCTTGCTGCCTTCGACCAAGAAGCGGTTG
 TTGGCGCTCTCGCGGCTTACGTTCTGCCCAAGTTTGAGCAGCCGCGTAGT
 GAGATCTATATCTATGATCTCGCAGTCTCCGGCGAGCACCGGAGGCAGGG
 CATTGCCACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAACGCGCTTG
 GTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACGATCCCGCAGTG
 GCTCTCTATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGATAT
 CGACCAAGTACCGCCACCTAAGAGCTGGAGCGGCGTAACCGTCGCACAG
 GAAGGACAGAGAAAGCGCGGATCTGGGAAGTGACGGACAGAACGGTCAGG
 ACCTGGATTGGGGAGGCGGTTGCCGCGCTGCTGCTGACGGTGTGACGTT
 CTCTGTTCCGGTCACACCACATACGTTCCGCCATTCTATGCGATGCACA
 TGCTGTATGCCGGTATACCGCTGAAAGTTCTGCAAAGCCTGATGGGACAT
 AAGTCCATCAGTTCAACGGAAGTCTACACGAAGGTTTTTTCGCTGGATGT
 GGCTGCCCGGCTCCGGGTGCAGTTTGCGATGCCGGAGTCTGATGCGGTTG

CGATGCTGAAACAATTATCCTGAGAATAAATGCCTTGGCCTTTATATGGA
 AATGTGGAAGTGGATATGCTGTTTTTGTCTGTTAAACAGAGAAGCT
 GGCTGTTATCCACTGAGAAGCGAACGAAACAGTCGGGAAAATCTCCCATT
 ATCGTAGAGATCCGCATTATTAATCTCAGGAGCCTGTGTAGCGTTTATAG
 GAAGTAGTGTCTGTTCATGATGCCTGCAAGCGGTAACGAAAACGATTTGA
 ATATGCCCTTCAGGAACAATAGAAATCTTCGTGCGGTGTTACGTTGAAGTG
 GAGCGGATTATGTCAGCAATGGACAGAACCACTAATGAACACAGAACCA
 TGATGTGGTCTGTCCCTTTTACAGCCAGTAGTGCTCGCCGCAGTCGAGCGA
 CAGGGCGAAGCCCTCGAGTGAGCGAGGAAGCACCAGGGAACAGCATTAT
 ATATTCTGCTTACACACGATGCCTGAAAAAACTTCCCTTGGGGTTATCCA
 CTTATCCACGGGGATATTTTTTATAATTATTTTTTTTATAGTTTTTAGATC
 TTCTTTTTTTAGAGCGCCTTGTAGGCCTTTATCCATGCTGGTTCAGAGAA
 GGTGTTGTGACAAATTGCCCTTTTCACTGTGACAAATCACCTCAAATGAC
 AGTCCGTGTCTGTGACAAATTGCCCTTAACCCCTGTGACAAATTGCCCTCAG
 AAGAAGCTGTTTTTTCACAAAGTTATCCCTGCTTATTGACTCTTTTTTAT
 TTAGTGTGACAATCTAAAACTTGTACACTTCACATGGATCTGTCATGG
 CGGAAACAGCGGTTATCAATCACAAGAAACGTAATAATAGCCCGCAATC
 GTCCAGTCAAACGACCTCACTGAGGCGGCATATAGTCTCTCCCGGGATCA
 AAAACGTATGCTGTATCTGTTGCTTGACCAGATCAGAAAATCTGATGGCA
 CCCTACAGGAACATGACGGTATCTGCGAGATCCATGTTGCTAAATATGCT
 GAAATATTCGGATTGACCTCTGCGGAAGCCAGTAAGGATATACGGCAGGC
 ATTGAAGAGTTTTCGCGGGGAAGGAAGTGGTTTTTTATCGCCCTGAAGAGG
 ATGCCGGCGATGAAAAAGGCTATGAATCTTTTCCCTTGGTTTTATCAAACGT
 GCGCACAGTCCATCCAGAGGGCTTTACAGTGTACATATCAACCCATATCT
 CATTCCCTTCTTTATCGGGTTACAGAACCGGTTTACGCAGTTTCGGCTTA
 GTGAAACAAAAGAAATCACCAATCCGTATGCCATGCGTTTATACGAATCC
 CTGTGTCAGTATCGTAAGCCGGATGGCTCAGGCATCGTGTCTCTGAAAAAT
 CGACTGGATCATAGAGCGTTACCAGCTGCCTCAAAGTTACCAGCGTATGC
 CTGACTTCCGCCGCCGCTTCCCTGCAGGTCTGTGTTAATGAGATCAACAGC
 AGAACTCCAATGCGCCTCTCATAACATTGAGAAAAAGAAAGGCCGCCAGAC
 GACTCATATCGTATTTTCTTCCGCGATATCACTTCCATGACGACAGGAT
 AGTCTGAGGGTTATCTGTACAGATTTGAGGGTGGTTCGTACATTTGTT
 CTGACCTACTGAGGGTAATTTGTACAGTTTTGCTGTTTCCCTTCAGCCTG
 CATGGATTTTCTCATACTTTTGAAGTGAATTTTAAAGGAAGCCAAATT
 TGAGGGCAGTTTGTACAGTTGATTTCTTCTCTTCCCTTCGTATGTG
 ACCGTATATCGGGGGTTAGTTTCGTATCATGATGAGGGTTGATTATCAC
 AGTTTATTACTCTGAATTGGCTATCCGCGTGTGTACCTCTACCTGGAGTT
 TTTCCACGGTGGATATTTCTTCTTGCCTGAGCGTAAGAGCTATCTGAC
 AGAACAGTTCTTCTTTGCTTCCCTCGCCAGTTTCGCTCGCTATGCTCGGTTA
 CACGGCTGCGGCGAGCGCTAGTGATAATAAGTGAAGTATGAGGTATGTGGTGT
 TCTTATCTCCTTTTGTAGTGTTGCTCTTATTTTAAACAACCTTTGCGGTTT
 TTTGATGACTTTGCGATTTTGTGTTGCTTTGCAGTAAATTGCAAGATTT
 AATAAAAAAACGCAAGCAATGATTAAAGGATGTTTCAAGATGAACTCAT
 GGAAACACTTAACAGTGCATAAACGCTGGTCATGAAATGACGAAGGCTA
 TCGCCATTGCACAGTTTAAATGATGACAGCCCGAAGCGAGGAAAATAAC
 CGGCGCTGGAGAATAGGTGAAGCAGCGGATTTAGTTGGGGTTTCTTCTCA
 GGCTATCAGAGATGCCGAGAAAGCAGGGCGACTACCGCACCCGGATATGG
 AAATTCGAGGACGGGTTGAGCAACGTGTTGGTTATACAATTGAACAAATT
 AATCATATGCGTGATGTGTTTGGTACCGGATTGCGACGTGCTGAGGACGT
 ATTTCCACCGGTGATCGGGGTTGCTGCCCATAAAGGTGGCGTTTACAAAA
 CCTCAGTTTCTGTTTCTGCTCAGGATCTGGCTCTGAAGGGGCTACGT
 GTTTTGTCTCGTGAAGGTAACGACCCCGAGGGAACAGCCTCAATGTATCA
 CGGATGGGTACCAGATCTTCATATTCATGCAGAAGATACTCTCCTGCCTT

TCTATCTTGGGGAAAAGGACGATGTCACTTATGCAATAAAGCCCACTTGC
TGGCCGGGGCTTGACATTATTCCTTCCTGTCTGGCTCTGCACCGTATTGA
AACTGAGTTAATGGGCAAATTTGATGAAGGTAAACTGCCACCGATCCAC
ACCTGATGCTCCGACTGGCCATTGAAACTGTTGCTCATGACTATGATGTC
ATAGTTATTGACAGCGCGCCTAACCTGGGTATCGGCACGATTAATGTCGT
ATGTGCTGCTGATGTGCTGATTGTTCCCACGCCTGCTGAGTTGTTTGA
ACACCTCCGCACTGCAGTTTTTCGATATGCTTCGTGATCTGCTCAAGA
ACGTTGATCTTAAAGGGTTCGAGCCTGATGTACGTATTTGCTTACCAA
ATACAGCAATAGTAATGGCTCTCAGTCCCCGTGGATGGAGGAGCAAATTC
GGGATGCCCTGGGGAAGCATGGTTCTAAAAAATGTTGTACGTGAAACGG
ATGAA GTTGGTAAAGGTCAGATCCGGATGAGAACTGTTTTTGAACAGG
CCATTGATCAACGCAGTTCAACTGGTGCCTGGAGAAATGCTCTTTCTAT
TTGGGAACCTGTCTGCAATGAAATTTTCGATCGTCTGATTAAACCACGCT
GGGAGATTAGATAATGAAGCGTGCGCCTGTTATTCCAAAACATACGCTCA
ATACTCAACCGGTTGAAGATACTTCGTTATCGACACCAGCTGCCCCGAT
GGTGGATTCTGTTAATTGCGCGCGTAGGAGTAATGGCTCGCGGTAATGCC
ATTACTTTGCTGTATGTGGTTCGGGATGTGAAGTTTACTCTTGAAGTGCT
CCGGGGATAGTGTGAGAAAACCTCTCGGGTATGGTCAGGTAATGAACGTG
ACCAGGAGCTGCTGCTTACTGAGGACGCACTGGATGATCTCATCCCTTCT
TTTTCTACTGACTGGTCAACAGACACCGGCGTTTCGGTCAAGAGTATCTGG
TGTCTATAGAAATTGCCGATGGGAGTCGCCGTCGTAAAGCTGCTGCACTT
ACCAGAAAGTGATTATCGATCGTGTCTGGTGGCGAGCTGGATGATGAGCA
GATGGCTGCATTATCCAGATTGGGTAACGATTATCGCCCAACAAGTGCTT
ATGAACGTGGTCAGCGTTATGCAAGCCGATTGCAGAATGAATTTGCTGGAA
ATATTTCTGCGCTGGCTGATGCGGAAAATATTTACGTAAGATTATTACCCG
CTGTATCAACACCGCCAAATTGCCCTAAATCAGTTGTTGCTCTTTTTTCTC
ACCCCGGTGAACCTATCTGCCCCGGTCAGGTGATGCACCTTCAAAAAGCCTT
TACAGATAAAGAGGAATTACTTAAGCAGCAGGCATCTAACCTTCATGAGCA
GAAAAAAGCTGGGGTGATATTTGAAGCTGAAGAAGTTATCACTCTTTTAACT
TCTGTGCTTAAAACGTCATCTGCATCAAGAACTAGTTTAAAGCTCACGACAT
CAGTTTGCCTCCTGGAGCGACAGTATTGTATAAGGGCGATAAAAATGGTGCT
TAACTGGACAGGTCACGTGTTCCAACTGAGTGTATAGAGAAAAATTGAGGCC
ATTCTTAAGGAACTTGAAAAGCCAGCACCCCTGATGCGACCACGTTTTAGTCT
ACGTTATCTGTCTTTACTTAATGTCCTTTGTTACAGGCCAGAAAGCATAACTG
GCCTGAATATTTCTCTCTGGGCCCCACTGTTCCACTTGTATCGTCGGTCTGA
TAATCAGACTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCT
GGGATCTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGA
CCACGGTCCCACTCGTATCGTCGGTCTGATAATCAGACTGGGACCACGGTCCC
ACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCATGGTCCCCTCGTATCGT
CGGTCTGATTATTAGTCTGGAACCACGGTCCCACTCGTATCGTCGGTCTGATT
ATTAGTCTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGG
ACCACGGTCCCACTTGTATTGTGCGATCAGACTATCAGCGTGAGACTACGATT
CCATCAATGCCCTGTCAAGGGCAAGTATTGACATGTCGTGCTAACCTGTAGA
ACGGAGTAACCTCGGTGTGCGGTTGTATGCCGTGCTGTGGATTGCTGTGTGTC
CTGCTTATCCACAACATTTTGCACGCGTTATGTGGACAAAATACCTGGTTACCC
AGGCCGTGCCGGCACGTTAACCGGGCTGCATCCGATGCAAGTGTGTCGCTGT
CGACGCAAGCGCCGCA

>pMXBG_I7G (SEQ ID NO: 44)

CACTCTGT GACTTGGTCTCA CTTC TGAGACC GTCAGACCAAGTTTACT
CATATATACTTTAGATTGATTTAAAACCTTCATTTTTTAATTTAAAAGGATC
TAGGTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGA
GTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTT
CTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAA
CCACCGCTACCAGCGGTGGTTTTGTTTGCCGGATCAAGAGCTACCAACTCT
TTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCC
TTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACC
CCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGG
CGATAAGTCGTGTCTTACCAGGTTGGACTCAAGACGATAGTTACCGGATA
AGGCGCAGCGGTGCGGCTGAACGGGGGGTTTCGTGCACACAGCCAGCTTG
GAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGA
AAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCG
GCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCC
TGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCG
ATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCA
ACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATG
TTCTTTCCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTT
TGAGTGAGCTGATACCGCTCGCCGACCGGAACGACCGAGCGCAGGAGT
CAGTGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACCGCCTCTCCCC
GCGCGTTGGCCGATTCAATTAATGCAGCTGGCAGCAGAGGTTCCCGACTG
GAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATT
AGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGA
ATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTA
CGCCAAGCTTGCAATGCCTGCAGGTGCACTCTAGAGGATCCCCGGGTACCG
AGCTCGAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCC
TGGCGTTACCCAACCTAATCGCCTTGACGACATCCCCCTTTCGCCAGCT
GGCGTAATAGCGAAGAGGCGCCGACCGATCGCCCTTCCCAACAGTTGCGC
AGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCT
GTGCGGTATTTACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTG
ATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCG
CCCTGACGGGCTTGTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGAC
GGTCTCA GAGC TGAGACCAAGTC ACGAAGTGTGCGGCCGCCTTCAGCG
GCCGCACGAATCCATGTGGGAGTTTATTTCTTGACACAGATATTTATGATA
TAATAACTGAGTAAGCTTAACATAAGGAGGAAAAACATATGTTACGCAGC
AGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGGTGG
CTCAAGTATGGGCATCATTGCGACATGTAGGCTCGGCCCTGACCAAGTCA
AATCCATGCGGGCTGCTCTTGATCTTTTCGGTCGTGAGTTCGGAGACGTA
GCCACCTACTCCCAACATCAGCCGGAATCCGATTACCTCGGGAACTTGCT
CCGTAGTAAGACATTCATCGCGCTTGCTGCCTTCGACCAAGAAGCGGTTG
TTGGCGCTCTCGCGGCTTACGTTCTGCCCAAGTTTGAGCAGCCGCGTAGT
GAGATCTATATCTATGATCTCGCAGTCTCCGGCGAGCACCGGAGGCAGGG
CATTGCCACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAACGCGCTTG
GTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACGATCCCGCAGTG
GCTCTCTATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGATAT
CGACCCAAGTACCGCCACCTAAGAGCTGGAGCGCGTAACCGTCGCACAG
GAAGGACAGAGAAAGCGCGGATCTGGGAAGTGACGGACAGAACGGTCAGG
ACCTGGATTGGGGAGGCGGTTGCCGCGCTGCTGCTGACGGTGTGACGTT
CTCTGTTCCGGTCACACCACATACGTTCCGCCATTCTATGCGATGCACA
TGCTGTATGCCGGTATACCGCTGAAAGTTCTGCAAAGCCTGATGGGACAT
AAGTCCATCAGTTCAACGGAAGTCTACACGAAGGTTTTTGCCTGGATGT
GGCTGCCCGGCTCCGGGTGCAGTTTGCGATGCCGGAGTCTGATGCGGTTG

CGATGCTGAAACAATTATCCTGAGAATAAATGCCTTGGCCTTTATATGGA
 AATGTGGAAGTGGATATGCTGTTTTTGTCTGTTAAACAGAGAAGCT
 GGCTGTTATCCACTGAGAAGCGAACGAAACAGTCGGGAAAATCTCCCATT
 ATCGTAGAGATCCGCATTATTAATCTCAGGAGCCTGTGTAGCGTTTATAG
 GAAGTAGTGTCTGTTCATGATGCCTGCAAGCGGTAACGAAAACGATTTGA
 ATATGCCCTTCAGGAACAATAGAAATCTTCGTGCGGTGTTACGTTGAAGTG
 GAGCGGATTATGTCAGCAATGGACAGAACCACTAATGAACACAGAACCA
 TGATGTGGTCTGTCCCTTTTACAGCCAGTAGTGCTCGCCGCAGTCGAGCGA
 CAGGGCGAAGCCCTCGAGTGAGCGAGGAAGCACCAGGGAACAGCATTAT
 ATATTCTGCTTACACACGATGCCTGAAAAAACTTCCCTTGGGGTTATCCA
 CTTATCCACGGGGATATTTTTTATAATTATTTTTTTTATAGTTTTTAGATC
 TTCTTTTTTTAGAGCGCCTTGTAGGCCTTTATCCATGCTGGTTCTAGAGAA
 GGTGTTGTGACAAATTGCCCTTTTCACTGTGACAAATCACCTCAAATGAC
 AGTCTGTCTGTGACAAATTGCCCTTAACCTGTGACAAATTGCCCTCAG
 AAGAAGCTGTTTTTTCACAAAGTTATCCCTGCTTATTGACTCTTTTTTAT
 TTAGTGTGACAATCTAAAACTTGTCACTTCACATGGATCTGTCATGG
 CGGAAACAGCGGTTATCAATCACAAGAAACGTAATAAGCCGCGAATC
 GTCCAGTCAAACGACCTCACTGAGGCGGCATATAGTCTCTCCCGGGATCA
 AAAACGTATGCTGTATCTGTTGCTTGACCAGATCAGAAAATCTGATGGCA
 CCCTACAGGAACATGACGGTATCTGCGAGATCCATGTTGCTAAATATGCT
 GAAATATTCGGATTGACCTCTGCGGAAGCCAGTAAGGATATACGGCAGGC
 ATTGAAGAGTTTTCGCGGGGAAGGAAGTGGTTTTTTATCGCCCTGAAGAGG
 ATGCCGGCGATGAAAAAGGCTATGAATCTTTTCTTGGTTTTATCAAACGT
 GCGCACAGTCCATCCAGAGGGCTTTACAGTGTACATATCAACCCATATCT
 CATTCCCTTCTTTATCGGGTTACAGAACCGGTTTACGCAGTTTCGGCTTA
 GTGAAACAAAAGAAATCACCAATCCGTATGCCATGCGTTTATACGAATCC
 CTGTGTCAGTATCGTAAGCCGGATGGCTCAGGCATCGTGTCTCTGAAAAAT
 CGACTGGATCATAGAGCGTTACCAGCTGCCTCAAAGTTACCAGCGTATGC
 CTGACTTCCGCCGCCGCTTCTGCGAGGTCTGTGTTAATGAGATCAACAGC
 AGAACTCCAATGCGCTCTCATACTTGAAGAAAAGAAAGGCCGCCAGAC
 GACTCATATCGTATTTTCTTCCGCGATATCACTTCCATGACGACAGGAT
 AGTCTGAGGGTTATCTGTACAGATTTGAGGGTGGTTCGTACATTTGTT
 CTGACCTACTGAGGGTAATTTGTACAGTTTTGTCTGTTTCTTCCAGCCTG
 CATGGATTTTCTCATACTTTTGAAGTGAATTTTAAAGGAAGCCAAATT
 TGAGGGCAGTTTGTACAGTTGATTTCTTCTCTTCCCTTCGTGATGTG
 ACCGTATATCGGGGGTTAGTTTCGTATCATGATGAGGGTTGATTATCAC
 AGTTTATTACTCTGAATTGGCTATCCGCGTGTGTACCTCTACCTGGAGTT
 TTTCCACGGTGGATATTTCTTCTTGCCTGAGCGTAAGAGCTATCTGAC
 AGAACAGTTCTTCTTGTCTCTCGCCAGTTGCTCGCTATGCTCGGTTA
 CACGGCTGCGGCGAGCGCTAGTGATAATAAGTGAAGTATGAGGTATGTGGTGT
 TCTTATCTCCTTTTGTAGTGTGCTCTTATTTTAAACAACTTTGCGGTTT
 TTTGATGACTTTGCGATTTTGTGTTGCTTTGCAGTAAATTGCAAGATTT
 AATAAAAAAACGCAAGCAATGATTAAAGGATGTTTCAAGATGAACTCAT
 GGAAACACTTAACAGTGCATAAACGCTGGTCATGAAATGACGAAGGCTA
 TCGCCATTGCACAGTTTAAATGATGACAGCCCGAAGCGAGGAAAATAAC
 CGGCGCTGGAGAATAGGTGAAGCAGCGGATTTAGTTGGGGTTTCTTCTCA
 GGCTATCAGAGATGCCGAGAAAGCAGGGCGACTACCGCACCCGGATATGG
 AAATTCGAGGACGGGTTGAGCAACGTGTTGGTTATACAATTGAACAAATT
 AATCATATGCGTGATGTGTTTGGTACGCGATTGCGACGTGCTGAGGACGT
 ATTTCCACCGGTGATCGGGGTTGCTGCCCATAAAGGTGGCGTTTACAAAA
 CCTCAGTTTCTGTTTCTGCTCAGGATCTGGCTCTGAAGGGGCTACGT
 GTTTTGTCTGTGGAAGGTAACGACCCCGAGGGAACAGCCTCAATGTATCA
 CGGATGGGTACCAGATCTTCATATTCATGCAGAAGATACTCTCCTGCCTT

TCTATCTTGGGGAAAAGGACGATGTCACTTATGCAATAAAGCCCACTTGC
TGGCCGGGGCTTGACATTATTCCTTCCTGTCTGGCTCTGCACCGTATTGA
AACTGAGTTAATGGGCAAATTTGATGAAGGTAAACTGCCACCGATCCAC
ACCTGATGCTCCGACTGGCCATTGAAACTGTTGCTCATGACTATGATGTC
ATAGTTATTGACAGCGCGCCTAACCTGGGTATCGGCACGATTAATGTCGT
ATGTGCTGCTGATGTGCTGATTGTTCCACGCCTGCTGAGTTGTTTGA
ACACCTCCGCACTGCAGTTTTTCGATATGCTTCGTGATCTGCTCAAGA
ACGTTGATCTTAAAGGGTTCGAGCCTGATGTACGTATTTGCTTACCAA
ATACAGCAATAGTAATGGCTCTCAGTCCCCGTGGATGGAGGAGCAAATTCGGG
ATGCCTGGGGAAGCATGGTTCTAAAAAATGTTGTACGTGAAACGGATGAA
GTTGGTAAAGGTCAGATCCGGATGAGAACTGTTTTTGAACAGGCCATTGA
TCAACGCAGTTCAACTGGTGCCTGGAGAAATGCTCTTTCTATTTGGGAAC
CTGTCTGCAATGAAATTTTCGATCGTCTGATTAAACCACGCTGGGAGATT
AGATAATGAAGCGTGCGCCTGTTATTCCAAAACATACGCTCAATACTCAA
CCGGTTGAAGATACTTCGTTATCGACACCAGCTGCCCCGATGGTGGATT
GTTAATTGCGCGCGTAGGAGTAATGGCTCGCGGTAATGCCATTACTTTGC
CTGTATGTGGTTCGGGATGTGAAGTTTACTCTTGAAGTGCTCCGGGGAT
AGTGTTGAGAAAACCTCTCGGGTATGGTCAGGTAATGAACGTGACCAGGA
GCTGCTTACTGAGGACGCACTGGATGATCTCATCCCTTCTTTTCTACTGA
CTGGTCAACAGACACCGGCGTTTCGGTCAAGAGTATCTGGTGTCTATAGAA
ATTGCCGATGGGAGTCGCCGTCGTAAAGCTGCTGCACTTACCGAAAGTGA
TTATCGTGTCTGGTTGGCGAGCTGGATGATGAGCAGATGGCTGCATTAT
CCAGATTGGGTAACGATTATCGCCCAACAAGTGCTTATGAACGTGGTCAG
CGTTATGCAAGCCGATTGCAGAATGAATTTGCTGGAAATATTTCTGCGCT
GGCTGATGCGGAAAATATTTACGTAAGATTATTACCCGCTGTATCAACA
CCGCCAAATTGCCTAAATCAGTTGTTGCTCTTTTTTCTACCCCGGTGAA
CTATCTGCCCCGGTCAGGTGATGCACTTCAAAAAGCCTTTACAGATAAAGA
GGAATTACTTAAGCAGCAGGCATCTAACCTTCATGAGCAGAAAAAAGCTG
GGGTGATATTTGAAGCTGAAGAAGTTATCACTCTTTTAACTTCTGTGCTT
AAAACGTCATCTGCATCAAGAACTAGTTTAAAGCTCACGACATCAGTTTGC
TCCTGGAGCGACAGTATTGTATAAGGGCGATAAAATGGTGTAAACCTGG
ACAGGTCACGTGTTCCAACTGAGTGTATAGAGAAAATGAGGCCATTCTT
AAGGAACTTGAAAAGCCAGCACCCCTGATGCGACCACGTTTTAGTCTACGT
TTATCTGTCTTTACTTAATGTCCTTTGTTACAGGCCAGAAAGCATAACTG
GCCTGAATATTTCTCTCTGGGCCCCACTGTTCCACTTGTATCGTCGGTCTGA
TAATCAGACTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAG
TCTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGA
CCACGGTCCCACTCGTATCGTCGGTCTGATAATCAGACTGGGACCACGGT
CCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCATGGTCCCACTC
GTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGT
CGGTCTGATTATTAGTCTGGAACCACGGTCCCACTCGTATCGTCGGTCTG
ATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTA
GTCTGGGACCACGATCCCACTCGTGTGTGCGGTCTGATTATCGGTCTGGG
ACCACGGTCCCACTTGTATTGTGCGATCAGACTATCAGCGTGAGACTACGA
TTCCATCAATGCCTGTCAAGGGCAAGTATTGACATGTCGTGTAACCTGT
AGAACGGAGTAACCTCGGTGTGCGGTTGTATGCCTGCTGTGGATTGCTGC
TGTTGCTCTGCTTATCCACAACATTTTGCACGCGTTATGTGGACAAAATA
CCTGGTTACCCAGGCCGTGCCGGCACGTTAACCGGGCTGCATCCGATGCA
AGTGTGTCGCTGTGACGCAAGCGGCCGCA

>pMXBG_G7F (SEQ ID NO: 45)

CACTCTGT GACTTGGTCTCA GAGC TGAGACC GTCAGACCAAGTTTACT
CATATATACTTTAGATTGATTTAAACTTCATTTTTTAATTTAAAAGGATC
TAGGTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGA
GTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTT
CTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAA
CCACCGCTACCAGCGGTGGTTTTGTTTGCCGGATCAAGAGCTACCAACTCT
TTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCC
TTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACC
CCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGG
CGATAAGTCGTGTCTTACCAGGTTGGACTCAAGACGATAGTTACCGGATA
AGGCGCAGCGGTGCGGCTGAACGGGGGGTTTCGTGCACACAGCCAGCTTG
GAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGA
AAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCG
GCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCC
TGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCG
ATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCA
ACGCGGCCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATG
TTCTTTCCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTT
TGAGTGAGCTGATACCGCTCGCCGACGCCGAACGACCGAGCGCAGGAGT
CAGTGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACCGCCTCTCCCC
GCGCGTTGGCCGATTCAATTAATGCAGCTGGCAGCAGAGGTTCCCGACTG
GAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATT
AGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGA
ATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTA
CGCCAAAGCTTGCAATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCG
AGCTCGAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCC
TGGCGTTACCCAACCTAATCGCCTTGACGACATCCCCCTTTCGCCAGCT
GGCGTAATAGCGAAGAGGCGCCGACCGATCGCCCTTCCCAACAGTTGCGC
AGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCT
GTGCGGTATTTACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTG
ATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCG
CCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC
GGTCTCA TGTG TGAGACCAAGTC ACGAAGTGTGCGGCCGCCTTCAGCG
GCCGCACGAATCCATGTGGGAGTTTATTTCTTGACACAGATATTTATGATA
TAATAACTGAGTAAGCTTAACATAAGGAGGAAAAACATATGTTACGCAGC
AGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGGTGG
CTCAAGTATGGGCATCATTTCGCACATGTAGGCTCGGCCCTGACCAAGTCA
AATCCATGCGGGCTGCTCTTGATCTTTTCGGTCGTGAGTTCGGAGACGTA
GCCACCTACTCCCAACATCAGCCGGAATCCGATTACCTCGGGAACTTGCT
CCGTAGTAAGACATTCATCGCGCTTGCTGCCTTCGACCAAGAAGCGGTTG
TTGGCGCTCTCGCGGCTTACGTTCTGCCCAAGTTTGAGCAGCCGCGTAGT
GAGATCTATATCTATGATCTCGCAGTCTCCGGCGAGCACCGGAGGCAGGG
CATTGCCACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAACGCGCTTG
GTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACGATCCCGCAGTG
GCTCTCTATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGATAT
CGACCCAAGTACCGCCACCTAAGAGCTGGAGCGCGTAACCGTCGCACAG
GAAGGACAGAGAAAGCGCGGATCTGGGAAGTGACGGACAGAACGGTCAGG
ACCTGGATTGGGGAGGCGGTTGCCGCGCTGCTGCTGACGGTGTGACGTT
CTCTGTTCCGGTCACACCACATACGTTCCGCCATTCTATGCGATGCACA
TGCTGTATGCCGGTATACCGCTGAAAGTTCTGCAAAGCCTGATGGGACAT
AAGTCCATCAGTTCAACGGAAGTCTACACGAAGGTTTTTTCGCTGGATGT
GGCTGCCCGGCTCCGGGTGCAGTTTGCGATGCCGGAGTCTGATGCGGTTG

CGATGCTGAAACAATTATCCTGAGAATAAATGCCTTGGCCTTTATATGGA
 AATGTGGAAGTGGATATGCTGTTTTTGTCTGTTAAACAGAGAAGCT
 GGCTGTTATCCACTGAGAAGCGAACGAAACAGTCGGGAAAATCTCCCATT
 ATCGTAGAGATCCGCATTATTAATCTCAGGAGCCTGTGTAGCGTTTATAG
 GAAGTAGTGTCTGTTCATGATGCCTGCAAGCGGTAACGAAAACGATTTGA
 ATATGCCCTTCAGGAACAATAGAAATCTTCGTGCGGTGTTACGTTGAAGTG
 GAGCGGATTATGTCAGCAATGGACAGAACCACTAATGAACACAGAACCA
 TGATGTGGTCTGTCCCTTTTACAGCCAGTAGTGCTCGCCGCAGTCGAGCGA
 CAGGGCGAAGCCCTCGAGTGAGCGAGGAAGCACCAGGGAACAGCATTAT
 ATATTCTGCTTACACACGATGCCTGAAAAAACTTCCCTTGGGGTTATCCA
 CTTATCCACGGGGATATTTTTTATAATTATTTTTTTTATAGTTTTTAGATC
 TTCTTTTTTTAGAGCGCCTTGTAGGCCTTTATCCATGCTGGTTCAGAGAA
 GGTGTTGTGACAAATTGCCCTTTTCACTGTGACAAATCACCTCAAATGAC
 AGTCCGTGTCTGTGACAAATTGCCCTTAACCCCTGTGACAAATTGCCCTCAG
 AAGAAGCTGTTTTTTCACAAAGTTATCCCTGCTTATTGACTCTTTTTTAT
 TTAGTGTGACAATCTAAAACTTGTACACTTCACATGGATCTGTCATGG
 CGGAAACAGCGGTTATCAATCACAAGAAACGTAATAATAGCCCGCAATC
 GTCCAGTCAAACGACCTCACTGAGGCGGCATATAGTCTCTCCCGGGATCA
 AAAACGTATGCTGTATCTGTTGCTTGACCAGATCAGAAAATCTGATGGCA
 CCCTACAGGAACATGACGGTATCTGCGAGATCCATGTTGCTAAATATGCT
 GAAATATTCGGATTGACCTCTGCGGAAGCCAGTAAGGATATACGGCAGGC
 ATTGAAGAGTTTTCGCGGGGAAGGAAGTGGTTTTTTATCGCCCTGAAGAGG
 ATGCCGCGCATGAAAAAGGCTATGAATCTTTTCCCTTGGTTTTATCAAACGT
 GCGCACAGTCCATCCAGAGGGCTTTACAGTGTACATATCAACCCATATCT
 CATTCCCTTCTTTATCGGGTTACAGAACCGGTTTACGCAGTTTCGGCTTA
 GTGAAACAAAAGAAATCACCAATCCGTATGCCATGCGTTTATACGAATCC
 CTGTGTCAGTATCGTAAGCCGGATGGCTCAGGCATCGTGTCTCTGAAAAAT
 CGACTGGATCATAGAGCGTTACCAGCTGCCTCAAAGTTACCAGCGTATGC
 CTGACTTCCGCCGCCGCTTCCCTGCAGGTCTGTGTTAATGAGATCAACAGC
 AGAACTCCAATGCGCCTCTCATACTTGAAGAAAAGAAAGGCCGCCAGAC
 GACTCATATCGTATTTTCTTCCGCGATATCACTTCCATGACGACAGGAT
 AGTCTGAGGGTTATCTGTACAGATTTGAGGGTGGTTCGTACATTTGTT
 CTGACCTACTGAGGGTAATTTGTACAGTTTTGTCTGTTTCCCTTCAGCCTG
 CATGGATTTTCTCATACTTTTGAAGTGAATTTTAAAGGAAGCCAAATT
 TGAGGGCAGTTTGTACAGTTGATTTCTTCTCTTCCCTTCGTATGTG
 ACCGTATATCGGGGGTTAGTTTCGTATCATGATGAGGGTTGATTATCAC
 AGTTTATTACTCTGAATTGGCTATCCGCGTGTGTACCTCTACCTGGAGTT
 TTTCCACGGTGGATATTTCTTCTTGCCTGAGCGTAAGAGCTATCTGAC
 AGAACAGTTCTTCTTTGCTTCCCTCGCCAGTTTCGCTCGCTATGCTCGGTTA
 CACGGCTGCGGCGAGCGCTAGTGATAATAAGTGAAGTATGTGGTGT
 TCTTATCTCCTTTTGTAGTGTGCTCTTATTTTAAACAACCTTTCGGGTTT
 TTTGATGACTTTGCGATTTTGTGTTGCTTTGCAGTAAATTGCAAGATTT
 AATAAAAAAACGCAAGCAATGATTAAAGGATGTTTCAAGATGAACTCAT
 GGAAACACTTAACAGTGCATAAACGCTGGTCATGAAATGACGAAGGCTA
 TCGCCATTGCACAGTTTAAATGATGACAGCCCGGAAGCGAGGAAAATAAC
 CGGCGCTGGAGAATAGGTGAAGCAGCGGATTTAGTTGGGGTTTCTTCTCA
 GGCTATCAGAGATGCCGAGAAAGCAGGGCGACTACCGCACCCGGATATGG
 AAATTCGAGGACGGGTTGAGCAACGTGTTGGTTATACAATTGAACAAATT
 AATCATATGCGTGATGTGTTTGGTACCGGATTGCGACGTGCTGAGGACGT
 ATTTCCACCGGTGATCGGGGTTGCTGCCCATAAAGGTGGCGTTTACAAAA
 CCTCAGTTTCTGTTTCTGCTCAGGATCTGGCTCTGAAGGGGCTACGT
 GTTTTGTCTGTGGAAGGTAACGACCCCGAGGGAACAGCCTCAATGTATCA
 CGGATGGGTACCAGATCTTCATATTCATGCAGAAGATACTCTCCTGCCTT

TCTATCTTGGGGAAAAGGACGATGTCACTTATGCAATAAAGCCCACTTGC
TGGCCGGGGCTTGACATTATTCCTTCCTGTCTGGCTCTGCACCGTATTGA
AACTGAGTTAATGGGCAAATTTGATGAAGGTAAACTGCCACCGATCCAC
ACCTGATGCTCCGACTGGCCATTGAAACTGTTGCTCATGACTATGATGTC
ATAGTTATTGACAGCGCGCCTAACCTGGGTATCGGCACGATTAATGTCGT
ATGTGCTGCTGATGTGCTGATTGTTCCCACGCCTGCTGAGTTGTTTGA
ACACCTCCGCACTGCAGTTTTTCGATATGCTTCGTGATCTGCTCAAGA
GTTGATCTTAAAGGGTTCGAGCCTGATGTACGTATTTGCTTACCAAATA
CAGCAATAGTAATGGCTCTCAGTCCCCGTGGATGGAGGAGCAAATTCGGG
ATGCCTGGGGAAGCATGGTTCTAAAAAATGTTGTACGTGAAACGGATGAA
GTTGGTAAAGGTCAGATCCGGATGAGAACTGTTTTTGAACAGGCCATTGA
TCAACGCAGTTCAACTGGTGCTGGAGAAATGCTCTTTCTATTTGGGAAC
CTGTCTGCAATGAAATTTTCGATCGTCTGATTAAACCACGCTGGGAGATT
AGATAATGAAGCGTGCGCCTGTTATTTCCAAAACATACGCTCAATACTCAA
CCGGTTGAAGATACTTCGTTATCGACACCAGCTGCCCCGATGGTGGATTCT
GTTAATTGCGCGCGTAGGAGTAATGGCTCGCGGTAATGCCATTACTTTGC
CTGTATGTGGTTCGGGATGTGAAGTTTACTCTTGAAGTGCTCCGGGGAT
AGTGTTGAGAAAACCTCTCGGGTATGGTCAGGTAATGAACGTGACCAGGA
GCTGCTTACTGAGGACGCACTGGATGATCTCATCCCTTCTTTTCTACTGA
CTGGTCAACAGACACCGGCGTTTCGGTCAAGAGTATCTGGTGTCTATAGAA
ATTGCCGATGGGAGTCGCCGTCGTAAAGCTGCTGCACTTACCGAAAGTGA
TTATCGTGTCTGGTTGGCGAGCTGGATGATGAGCAGATGGCTGCATTAT
CCAGATTGGGTAACGATTATCGCCCAACAAGTGCTTATGAACGTGGTCAG
CGTTATGCAAGCCGATTGCAGAATGAATTTGCTGGAAATATTTCTGCGCT
GGCTGATGCGGAAAATATTTACGTAAGATTATTACCCGCTGTATCAACA
CCGCCAAATTGCCTAAATCAGTTGTTGCTCTTTTTTCTCACCCTGGTGAA
CTATCTGCCCCGGTCAGGTGATGCACTTCAAAAAGCCTTTACAGATAAAGA
GGAATTACTTAAGCAGCAGGCATCTAACCTTCATGAGCAGAAAAAAGCTG
GGGTGATATTTGAAGCTGAAGAAGTTATCACTCTTTTAACTTCTGTGCTT
AAAACGTCATCTGCATCAAGAACTAGTTTAAAGCTCACGACATCAGTTTGC
TCCTGGAGCGACAGTATTGTATAAGGGCGATAAAATGGTGCTTAACCTGG
ACAGGTCACGTGTTCCAACTGAGTGTATAGAGAAAAATTGAGGCCATTCTT
AAGGAACTTGAAAAGCCAGCACCCCTGATGCGACCACGTTTTAGTCTACGT
TTATCTGTCTTTACTTAATGTCCTTTGTTACAGGCCAGAAAGCATAACTG
GCCTGAATATTTCTCTCTGGGCCCCACTGTTCCACTTGTATCGTCGGTCTGA
TAATCAGACTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAG
TCTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGA
CCACGGTCCCACTCGTATCGTCGGTCTGATAATCAGACTGGGACCACGGT
CCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCATGGTCCCACTC
GTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGT
CGGTCTGATTATTAGTCTGGAACCACGGTCCCACTCGTATCGTCGGTCTG
ATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTA
GTCTGGGACCACGATCCCACTCGTGTTGTGCGGTCTGATTATCGGTCTGGG
ACCACGGTCCCACTTGTATTGTGCGATCAGACTATCAGCGTGAGACTACGA
TTCCATCAATGCCTGTCAAGGGCAAGTATTGACATGTCGTGTAACCTGT
AGAACGGAGTAACCTCGGTGTGCGGTTGTATGCCTGCTGTGGATTGCTGC
TGTGTCTGCTTATCCACAACATTTTGGCAGCGGTTATGTGGACAAAATA
CCTGGTTACCCAGGCCGTGCCGGCACGTTAACCGGGCTGCATCCGATGCA
AGTGTGTCGCTGTGACGCAAGCGGCCGCA

>pMXBG_F7E (SEQ ID NO: 46)

CACTCTGT GACTTGGTCTCA TGTG TGAGACC GTCAGACCAAGTTTACT
CATATATACTTTAGATTGATTTAAACTTCATTTTTTAATTTAAAAGGATC
TAGGTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGA
GTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTT
CTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAA
CCACCGCTACCAGCGGTGGTTTGGTTGCCGGATCAAGAGCTACCAACTCT
TTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCC
TTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCG
CCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGG
CGATAAGTCGTGTCTTACC GGTTGGACTCAAGACGATAGTTACCGGATA
AGGCGCAGCGGTGCGGCTGAACGGGGGGTTTCGTGCACACAGCCCAGCTTG
GAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGA
AAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCG
GCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCC
TGGTATCTTTATAGTCCTGTGCGGTTTTCGCCACCTCTGACTTGAGCGTCG
ATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCA
ACGCGGCCCTTTTTACGGTTCTTGGCCTTTTGCTGGCCTTTTGCTCACATG
TTCTTTTCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTT
TGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGT
CAGTGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAACCGCCTCTCCCC
GCGCGTTGGCCGATTCAATTAATGCAGCTGGCAGCAGAGTTTCCCGACTG
GAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATT
AGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGA
ATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTA
CGCCAAAGCTTGCAATGCCTGCAGGTGCACTCTAGAGGATCCCCGGGTACCG
AGCTCGAATTCAGTGGCCGTGTTTTACAACGTCGTGACTGGGAAAACCC
TGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCT
GGCGTAATAGCGAAGAGGGCCCGCACCAGATCGCCCTTCCCAACAGTTGCGC
AGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCT
GTGCGGTATTTACACCGCATATGGTGCAGTCTCAGTACAATCTGCTCTG
ATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCG
CCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC
GGTCTCA CAGA TGAGACCAAGTC ACGAAGTGTGCGGCCGCCTTCAGCG
GCCGCACGAATCCATGTGGGAGTTTATTTCTTGACACAGATATTTATGATA
TAATAACTGAGTAAGCTTAACATAAGGAGGAAAAACATATGTTACGCAGC
AGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGGTGG
CTCAAGTATGGGCATCATTTCGCACATGTAGGCTCGGCCCTGACCAAGTCA
AATCCATGCGGGCTGCTCTTGATCTTTTCGGTCGTGAGTTTCGGAGACGTA
GCCACCTACTCCCAACATCAGCCGGAAGTCCGATTACCTCGGGAACTTGCT
CCGTAGTAAGACATTCATCGCGCTTGCTGCCTTCGACCAAGAAGCGGTTG
TTGGCGCTCTCGCGGCTTACGTTCTGCCCAAGTTTGAGCAGCCGCGTAGT
GAGATCTATATCTATGATCTCGCAGTCTCCGGCGAGCACCGGAGGCAGGG
CATTGCCACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAACGCGCTTG
GTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACGATCCCGCAGTG
GCTCTCTATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGATAT
CGACCCAAGTACCGCCACCTAAGAGCTGGAGCGCGTAACCGTCGCACAG
GAAGGACAGAGAAAGCGCGGATCTGGGAAGTGACGGACAGAACGGTCAGG
ACCTGGATTGGGGAGGCGGTTGCCGCCGCTGCTGCTGACGGTGTGACGTT
CTCTGTTCCGGTCACACCACATACGTTCCGCCATTCTATGCGATGCACA
TGCTGTATGCCGGTATACCGCTGAAAGTTCTGCAAAGCCTGATGGGACAT
AAGTCCATCAGTTCAACGGAAGTCTACACGAAGGTTTTTGCGCTGGATGT
GGCTGCCCGGCTCCGGGTGCAGTTTGCGATGCCGGAGTCTGATGCGGTTG

CGATGCTGAAACAATTATCCTGAGAATAAATGCCTTGGCCTTTATATGGA
 AATGTGGAAGTGGATATGCTGTTTTTGTCTGTTAAACAGAGAAGCT
 GGCTGTTATCCACTGAGAAGCGAACGAAACAGTCGGGAAAATCTCCCATT
 ATCGTAGAGATCCGCATTATTAATCTCAGGAGCCTGTGTAGCGTTTATAG
 GAAGTAGTGTCTGTTCATGATGCCTGCAAGCGGTAACGAAAACGATTTGA
 ATATGCCCTTCAGGAACAATAGAAATCTTCGTGCGGTGTTACGTTGAAGTG
 GAGCGGATTATGTCAGCAATGGACAGAACCACTAATGAACACAGAACCA
 TGATGTGGTCTGTCCCTTTTACAGCCAGTAGTGCTCGCCGCAGTCGAGCGA
 CAGGGCGAAGCCCTCGAGTGAGCGAGGAAGCACCAGGGAACAGCATTAT
 ATATTCTGCTTACACACGATGCCTGAAAAAACTTCCCTTGGGGTTATCCA
 CTTATCCACGGGGATATTTTTTATAATTATTTTTTTTATAGTTTTTAGATC
 TTCTTTTTTTAGAGCGCCTTGTAGGCCTTTATCCATGCTGGTTCAGAGAA
 GGTGTTGTGACAAATTGCCCTTTTCACTGTGACAAATCACCTCAAATGAC
 AGTCTGTCTGTGACAAATTGCCCTTAACCCCTGTGACAAATTGCCCTCAG
 AAGAAGCTGTTTTTTCACAAAGTTATCCCTGCTTATTGACTCTTTTTTAT
 TTAGTGTGACAATCTAAAACTTGTACACTTCACATGGATCTGTCATGG
 CGGAAACAGCGGTTATCAATCACAAGAAACGTAATAATAGCCCGCAATC
 GTCCAGTCAAACGACCTCACTGAGGCGGCATATAGTCTCTCCCGGGATCA
 AAAACGTATGCTGTATCTGTTGCTTGACCAGATCAGAAAATCTGATGGCA
 CCCTACAGGAACATGACGGTATCTGCGAGATCCATGTTGCTAAATATGCT
 GAAATATTCGGATTGACCTCTGCGGAAGCCAGTAAGGATATACGGCAGGC
 ATTGAAGAGTTTTCGCGGGGAAGGAAGTGGTTTTTTATCGCCCTGAAGAGG
 ATGCCGGCGATGAAAAAGGCTATGAATCTTTTCCCTTGGTTTTATCAAACGT
 GCGCACAGTCCATCCAGAGGGCTTTACAGTGTACATATCAACCCATATCT
 CATTCCCTTCTTTATCGGGTTACAGAACCGGTTTACGCAGTTTCGGCTTA
 GTGAAACAAAAGAAATCACCAATCCGTATGCCATGCGTTTATACGAATCC
 CTGTGTCAGTATCGTAAGCCGGATGGCTCAGGCATCGTGTCTCTGAAAAAT
 CGACTGGATCATAGAGCGTTACCAGCTGCCTCAAAGTTACCAGCGTATGC
 CTGACTTCCGCCGCCGCTTCCCTGCAGGTCTGTGTTAATGAGATCAACAGC
 AGAACTCCAATGCGCCTCTCATAACATTGAGAAAAAGAAAGGCCGCCAGAC
 GACTCATATCGTATTTTCTTCCGCGATATCACTTCCATGACGACAGGAT
 AGTCTGAGGGTTATCTGTACAGATTTGAGGGTGGTTCGTACATTTGTT
 CTGACCTACTGAGGGTAATTTGTACAGTTTTGCTGTTTCCCTTCAGCCTG
 CATGGATTTTCTCATACTTTTGAAGTGAATTTTAAAGGAAGCCAAATT
 TGAGGGCAGTTTGTACAGTTGATTTCTTCTCTTCCCTTCGTCATGTG
 ACCGTATATCGGGGGTTAGTTTCGTACATCATTGATGAGGGTTGATTATCA
 AGTTTATTACTCTGAATTGGCTATCCGCGTGTGTACCTCTACCTGGAGTT
 TTTCCACGGTGATATTTCTTCTTGCCTGAGCGTAAGAGCTATCTGAC
 AGAACAGTTCTTCTTTGCTTCCCTCGCCAGTTTCGCTCGCTATGCTCGGTTA
 CACGGCTGCGGCGAGCGCTAGTGATAATAAGTGAAGTATGAGGTATGTGGTGT
 TCTTATCTCCTTTTGTAGTGTTGCTCTTATTTTAAACAACTTTGCGGTTT
 TTTGATGACTTTGCGATTTTGTGTTGCTTTGCAGTAAATTGCAAGATTT
 AATAAAAAAACGCAAGCAATGATTAAAGGATGTTTCAAGATGAACTCAT
 GGAAACACTTAACAGTGCATAAACGCTGGTCATGAAATGACGAAGGCTA
 TCGCCATTGCACAGTTTAAATGATGACAGCCCGAAGCGAGGAAAATAAC
 CGGCGCTGGAGAATAGGTGAAGCAGCGGATTTAGTTGGGGTTTCTTCTCA
 GGCTATCAGAGATGCCGAGAAAGCAGGGCGACTACCGCACCCGGATATGG
 AAATTCGAGGACGGGTTGAGCAACGTGTTGGTTATACAATTGAACAAATT
 AATCATATGCGTGATGTGTTTGGTACCGGATTGCGACGTGCTGAGGACGT
 ATTTCCACCGGTGATCGGGGTTGCTGCCCATAAAGGTGGCGTTTACAAAA
 CCTCAGTTTCTGTTTCTGCTCAGGATCTGGCTCTGAAGGGGCTACGT
 GTTTTGTCTGTGGAAGGTAACGACCCCGAGGGAACAGCCTCAATGTATCA
 CGGATGGGTACCAGATCTTCATATTCATGCAGAAGATACTCTCCTGCCTT

TCTATCTTGGGGAAAAGGACGATGTCACTTATGCAATAAAGCCCACTTGC
TGGCCGGGGCTTGACATTATTCCTTCCTGTCTGGCTCTGCACCGTATTGA
AACTGAGTTAATGGGCAAATTTGATGAAGGTAAACTGCCACCGATCCAC
ACCTGATGCTCCGACTGGCCATTGAAACTGTTGCTCATGACTATGATGTC
ATAGTTATTGACAGCGCGCCTAACCTGGGTATCGGCACGATTAATGTCGT
ATGTGCTGCTGATGTGCTGATTGTTCCCACGCCTGCTGAGTTGTTTGA
ACACCTCCGCACTGCAGTTTTTCGATATGCTTCGTGATCTGCTCAAGA
GTTGATCTTAAAGGGTTCGAGCCTGATGTACGTATTTGCTTACCAAATA
CAGCAATAGTAATGGCTCTCAGTCCCCGTGGATGGAGGAGCAAATTCGGG
ATGCCTGGGGAAGCATGGTTCTAAAAAATGTTGTACGTGAAACGGATGAA
GTTGGTAAAGGTCAGATCCGGATGAGAACTGTTTTTGAACAGGCCATTGA
TCAACGCAGTTCAACTGGTGCCTGGAGAAATGCTCTTTCTATTTGGGAAC
CTGTCTGCAATGAAATTTTCGATCGTCTGATTAAACCACGCTGGGAGATT
AGATAATGAAGCGTGCGCCTGTTATTTCCAAAACATACGCTCAATACTCAA
CCGGTTGAAGATACTTCGTTATCGACACCAGCTGCCCCGATGGTGGATTCT
GTTAATTGCGCGCGTAGGAGTAATGGCTCGCGGTAATGCCATTACTTTGC
CTGTATGTGGTTCGGGATGTGAAGTTTACTCTTGAAGTGCTCCGGGGAT
AGTGTTGAGAAAACCTCTCGGGTATGGTCAGGTAATGAACGTGACCAGGA
GCTGCTTACTGAGGACGCACTGGATGATCTCATCCCTTCTTTTCTACTGA
CTGGTCAACAGACACCGGCGTTTCGGTCAAGAGTATCTGGTGTCTATAGAA
ATTGCCGATGGGAGTCGCCGTCGTAAAGCTGCTGCACTTACCGAAAGTGA
TTATCGTGTCTGGTTGGCGAGCTGGATGATGAGCAGATGGCTGCATTAT
CCAGATTGGGTAACGATTATCGCCCAACAAGTGCTTATGAACGTGGTCAG
CGTTATGCAAGCCGATTGCAGAATGAATTTGCTGGAAATATTTCTGCGCT
GGCTGATGCGGAAAATATTTACGTAAGATTATTACCCGCTGTATCAACA
CCGCCAAATTGCCTAAATCAGTTGTTGCTCTTTTTTCTCACCCTGGTGAA
CTATCTGCCCCGTCAGGTGATGCACTTCAAAAAGCCTTTACAGATAAAGA
GGAATTACTTAAGCAGCAGGCATCTAACCTTCATGAGCAGAAAAAAGCTG
GGGTGATATTTGAAGCTGAAGAAGTTATCACTCTTTTAACTTCTGTGCTT
AAAACGTCATCTGCATCAAGAACTAGTTTAAAGCTCACGACATCAGTTTGC
TCCTGGAGCGACAGTATTGTATAAGGGCGATAAAATGGTGCTTAACCTGG
ACAGGTCACGTGTTCCAAGTGTATAGAGAAAAATTGAGGCCATTCTT
AAGGAACTTGAAAAGCCAGCACCTGATGCGACCACGTTTTAGTCTACGT
TTATCTGTCTTTACTTAATGTCCTTTGTTACAGGCCAGAAAGCATAACTG
GCCTGAATATTTCTCTCTGGGCCCCACTGTTCCACTTGATCGTCGGTCTGA
TAATCAGACTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAG
TCTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGA
CCACGGTCCCACTCGTATCGTCGGTCTGATAATCAGACTGGGACCACGGT
CCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCATGGTCCCACTC
GTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGT
CGGTCTGATTATTAGTCTGGAACCACGGTCCCACTCGTATCGTCGGTCTG
ATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTA
GTCTGGGACCACGATCCCACTCGTGTTGTGCGGTCTGATTATCGGTCTGGG
ACCACGGTCCCACTTGTATTGTGCGATCAGACTATCAGCGTGAGACTACGA
TTCCATCAATGCCTGTCAAGGGCAAGTATTGACATGTCGTGTAACCTGT
AGAACGGAGTAACCTCGGTGTGCGGTTGTATGCCTGCTGTGGATTGCTGC
TGTTGCTGCTTATCCACAACATTTTGGGCACGGTTATGTGGACAAAATA
CCTGGTTACCCAGGCCGTGCCGGCACGTTAACCGGGCTGCATCCGATGCA
AGTGTGTCGCTGTGACGCAAGCGGCCGCA

>pMXBK A7E (SEQ ID NO: 47)

CACTCTGT GACTTGGTCTCA TGCC TGAGACC GTCAGACCAAGTTTACT
 CATATATACTTTAGATTGATTTAAACTTTCATTTTTTAATTTAAAAGGATC
 TAGGTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGA
 GTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTT
 CTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAA
 CCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCAACTCT
 TTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCC
 TTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCG
 CCTACATAACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGG
 CGATAAGTCGTGTCTTACCGGTTGGACTCAAGACGATAGTTACCGGATA
 AGGCGCAGCGGTGCGGCTGAACGGGGGTTTCGTGCACACAGCCAGCTTG
 GAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGA
 AAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCG
 GCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCC
 TGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCG
 ATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCA
 ACGCGGCCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATG
 TTCTTTCCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTT
 TGAGTGAGCTGATACCGCTCGCCGACGCCGAACGACCGAGCGCAGCGAGT
 CAGTGAGCGAGGAAGCGGATGAGCGCCCAATACGCAAAACCGCTCTCCCC
 GCGCGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTTCCCGACTG
 GAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATT
 AGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGA
 ATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTA
 CGCCAAGCTTGATGCCTGCAGGTGCACTCTAGAGGATCCCCGGGTACCG
 AGCTCGAATTCAGTGGCCGTGTTTTACAACGTCGTGACTGGGAAAACCC
 TGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCT
 GGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGC
 AGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCT
 GTGCGGTATTTACACCCGCATATGGTGCACCTCTCAGTACAATCTGCTCTG
 ATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACCGG
 CCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC
GGTCTCA CAGA TGAGACCAAGTC ACGAAGTGTGCGGCCGCCTTCGTGC
 GCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCG
 CTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAA
 GAGTATGGCTAAAATGAGAATATCACCGGAATTGAAAAAACTGATCGAAA
 AATACCGCTGCGTAAAAGATACGGAAGGAATGTCTCCTGCTAAGGTATAT
 AAGCTGGTGGGAGAAAATGAAAACCTATATTTAAAAATGACGGACAGCCG
 GTATAAAGGGACCACCTATGATGTGGAACGGGAAAAGGACATGATGCTAT
 GGCTGGAAGGAAAGCTGCCTGTTCCAAAGGTCTGCACTTTGAACGGCAT
 GATGGCTGGAGCAATCTGCTCATGAGTGAGGCCGATGGCGTCTTTGCTC
 GGAAGAGTATGAAGATGAACAAAGCCCTGAAAAGATTATCGAGCTGTATG
 CGGAGTGCATCAGGCTCTTTCACTCCATCGACATATCGGATTGTCCCTAT
 ACGAATAGCTTAGACAGCCGCTTAGCCGAATTGGATTACTTACTGAATAA
 CGATCTGGCCGATGTGGATTGCGAAAACCTGGGAAGAGGACACTCCATTTA
 AAGATCCGCGCGAGCTGTATGATTTTTTAAAGACGGAAAAGCCGAAGAG
 GAACTTGTCTTTTCCACGGCGACCTGGGAGACAGCAACATCTTTGTGAA
 AGATGGCAAAGTAAGTGGCTTTATTGATCTTGGGAGAAGCGGCAGGGCGG
 ACAAGTGGTATGACATTGCCTTCTGCGTCCGGTCGCTCAGGGAGGATATC
 GGGGAAGAACAGTATGTCGAGCTATTTTTTGACTTACTGGGGATCAAGCC
 TGATTGGGAGAAAATAAATATTATATTTTACTGGATGAATTGTTTTAGG
 AGCTGGAGCGGCGTAACCGTCGCACAGGAAGGACAGAGAAAAGCGCGGATC

TGGGAAGTGACGGACAGAACGGTCAGGACCTGGATTGGGGAGGCGGTTGC
 CGCCGCTGCTGCTGACGGTGTGACGTTCTCTGTTCCGGTCACACCACATA
 CGTTCGCGCCATTCCCTATGCGATGCACATGCTGTATGCCGGTATACCGCTG
 AAAGTTCCTGCAAAGCCTGATGGGACATAAGTCCATCAGTTCAACGGAAGT
 CTACACGAAGGTTTTTGGCGCTGGATGTGGCTGCCCGGCTCCGGGTGCAGT
 TTGCGATGCCGGAGTCTGATGCGGTTGCGATGCTGAAACAATTATCCTGA
 GAATAAATGCCCTTGGCCTTTATATGGAATGTGGAAGTGAAGTGGATATGC
 TGTTTTTGTCTGTTAAACAGAGAAGCTGGCTGTTATCCACTGAGAAGCGA
 ACGAAACAGTCGGGAAAATCTCCCATATCGTAGAGATCCGCATTATTAA
 TCTCAGGAGCCTGTGTAGCGTTTATAGGAAGTAGTGTTCTGTCTATGATGC
 CTGCAAGCGGTAACGAAAACGATTTGAATATGCCCTCAGGAACAATAGAA
 ATCTTCGTGCGGTGTTACGTTGAAGTGGAGCGGATTATGTCAGCAATGGA
 CAGAACAACCTAATGAACACAGAACCATGATGTGGTCTGTCCTTTTACAG
 CCAGTAGTGCTCGCCGCAGTCGAGCGACAGGGCGAAGCCCTCGAGTGAGC
 GAGGAAGCACCAGGGAACAGCACTTATATATTCTGCTTACACACGATGCC
 TGAAAAAACTTCCCTTGGGGTTATCCACTTATCCACGGGGATATTTTTAT
 AATTATTTTTTTTTATAGTTTTTTAGATCTTCTTTTTTAGAGCGCCTTGTA
 GCCTTTATCCATGCTGGTCTAGAGAAGGTGTTGTGACAAATTGCCCTTT
 CAGTGTGACAAATCACCTCAAATGACAGTCCTGTCTGTGACAAATTGCC
 CTTAACCCTGTGACAAATTGCCCTCAGAAGAAGCTGTTTTTTTACAAAGT
 TATCCCTGCTTATTGACTCTTTTTTATTTAGTGAGACAATCTAAAACTT
 GTCACACTTCACATGGATCTGTCTATGGCGGAAACAGCGGTTATCAATCAC
 AAGAAACGTAAAAATAGCCCGCGAATCGTCCAGTCAAACGACCTCACTGA
 GGCGGCATATAGTCTCTCCCGGGATCAAAAACGTATGCTGTATCTGTTTCG
 TTGACCAGATCAGAAAATCTGATGGCACCTTACAGGAACATGACGGTATC
 TGCGAGATCCATGTTGCTAAATATGCTGAAATATTCGGATTGACCTCTGC
 GGAAGCCAGTAAGGATATACGGCAGGCATTGAAGAGTTTCGCGGGGAAGG
 AAGTGGTTTTTTTATCGCCCTGAAGAGGATGCCGGCGATGAAAAAGGCTAT
 GAATCTTTTTCCTTGGTTTTATCAAACGTGCGCACAGTCCATCCAGAGGGCT
 TTACAGTGATACATATCAACCCATATCTCATTCCCTTCTTTATCGGGTTAC
 AGAACCGGTTTACGCAGTTTTCGGCTTAGTGAAACAAAAGAAATCACCAAT
 CCGTATGCCATGCGTTTATACGAATCCCTGTGTGAGTATCGTAAGCCGGA
 TGGCTCAGGCATCGTGTCTCTGAAAATCGACTGGATCATAGAGCGTTACC
 AGCTGCCCTCAAAGTTACCAGCGTATGCCTGACTTCCGCCGCCGCTTCCTG
 CAGGTCTGTGTTAATGAGATCAACAGCAGAACTCCAATGCGCCTCTCATA
 CATTTAGAAAAAGAAAGGCCGCCAGACGACTCATATCGTATTTTTCTTCC
 GCGATATCACTTCCATGACGACAGGATAGTCTGAGGGTTATCTGTACAG
 ATTTGAGGGTGGTTCGTACATTTGTTCTGACCTACTGAGGGTAATTTGT
 CACAGTTTTGCTGTTTTCTTTCAGCCTGCATGGATTTTCTCATACTTTTTG
 AACTGTAATTTTTAAGGAAGCCAAATTTGAGGGCAGTTTGTACAGTTGA
 TTTCTTCTCTTTCCCTTCGTCTATGTGACCTGATATCGGGGGTTAGTTTCG
 TCATCATTTGATGAGGGTTGATTATCACAGTTTATTAATCTGAATTGGCTA
 TCCGCGTGTGTACCTCTACCTGGAGTTTTTCCACGGTGGATATTTCTTC
 TTGCGCTGAGCGTAAGAGCTATCTGACAGAACAGTTCTTCTTTGCTTCCT
 CGCCAGTTCGCTCGCTATGCTCGGTTACACGGCTGCGGCGAGCGCTAGTG
 ATAATAAGTGACTGAGGTATGTGGTGTCTTATCTCCTTTTGTAGTGTG
 CTCTTATTTTAAACAACCTTTGCGGTTTTTTGATGACTTTGCGATTTTGT
 GTTGCTTTGTCAGTAAATTGCAAGATTTAATAAAAAAACGCAAAGCAATGA
 TTAAAGGATGTTTCAAGTGAAGTCAATGGAACACTTAACAGTGCATAA
 ACGCTGGTCATGAAATGACGAAGGCTATCGCCATTGCACAGTTTAAATGAT
 GACAGCCCGGAAGCGAGGAAAATAACCCGGCGCTGGAGAATAGGTGAAGC
 AGCGGATTTAGTTGGGGTTTCTTCTCAGGCTATCAGAGATGCCGAGAAAG
 CAGGGCGACTACCGCACCCGGATATGGAATTCGAGGACGGGTTGAGCAA

CGTGTGTTGGTTATACAATTGAACAAATTAATCATATGCGTGATGTGTTTGG
 TACGCGATTGCGACGTGCTGAGGACGTATTTCCACCGGTGATCGGGGTTG
 CTGCCCATAAAGGTGGCGTTTACAAAACCTCAGTTTCTGTTTCATCTTGCT
 CAGGATCTGGCTCTGAAGGGGCTACGTGTTTTGCTCGTGGAAGGTAACGA
 CCCCCAGGGAACAGCCTCAATGTATCACGGATGGGTACCAGATCTTCATA
 TTCATGCAGAAGATACTCTCCTGCCTTTCTATCTTGGGGAAAAGGACGAT
 GTCACTTATGCAATAAAGCCCACTTGCTGGCCGGGGCTTGACATTATTCC
 TTCCTGTCTGGCTCTGCACCGTATTGAAACTGAGTTAATGGGCAAATTTG
 ATGAAGGTAAACTGCCCACCGATCCACACCTGATGCTCCGACTGGCCATT
 GAAACTGTTGCTCATGACTATGATGTCATAGTTATTGACAGCGCGCCTAA
 CCTGGGTATCGGCACGATTAATGTCGTATGTGCTGCTGATGTGCTGATTG
 TTCCCACGCCTGCTGAGTTGTTTGACTACACCTCCGCACTGCAGTTTTTC
 GATATGCTTCGTGATCTGCTCAAGAACGTTGATCTTAAAGGGTTCGAGCC
 TGATGTACGTATTTTGCTTACCAAATACAGCAATAGTAATGGCTCTCAGT
 CCCCCTGGATGGAGGAGCAAATTCGGGATGCCTGGGGAAAGCATGGTTCTA
 AAAAAATGTTGTACGTGAAACGGATGAAGTTGGTAAAGGTCAGATCCGGAT
 GAGAATGTTTTTGAACAGGCCATTGATCAACGCAGTTCAACTGGTGCCCT
 GGAGAAATGCTCTTTCTATTTGGGAACCTGTCTGCAATGAAATTTTCGAT
 CGTCTGATTAAACCACGCTGGGAGATTAGATAATGAAGCGTGCGCCTGTT
 ATTCCAAAACATACGCTCAATACTCAACCGGTGAAGATACTTCGTTATC
 GACACCAGCTGCCCCGATGGTGGATTTCGTTAATTGCGCGCGTAGGAGTAA
 TGGCTCGCGGTAATGCCATTACTTTGCCTGTATGTGGTCGGGATGTGAAG
 TTTACTCTTGAAGTGCTCCGGGGTGATAGTGTTGAGAAAACCTCTCGGGT
 ATGGTCAGGTAATGAACGTGACCAGGAGCTGCTTACTGAGGACGCACTGG
 ATGATCTCATCCCTTCTTTTCTACTGACTGGTCAACAGACACCGGCGTTCT
 GGTGGAAGAGTATCTGGTGTATAGAAAATTGCCGATGGGAGTCGCGGTCTC
 TAAAGCTGCTGCACTTACCGAAAGTGATTATCGTGTTCTGGTTGGCGAGC
 TGGATGATGAGCAGATGGCTGCATTATCCAGATTGGGTAACGATTATCGC
 CCAACAAGTGCTTATGAACGTGGTCAGCGTTATGCAAGCCGATTGCAGAA
 TGAATTTGCTGGAAATATTTCTGCGCTGGCTGATGCGGAAAATATTTTAC
 GTAAGATTATTACCCGCTGTATCAACACCGCCAAATGCTTAAATCAGTT
 GTTGCTCTTTTTTCTCACCCCGGTGAACATCTGCCCCGTCAGGTGATGC
 ACTTCAAAAAGCCTTTACAGATAAAGAGGAATTACTTAAGCAGCAGGCAT
 CTAACCTTCATGAGCAGAAAAAAGCTGGGGTGATATTTGAAGCTGAAGAA
 GTTATCACTCTTTTAACTTCTGTGCTTAAAACGTCATCTGCATCAAGAAC
 TAGTTTTAAGCTCACGACATCAGTTTGCTCCTGGAGCGACAGTATTGTATA
 AGGGCGATAAAATGGTGCTTAACCTGGACAGGTCACGTGTTCCAAGTGAAG
 TGTATAGAGAAAATTGAGGCCATTCTTAAGGAACCTGAAAAGCCAGCACC
 CTGATGCGACCACGTTTTAGTCTACGTTTATCTGTCTTTACTTAAATGTCC
 TTTGTTACAGGCCAGAAAGCATAAAGTGCCTGAATATTCTCTCTGGGCCC
 ACTGTTCCACTTGTATCGTCGGTCTGATAATCAGACTGGGACCACGGTCC
 CACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACCTCGT
 ATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACCTCGTATCGTCG
 GTCTGATAATCAGACTGGGACCACGGTCCCACCTCGTATCGTCGGTCTGAT
 TATTAGTCTGGGACCATGGTCCCACCTCGTATCGTCGGTCTGATTATTAGT
 CTGGGACCACGGTCCCACCTCGTATCGTCGGTCTGATTATTAGTCTGGAAC
 CACGGTCCCACCTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCT
 CCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGATCCCACCTCG
 TGTGTCGGTCTGATTATCGGTCTGGGACCACGGTCCCACCTTGTATTGTC
 GATCAGACTATCAGCGTGAGACTACGATTCCATCAATGCCTGTCAAGGGC
 AAGTATTGACATGTCGTGTAACCTGTAGAACGGAGTAACCTCGGTGTGC
 GGTGTATGCCTGCTGTGGATTGCTGCTGTGCTGCTTATCCACAACAT
 TTTGCGCACGGTTATGTGGACAAAATACCTGGTTACCCAGGCCGTGCCGG
 CACGTAAACCGGGCTGCATCCGATGCAAGTGTGTCGGTGTGCGACGCAAGC
 GGCCGCA

REIVINDICACIONES

1. Un ácido nucleico para su uso en el ensamblaje de ADN, en donde el ácido nucleico comprende al menos un elemento de restricción metilable, el elemento de restricción metilable comprende:

- una secuencia de reconocimiento de enzimas de restricción de tipo IIS; y

- una secuencia de reconocimiento de la metilasa del ADN,

en donde la secuencia de reconocimiento de la enzima de restricción de tipo IIS y la secuencia de reconocimiento de la ADN metilasa se solapan de tal forma que la base metilable por la ADN metilasa se encuentra dentro de la secuencia de reconocimiento de la enzima de restricción de tipo IIS,

en donde la secuencia de reconocimiento de la metilasa del ADN no es idéntica ni está rodeada por la secuencia de reconocimiento de la enzima de restricción tipo IIS, y

en donde la secuencia de reconocimiento de la metilasa del ADN no se solapa con la secuencia que formaría la secuencia del extremo sobresaliente generada por la enzima de restricción de tipo IIS,

en donde el ácido nucleico comprende además una secuencia opuesta no protegible de reconocimiento de enzima de restricción de tipo IIS proporcionada en el lado opuesto del sitio de corte del elemento de restricción protegible por metilación, formando así un elemento compuesto truncado,

en donde la secuencia opuesta de reconocimiento de la enzima de restricción tipo IIS del elemento compuesto truncador está dispuesta para dirigir a la enzima de restricción tipo IIS a cortar el ácido nucleico en un sitio que está dentro de la secuencia de reconocimiento de la enzima de restricción tipo IIS del elemento de restricción metilable-protégible, y/o dentro de los nucleótidos entre el sitio de corte y la secuencia de reconocimiento de la enzima de restricción tipo IIS del elemento de restricción metilable-protégible.

2. El ácido nucleico según la reivindicación 1, en donde el ácido nucleico comprende un primer y un segundo elemento de restricción protegibles por metilación,

y en donde el ácido nucleico comprende además secuencias de reconocimiento de enzimas de restricción de tipo IIS opuestas proporcionadas en el lado opuesto del sitio de corte de cada uno de los elementos de restricción protegibles contra metilación primero y segundo para formar:

A i) un elemento compuesto de mantenimiento o un elemento compuesto de inserción, y ii) un elemento compuesto truncado; o

B) dos elementos compuestos truncados,

en donde el elemento compuesto de mantenimiento es una disposición en donde la secuencia de reconocimiento de la enzima de restricción de tipo IIS opuesta está dispuesta para dirigir la enzima de restricción de tipo IIS para cortar el ácido nucleico en el mismo sitio que la secuencia de reconocimiento de la enzima de restricción de tipo IIS del elemento de restricción metilación-protégible opuesto, de tal manera que se produce el mismo voladizo,

en donde el elemento compuesto insercional es un arreglo que comprende un inserto de secuencia funcional provisto entre el elemento de restricción protegible contra metilación y la secuencia opuesta de reconocimiento de enzima de restricción tipo IIS, y

en donde el elemento compuesto truncador es una disposición en donde la secuencia de reconocimiento de la enzima de restricción de tipo IIS opuesta del elemento compuesto truncador está dispuesta para dirigir la enzima de restricción de tipo IIS para cortar el ácido nucleico en un sitio que está dentro de la secuencia de reconocimiento de la enzima de restricción de tipo IIS del elemento de restricción metilable-protégible, y/o dentro de los nucleótidos entre el sitio de corte y la secuencia de reconocimiento de la enzima de restricción de tipo IIS del elemento de restricción metilable-protégible.

3. El ácido nucleico según la reivindicación 2, en donde el ácido nucleico comprende una secuencia de ácido nucleico entre los sitios de corte de los dos elementos de restricción protegibles por metilación, que es una secuencia de descarte; o

en donde el ácido nucleico es un vector linealizado que tiene un elemento de restricción protegible contra la metilación en cada extremo.

4. El ácido nucleico según la reivindicación 3, en donde la secuencia de descarte comprende un marcador seleccionable, un gen informador o una etiqueta.

5. El ácido nucleico según cualquiera de las reivindicaciones precedentes, en donde la secuencia de reconocimiento de la enzima de restricción del elemento de restricción protegible contra la metilación es reconocida por la misma especie de enzima de restricción que la secuencia de reconocimiento de la enzima de restricción opuesta no protegible; y/o

5 en donde las secuencias de reconocimiento de metilasas de ADN de cada elemento de restricción protegible contra la metilación son reconocidas por la misma metilasa.

6. El ácido nucleico según cualquier reivindicación 1 anterior, en donde el elemento de restricción protegible contra la metilación comprende o consiste en la secuencia GACNNGGTCTCNNNNN (BsaI/M.Osp807II - SEQ ID NO: 1) o GAAGACGCNNNNN (BpiI/M2.NmeMC58II - SEQ ID NO: 2) o GAAGCTTCNNNN (LglI/M.XmnI - SEQ ID NO: 3).

10 7. Un método de ensamblaje de fragmentos de ADN sin cicatriz que comprende los pasos de:

(A) proporcionar un primer ácido nucleico metilado linealizado mediante corte con enzima de restricción de un primer ácido nucleico, en donde el primer ácido nucleico comprende primero y segundo elementos de restricción protegibles contra metilación con una secuencia de descarte entre ellos, o proporcionar una versión previamente cortada con enzima de restricción del mismo que tiene la secuencia de descarte extirpada,

15 en donde el primer y segundo elementos de restricción protegibles por metilación del primer ácido nucleico comprenden cada uno:

- una secuencia de reconocimiento de enzima de restricción de tipo IIS; y
- una secuencia de reconocimiento de metilasa del ADN,

20 en donde la secuencia de reconocimiento de la enzima de restricción de tipo IIS y la secuencia de reconocimiento de la metilasa del ADN se solapan de manera que la base metilable por la metilasa del ADN se encuentra dentro de la secuencia de reconocimiento de la enzima de restricción de tipo IIS,

en donde la secuencia de reconocimiento de la metilasa del ADN no es idéntica ni está rodeada por la secuencia de reconocimiento de la enzima de restricción tipo IIS, y

25 en donde la secuencia de reconocimiento de metilasa del ADN no se solapa con la secuencia que formaría la secuencia del extremo sobresaliente generada por la enzima de restricción de tipo IIS, y

30 en donde el primer y segundo elementos de restricción protegibles por metilación están metilados con la metilasa del ADN que reconoce la secuencia de reconocimiento de la metilasa del ADN; y en donde el primer ácido nucleico comprende además secuencias de reconocimiento de enzimas de restricción de tipo IIS no protegibles opuestas proporcionadas en el lado opuesto del sitio de corte de cada uno de los elementos de restricción protegibles por metilación primero y segundo para proporcionar i) un elemento compuesto de mantenimiento y ii) un elemento compuesto truncado,

35 en donde el elemento compuesto de mantenimiento es una disposición en donde la secuencia de reconocimiento de la enzima de restricción de tipo IIS opuesta está dispuesta para dirigir la enzima de restricción de tipo IIS para cortar el ácido nucleico en el mismo sitio que la secuencia de reconocimiento de la enzima de restricción de tipo IIS del elemento de restricción opuesto protegible por metilación, de modo que se produzca el mismo saliente, y

40 en donde el elemento compuesto de truncamiento es una disposición en donde la secuencia de reconocimiento de la enzima de restricción de tipo IIS opuesta del elemento compuesto de truncamiento está dispuesta para dirigir la enzima de restricción de tipo IIS para cortar el ácido nucleico en un sitio que está dentro de la secuencia de reconocimiento de la restricción de tipo IIS secuencia de reconocimiento enzimático del elemento de restricción protegible por metilación, o dentro de los nucleótidos entre el sitio de corte y la secuencia de reconocimiento del elemento de restricción protegible por metilación secuencia de reconocimiento de enzima del elemento de restricción protegible por metilación;

45 en donde el primer ácido nucleico es una enzima de restricción cortada por una enzima de restricción de tipo IIS que reconoce las secuencias de reconocimiento de la enzima de restricción de tipo IIS opuestas, formando así el primer ácido nucleico metilado linealizado, y

50 en donde los elementos compuestos de mantenimiento y truncamiento están metilados con la metilasa del ADN; proporcionar un primer fragmento de ADN para ensamblaje que tiene extremos salientes que están adaptados para ligarse a los extremos salientes del primer ácido nucleico metilado linealizado;

ligar el primer fragmento de ADN para el ensamblaje y el primer ácido nucleico metilado linealizado con una ligasa para formar un primer vector intermedio metilado;

transformar el primer vector intermedio metilado en una cepa bacteriana que no expresa la(s) metilasa del ADN(s) que reconoce cualquiera de las secuencias de reconocimiento de metilasa del ADN de los elementos compuestos de mantenimiento y truncamiento, formando de este modo un primer vector intermedio que no está metilado;

5 aislar el primer vector intermedio;

(B) proporcionar un segundo ácido nucleico metilado linealizado mediante corte con enzimas de restricción de un segundo ácido nucleico,

10 en donde el segundo ácido nucleico comprende un primer y un segundo elementos de restricción protegibles por metilación con una secuencia de descarte entre ellos, o proporcionando una versión previamente cortada con enzima de restricción del mismo que tiene la secuencia de descarte escindida,

en donde el primer y segundo elementos de restricción protegibles por metilación del segundo ácido nucleico comprenden cada uno:

- una secuencia de reconocimiento de enzima de restricción de tipo IIS; y

- una secuencia de reconocimiento de metilasa del ADN,

15 en donde la secuencia de reconocimiento de la enzima de restricción de tipo IIS y la secuencia de reconocimiento de la metilasa del ADN se solapan de manera que la base metilable por la metilasa del ADN se encuentra dentro de la secuencia de reconocimiento de la enzima de restricción de tipo IIS,

en donde la secuencia de reconocimiento de la metilasa del ADN no es idéntica ni está rodeada por la secuencia de reconocimiento de la enzima de restricción tipo IIS, y

20 en donde la secuencia de reconocimiento de metilasa del ADN no se solapa con la secuencia que formaría la secuencia del extremo sobresaliente generada por la enzima de restricción de tipo IIS, y

en donde el primer y segundo elementos de restricción protegibles por metilación están metilados con la metilasa del ADN que reconoce la secuencia de reconocimiento de la metilasa del ADN; y

25 en donde el segundo ácido nucleico comprende además secuencias de reconocimiento de enzimas de restricción de tipo IIS no protegibles opuestas proporcionadas en el lado opuesto del sitio de corte de cada uno de los elementos de restricción protegibles por metilación primero y segundo para proporcionar i) un elemento compuesto de mantenimiento y ii) un elemento compuesto truncado,

30 en donde el elemento compuesto de mantenimiento es una disposición en donde la secuencia de reconocimiento de la enzima de restricción de tipo IIS opuesta está dispuesta para dirigir la enzima de restricción de tipo IIS para cortar el ácido nucleico en el mismo sitio que la secuencia de reconocimiento de la enzima de restricción de tipo IIS de la restricción opuesta protegible por metilación, de manera que se produzca el mismo saliente, y

35 en donde el elemento compuesto de truncamiento es una disposición en donde la secuencia de reconocimiento de la enzima de restricción de tipo IIS opuesta del elemento compuesto de truncamiento está dispuesta para dirigir la enzima de restricción de tipo IIS para cortar el ácido nucleico en un sitio que está dentro de la secuencia de reconocimiento de la restricción de tipo IIS secuencia de reconocimiento enzimático del elemento de restricción protegible por metilación, y/o dentro de los nucleótidos entre el sitio de corte y la secuencia de reconocimiento de la secuencia de reconocimiento de enzima de restricción tipo IIS del elemento de restricción protegible por metilación;

40 en donde el segundo ácido nucleico es una enzima de restricción cortada por una enzima de restricción de tipo IIS que reconoce las secuencias de reconocimiento de la enzima de restricción de tipo IIS opuestas, formando así el segundo ácido nucleico metilado linealizado; y en donde los elementos compuestos de mantenimiento y truncamiento están metilados con la metilasa del ADN, y en donde el saliente proporcionado por el elemento compuesto de mantenimiento del primer ácido nucleico metilado linealizado es diferente en secuencia al saliente proporcionado por el elemento de restricción protegible por metilación del segundo ácido nucleico metilado linealizado;

45 proporcionar un segundo fragmento de ADN para ensamblaje que tiene extremos salientes que están adaptados para ligarse a los extremos salientes del segundo ácido nucleico metilado linealizado;

50 ligar el segundo fragmento de ADN para el ensamblaje y el segundo ácido nucleico metilado linealizado con una ligasa para formar un segundo vector intermedio metilado;

transformar el segundo vector intermedio metilado en una cepa bacteriana que no expresa la(s) metilasa del ADN(s) que reconoce(n) cualquiera de las secuencia(s) de reconocimiento de metilasa del ADN de los elementos compuestos de mantenimiento y truncamiento, formando de este modo un segundo vector intermedio que no está metilado;

5 aislar el segundo vector intermedio;

(C) cortar el primer vector intermedio con una enzima de restricción de tipo IIS que reconoce la secuencia de reconocimiento de la enzima de restricción de tipo IIS del elemento compuesto de mantenimiento y una enzima de restricción de tipo IIS que reconoce la secuencia de reconocimiento de la enzima de restricción de tipo IIS del elemento compuesto truncado, formando de este modo un primer inserto de fragmento de ADN adaptado que
10 comprende una secuencia saliente mantenida que está determinada por el elemento compuesto de mantenimiento y una secuencia saliente nativa opuesta que está determinada por la secuencia nativa del primer fragmento de ADN para el ensamblaje;

(D) cortar el segundo vector intermedio con una enzima de restricción de tipo IIS que reconoce la secuencia de reconocimiento de la enzima de restricción de tipo IIS del elemento compuesto de mantenimiento y una enzima de restricción de tipo IIS que reconoce la secuencia de reconocimiento de la enzima de restricción de tipo IIS del elemento compuesto truncado, formando así un segundo inserto de fragmento de ADN adaptado que comprende una secuencia saliente mantenida que está determinada por el elemento compuesto de mantenimiento y una secuencia saliente nativa opuesta que está determinada por la secuencia nativa del segundo fragmento de ADN para el ensamblaje;

20 en donde

(i) el primer y segundo fragmentos de ADN adaptados son fragmentos terminales en donde sus secuencias salientes nativas son complementarias, de modo que están dispuestas para ligarse entre sí; o

(ii) se proporcionan uno o más fragmentos de ADN medios para ensamblaje en donde el primer y segundo fragmentos de ADN adaptados son fragmentos terminales respectivos en el ensamblaje, y el uno o más fragmentos de ADN medios están dispuestos para ligarse entre el primer y segundo fragmentos de ADN adaptados a través de secuencias nativas salientes complementarias;

que comprende además el paso de ligar juntos, con una ligasa, el primer y segundo fragmentos de ADN adaptados, o el primer y segundo fragmentos de ADN adaptados y uno o más fragmentos de ADN intermedios, para formar un fragmento de ADN ensamblado que tiene salientes mantenidos en cada extremo.

30 8. El método según la reivindicación 7, en donde el fragmento de ADN medio comprende una primera secuencia saliente nativa que es complementaria a la saliente nativa del primer fragmento de ADN adaptado y una segunda secuencia saliente nativa que es complementaria a la saliente nativa del segundo fragmento de ADN adaptado.

9. El método según la reivindicación 7, en donde el método comprende dos o más fragmentos de ADN medios para el ensamblaje, los fragmentos de ADN medios comprenden secuencias salientes nativas que son complementarias al saliente nativo de un fragmento de ADN medio vecino, de modo que puedan ligarse entre sí en un orden predeterminado; y en donde el primer y el último fragmento de ADN intermedio de la secuencia están dispuestos para ligarse al primer fragmento de ADN adaptado y al segundo fragmento de ADN adaptado respectivos mediante secuencias sobresalientes nativas complementarias.

40 10. El método según cualquiera de las reivindicaciones 7 a 9, en donde un fragmento de ADN medio para el ensamblaje es proporcionado por

proporcionar otro ácido nucleico metilado linealizado mediante corte con enzima de restricción de un ácido nucleico, en donde el ácido nucleico comprende un primer y un segundo elementos de restricción protegibles contra la metilación con una secuencia de descarte entre ellos, o proporcionar una versión del mismo previamente cortada con enzima de restricción a la que se le ha eliminado la secuencia de descarte,

45 en donde el primer y segundo elementos de restricción protegibles por metilación del ácido nucleico comprenden cada uno:

- una secuencia de reconocimiento de enzima de restricción de tipo IIS; y
- una secuencia de reconocimiento de metilasa del ADN,

50 en donde la secuencia de reconocimiento de la enzima de restricción de tipo IIS y la secuencia de reconocimiento de la metilasa de ADN se solapan de tal manera que la base metilable por la metilasa de ADN se encuentra dentro de la secuencia de reconocimiento de la enzima de restricción de tipo IIS,

en donde la secuencia de reconocimiento de la metilasa del ADN no es idéntica ni está rodeada por la secuencia de reconocimiento de la enzima de restricción tipo IIS, y

en donde la secuencia de reconocimiento de metilasa del ADN no se solapa con la secuencia que formaría la secuencia del extremo sobresaliente generada por la enzima de restricción de tipo IIS, y

en donde el primer y segundo elementos de restricción protegibles por metilación están metilados con la metilasa del ADN que reconoce la secuencia de reconocimiento de la metilasa del ADN; y

5 en donde el ácido nucleico comprende además secuencias opuestas no protegibles de reconocimiento de enzimas de restricción de tipo IIS proporcionadas en el lado opuesto del sitio de corte de cada uno de los elementos de restricción protegibles contra metilación primero y segundo para proporcionar dos elementos compuestos truncadores,

10 en donde el elemento compuesto truncador es una disposición en donde la secuencia de reconocimiento de la enzima de restricción de tipo IIS opuesta del elemento compuesto truncador está dispuesta para dirigir la enzima de restricción de tipo IIS para cortar el ácido nucleico en un sitio que está dentro de la secuencia de reconocimiento de la enzima de restricción de tipo IIS del elemento de restricción metilable-protégible, y/o dentro de los nucleótidos entre el sitio de corte y la secuencia de reconocimiento de la enzima de restricción de tipo IIS del elemento de restricción metilable-protégible;

15 en donde el ácido nucleico es cortado con enzima de restricción por una enzima de restricción tipo IIS que reconoce las secuencias opuestas de reconocimiento de enzimas de restricción tipo IIS, formando así el ácido nucleico metilado linealizado adicional;

proporcionar un fragmento de ADN medio adaptado para ensamblaje que tiene extremos salientes que están adaptados para ligarse a los extremos salientes del ácido nucleico metilado linealizado adicional;

20 ligar el fragmento de ADN medio adaptado para su ensamblaje y ácido nucleico metilado linealizado adicionalmente con una ligasa para formar un vector intermedio metilado;

transformar el vector intermedio metilado en una cepa bacteriana que no expresa la metilasa del ADN que reconoce cualquiera de las secuencias de reconocimiento de la metilasa del ADN de los elementos de restricción protegibles por metilación, formando de ese modo un vector intermedio que no está metilado;

25 aislar el vector intermedio;

cortar el vector intermedio con enzimas de restricción que reconocen la secuencia de reconocimiento de la enzima de restricción tipo IIS de los elementos de restricción protegibles por metilación, formando así un inserto de fragmento de ADN medio que comprende secuencias salientes nativas en cada extremo que están determinadas por la secuencia nativa del fragmento de ADN medio para el ensamblaje.

35 11. El método según cualquiera de las reivindicaciones 7 a 10, que comprende además el paso de proporcionar un vector de destino linealizado para la inserción de los fragmentos de ADN ensamblados, opcionalmente en donde el vector de destino linealizado se proporciona cortando un vector de destino circular con la misma(s) enzima(s) de restricción de tipo IIS que reconocen las secuencias de reconocimiento de la enzima de restricción de tipo IIS del elemento compuesto de mantenimiento y/o truncamiento.

12. Una composición que comprende el ácido nucleico según cualquiera de las reivindicaciones 1 a 6.

13. Un kit que comprende uno o más ácidos nucleicos según cualquiera de las reivindicaciones 1 a 6, y una enzima de restricción y/o una ligasa.

40 14. Una célula huésped que comprende el ácido nucleico según cualquiera de las reivindicaciones 1 a 6.

15. Uso del ácido nucleico según cualquiera de las reivindicaciones 1 a 6 para ensamblar fragmentos de ADN de interés, opcionalmente en donde el uso es *in vitro*.

Figura 1, continuación

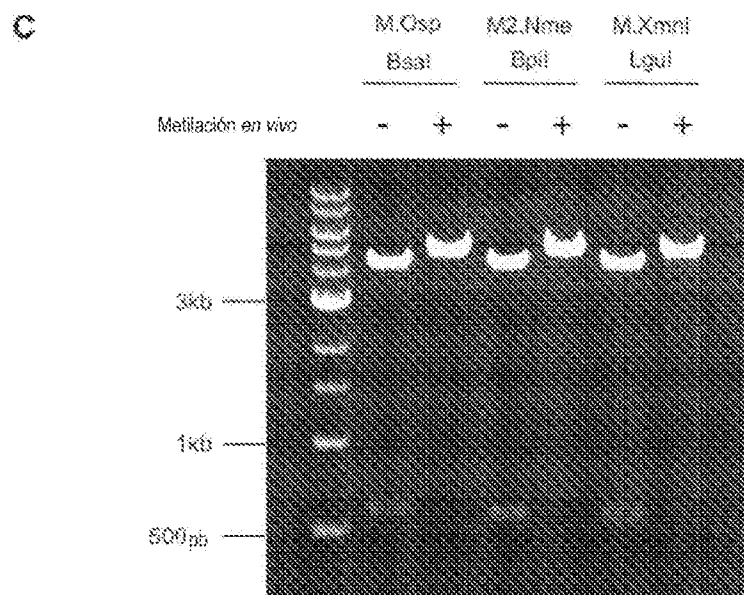


Figura 2

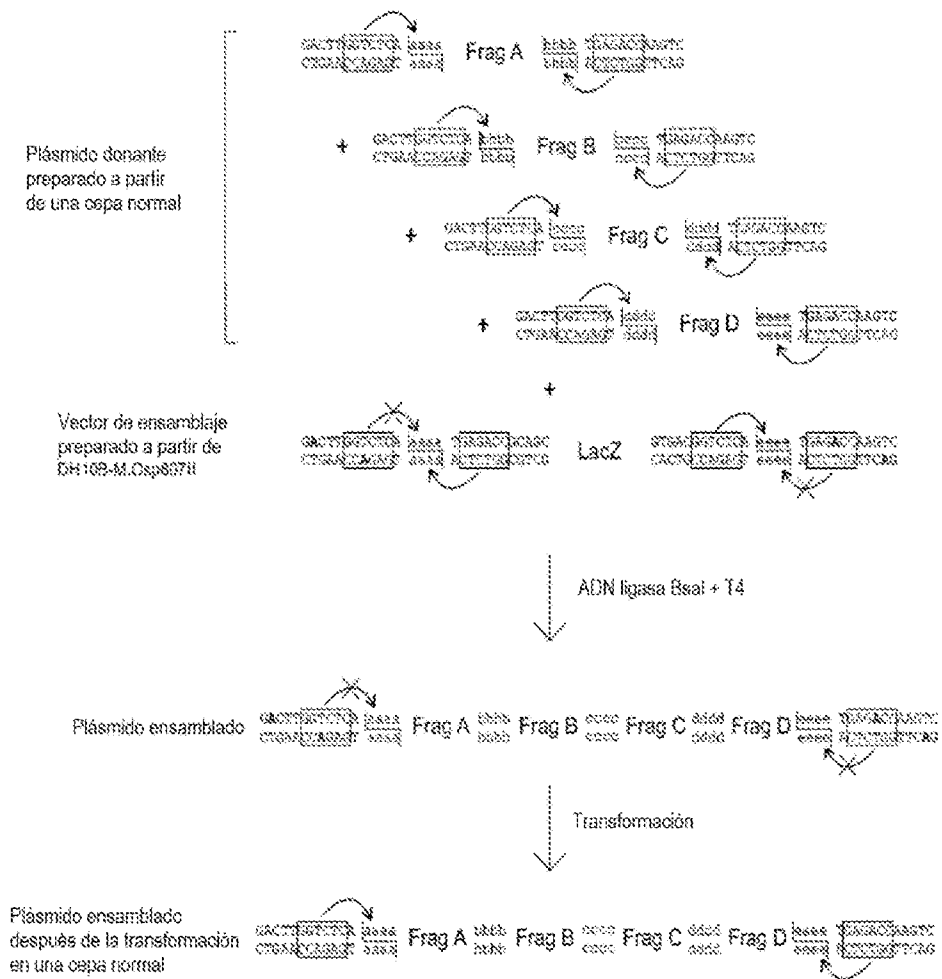


Figura 3

A

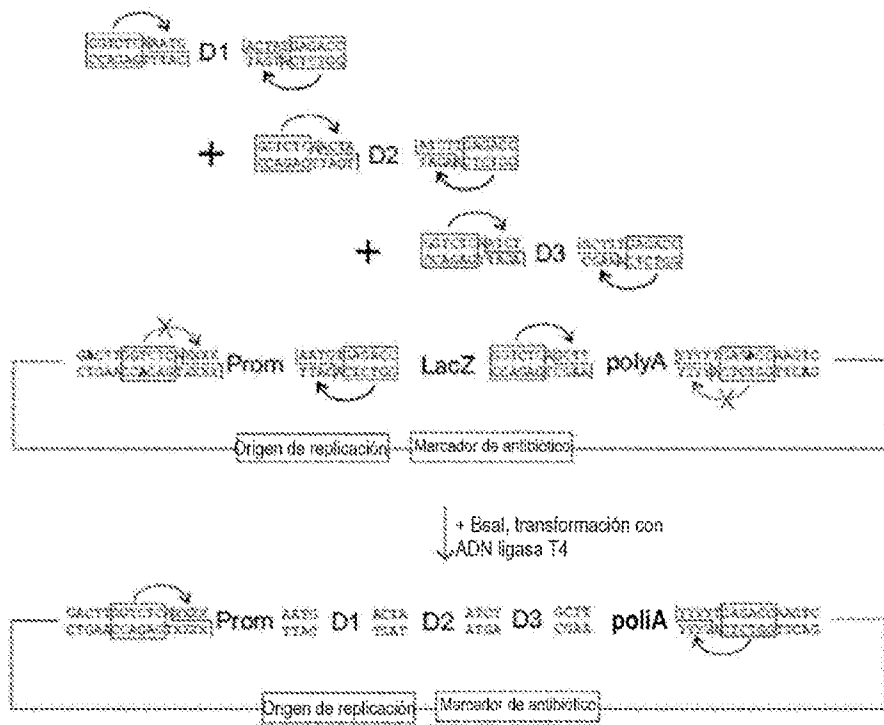


Figura 3, continuación

B

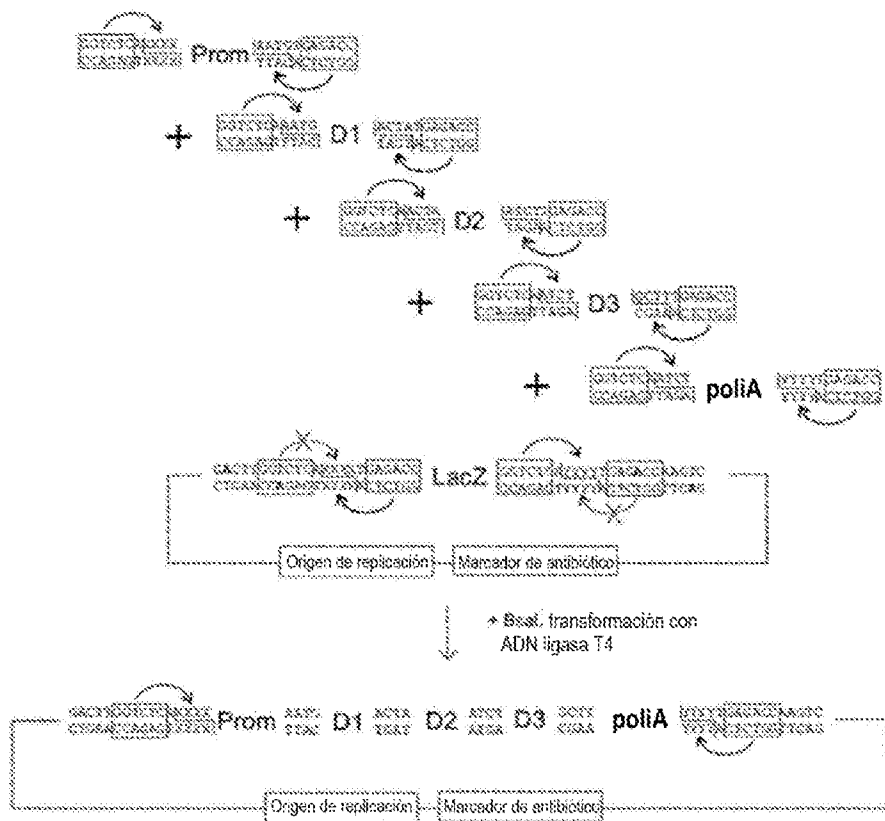


Figura 4

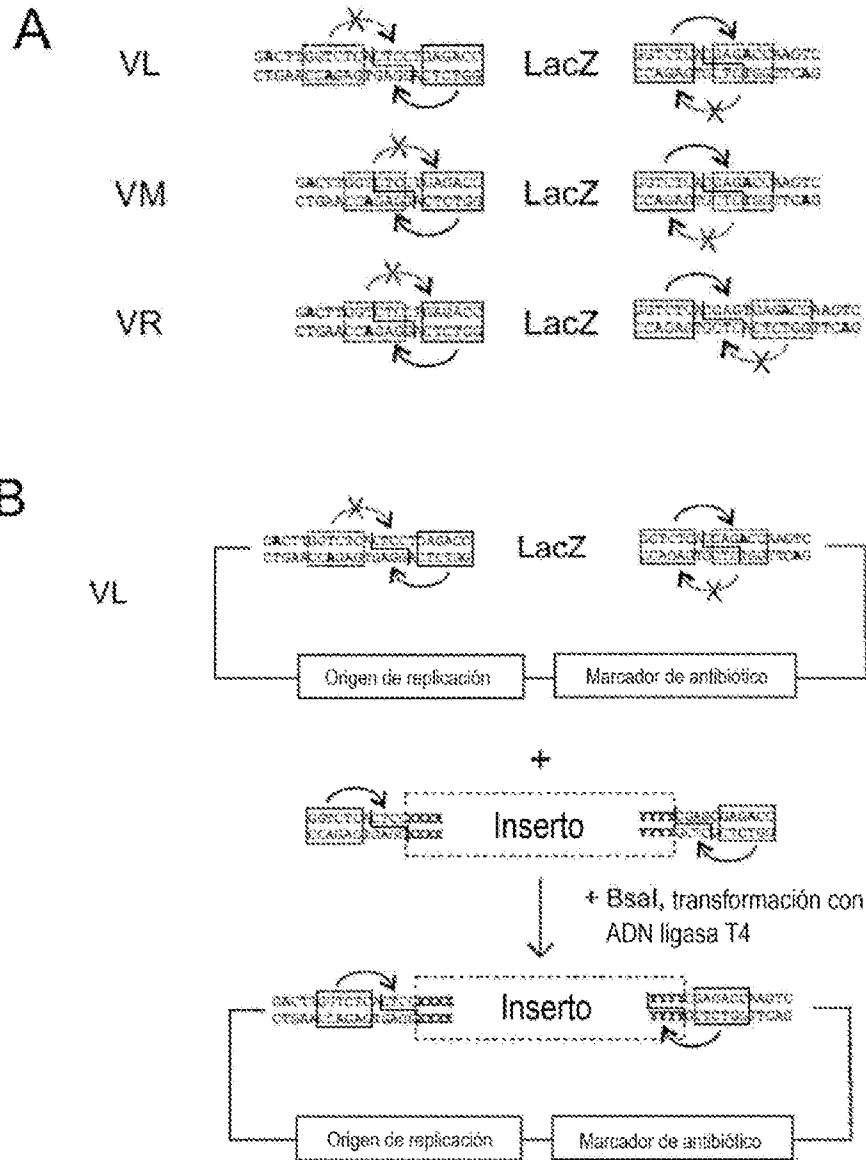


Figura 4, continuación

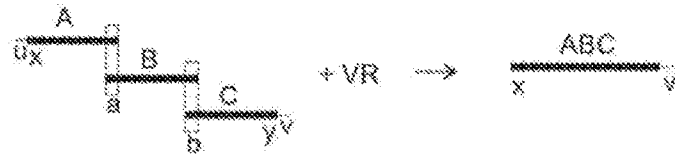
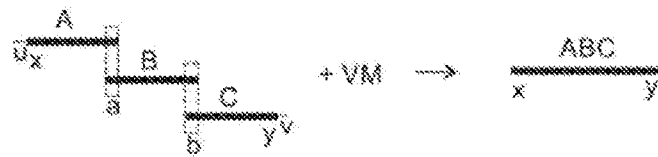
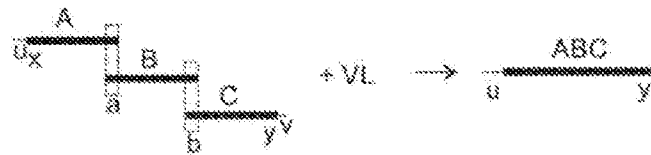
C

$$\text{Inserto}(u, v) + \text{VL} \longrightarrow \text{Inserto}(u, y)$$

$$\text{Inserto}(u, v) + \text{VM} \longrightarrow \text{Inserto}(x, y)$$

$$\text{Inserto}(u, v) + \text{VR} \longrightarrow \text{Inserto}(x, v)$$

D



$$A(u, a) + B(a, b) + C(b, v) + \text{VL} \longrightarrow \text{ABC}(u, y)$$

$$A(u, a) + B(a, b) + C(b, v) + \text{VM} \longrightarrow \text{ABC}(x, y)$$

$$A(u, a) + B(a, b) + C(b, v) + \text{VR} \longrightarrow \text{ABC}(x, v)$$

Figura 5

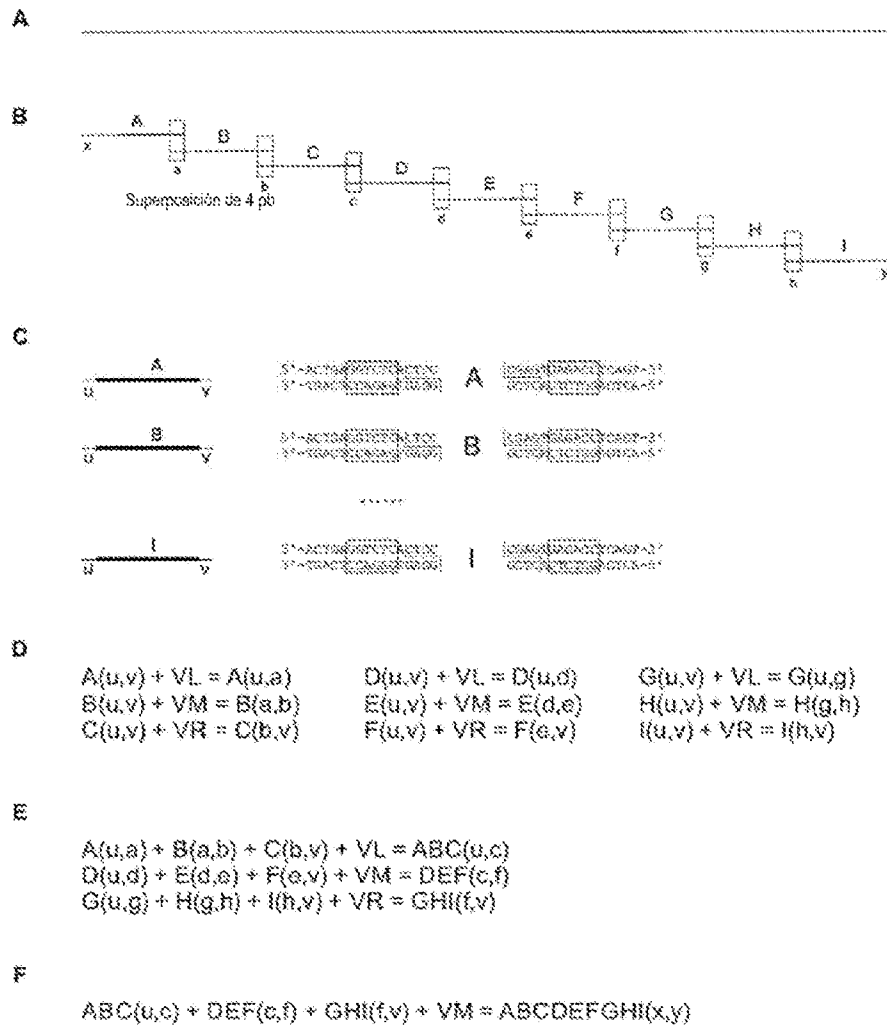


Figura 6

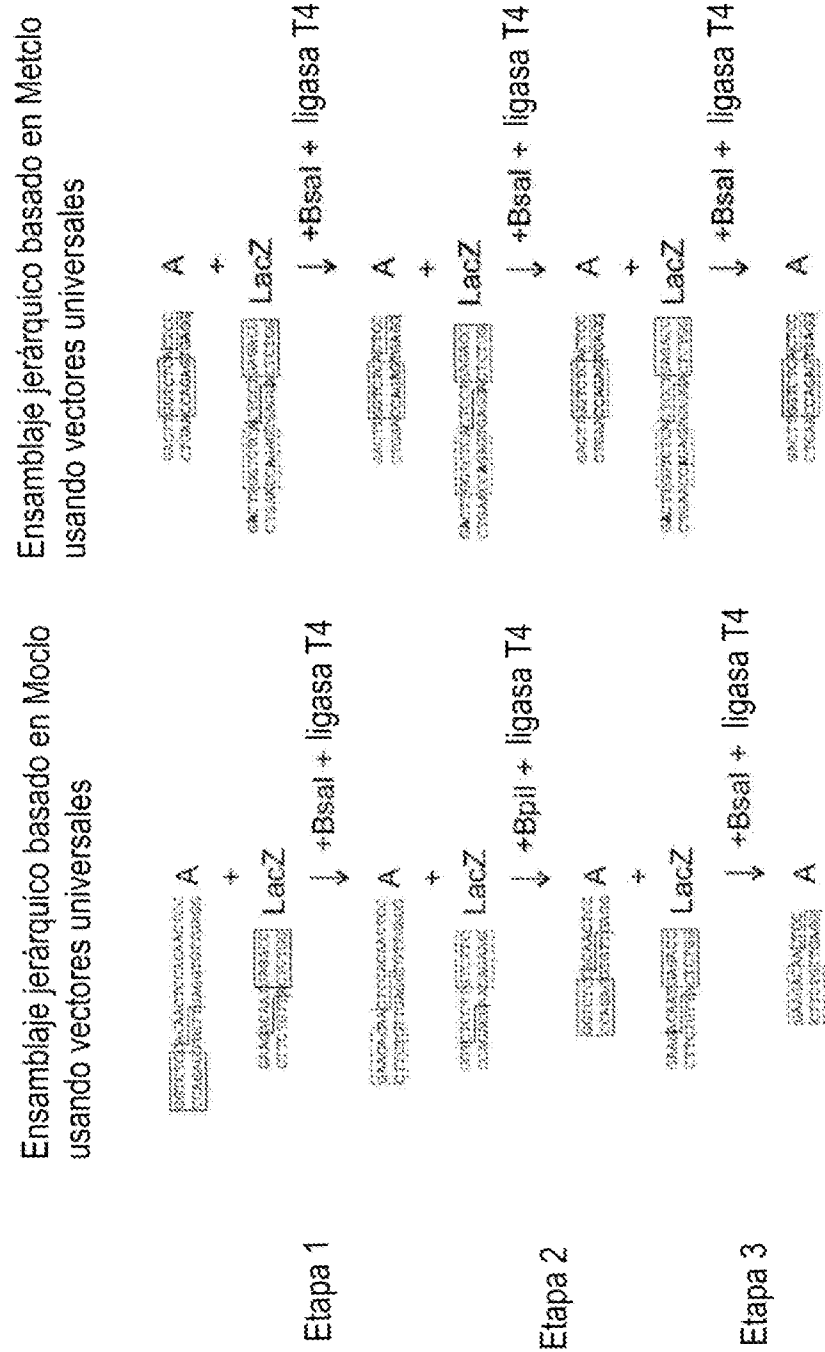


Figura 7

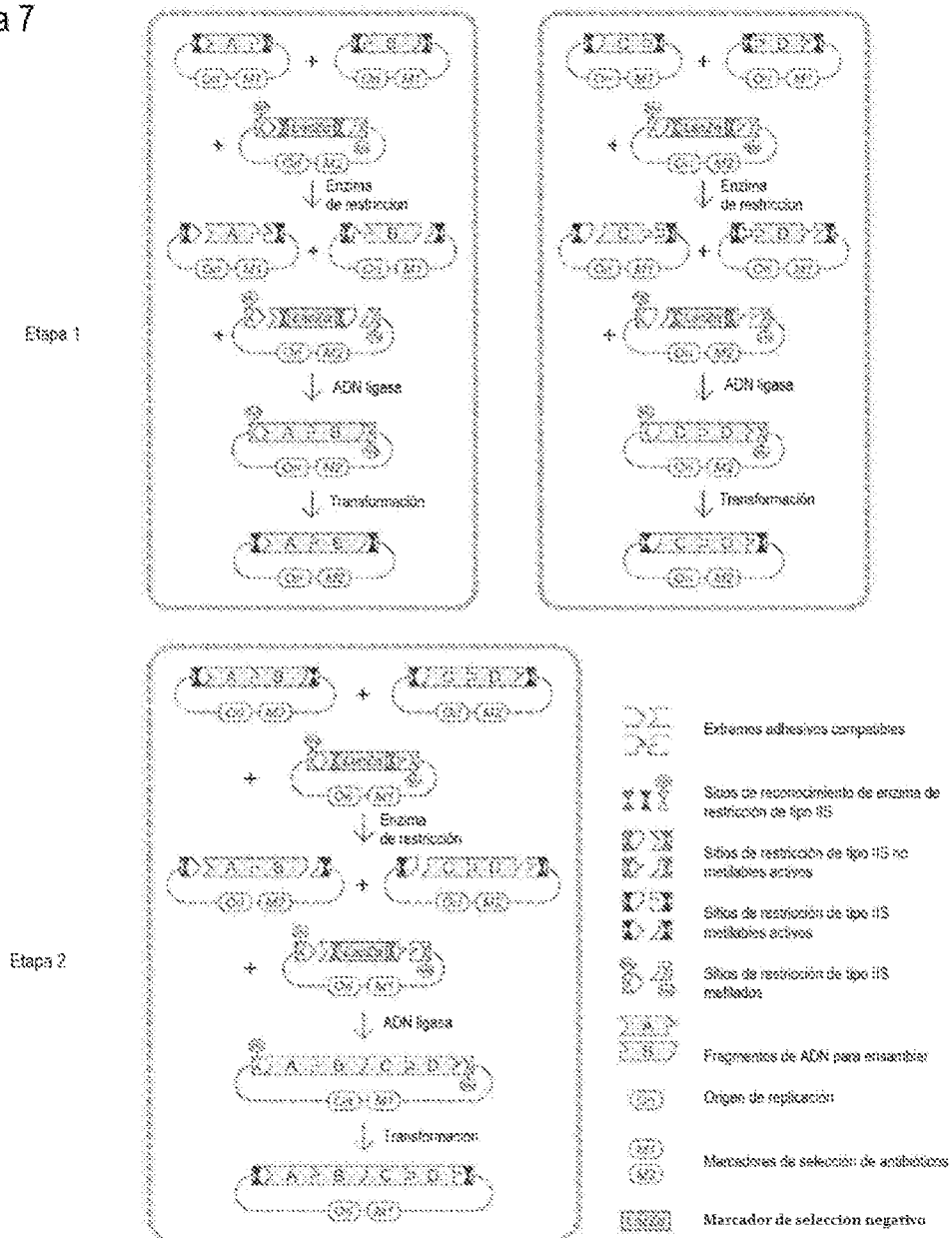


Figura 8

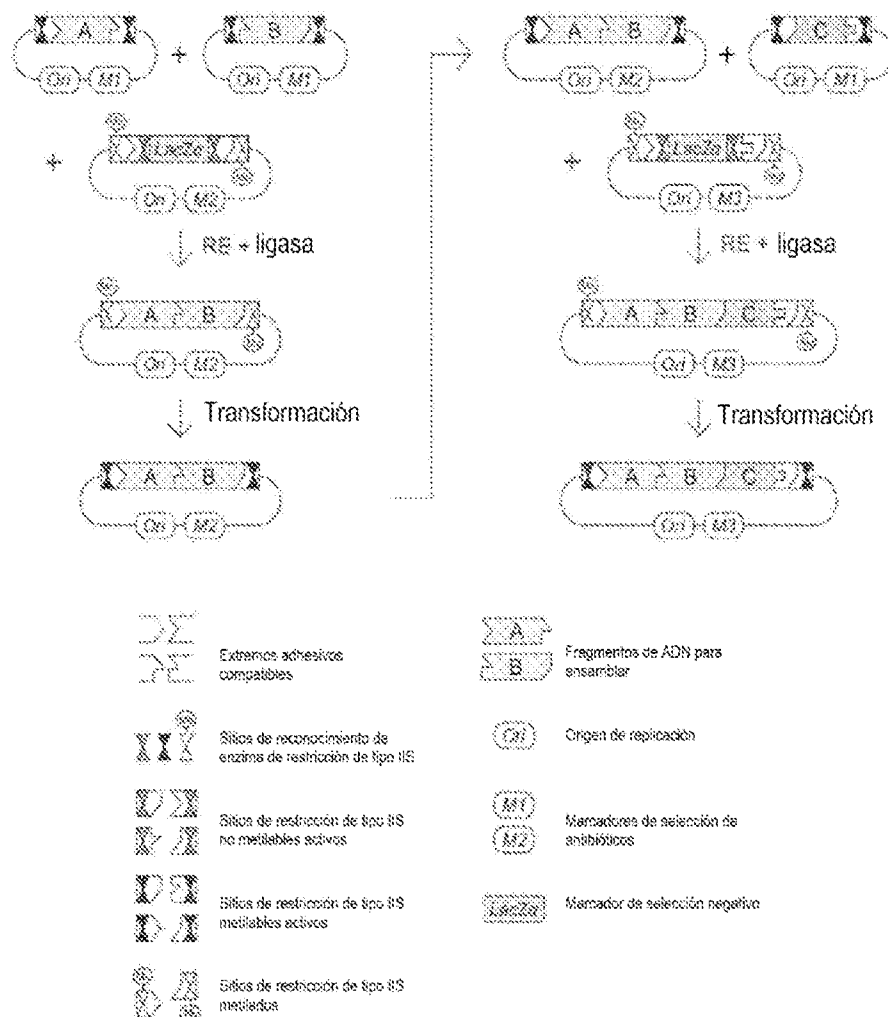


Figura 9

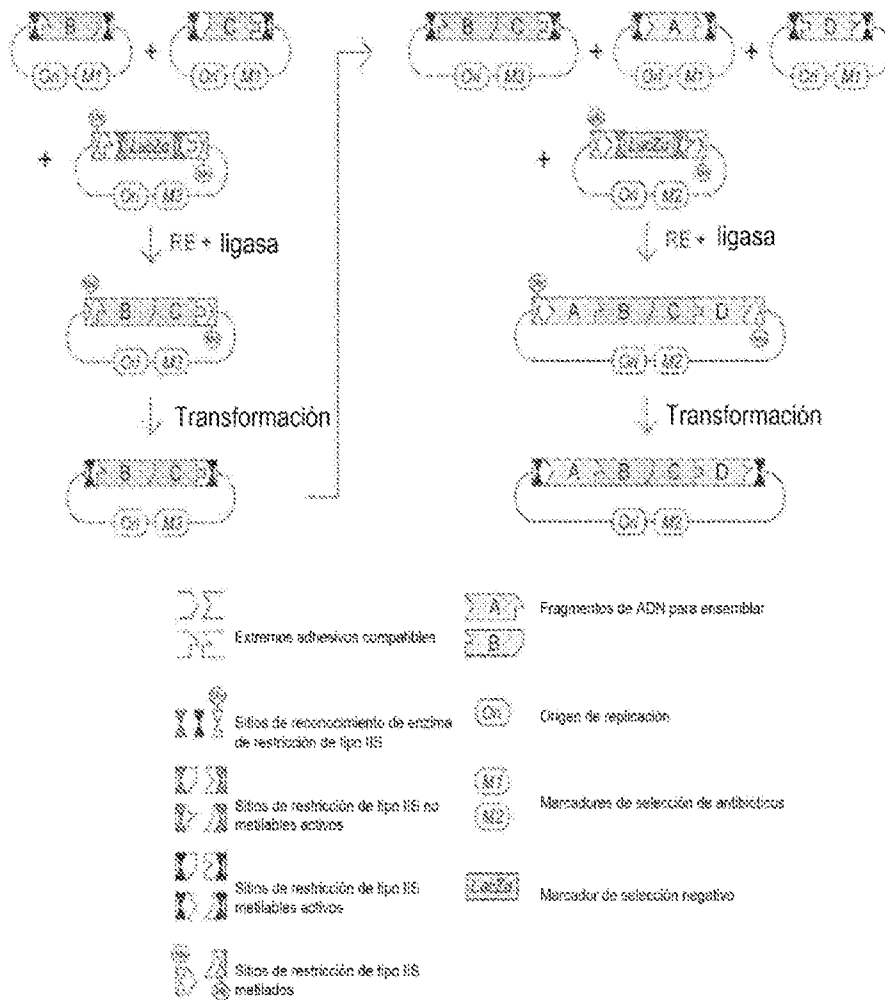


Figura 10

