

**GENE EDITING SYSTEMS COMPRISING AN RNA GUIDE TARGETING
HYDROXYACID OXIDASE 1 (HAO1) AND USES THEREOF**

CROSS REFERENCE TO RELATED APPLICATIONS

5 This application claims the benefit under 35 U.S.C. § 119(e) of U.S. Provisional Application No. 63/197,073, filed June 4, 2021, U.S. Provisional Application No. 63/225,046, filed July 23, 2021, U.S. Provisional Application No. 63/292,889, filed December 22, 2021, and U.S. Provisional Application No. 63/300,727, filed January 19, 2022, the contents of each of which are incorporated by reference herein in their entirety.

SEQUENCE LISTING

10 The instant application contains a Sequence Listing which has been filed electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on June 3, 2022, is named 116928-0040-0004WO00_SEQ.txt and is 367,354 bytes in size.

BACKGROUND

15 Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) genes, collectively known as CRISPR-Cas or CRISPR/Cas systems, are adaptive immune systems in archaea and bacteria that defend particular species against foreign genetic elements.

SUMMARY OF THE INVENTION

20 The present disclosure is based, at least in part, on the development of a system for genetic editing of a hydroxyacid oxidase 1 (HAO1) gene. The system involves a Cas12i CRISPR nuclease polypeptide (e.g., a Cas12i2 polypeptide) and an RNA guide mediating cleavage at a genetic site within the HAO1 gene by the CRISPR nuclease polypeptide. As reported herein, the gene editing system disclosed herein has achieved successful editing of HAO1 gene with high editing efficiency and accuracy.

25 Without being bound by theory, the gene editing system disclosed herein may further exhibit one or more of the following advantageous features. Compared to SpCas9 and Cas12a, Cas12i effectors are smaller (1033 to 1093aa), which, in conjunction with their short mature crRNA (40-43 nt), is preferable in terms of delivery and cost of synthesis. Cas12i cleavage results in larger deletions compared to the small deletions and +1 insertions induced by Cas9

cleavage. Cas12i PAM sequences also differ from those of Cas9. Therefore, larger and different portions of genetic sites of interest can be disrupted with a Cas12i polypeptide and RNA guide compared to Cas9. Using an unbiased approach of tagmentation-based tag integration site sequencing (TTISS), more potential off-target sites with a higher number of unique integration events were identified for SpCas9 compared to Cas12i2. *See* WO/2021/202800. Therefore, 5 Cas12i such as Cas12i2 may be more specific than Cas9.

Accordingly, provided herein are gene editing systems for editing HAO1 gene, pharmaceutical compositions or kits comprising such, methods of using the gene editing systems to produce genetically modified cells, and the resultant cells thus produced. Also provided herein 10 are uses of the gene editing systems disclosed herein, the pharmaceutical compositions and kits comprising such, and/or the genetically modified cells thus produced for treating primary hyperoxaluria (PH) in a subject.

In some aspects, the present disclosure features system for genetic editing of a hydroxyacid oxidase 1 (HAO1) gene, comprising (i) a Cas12i polypeptide or a first nucleic acid 15 encoding the Cas12i polypeptide, and (ii) an RNA guide or a second nucleic acid encoding the RNA guide. The RNA guide comprises a spacer sequence specific to a target sequence within an HAO1 gene, the target sequence being adjacent to a protospacer adjacent motif (PAM) comprising the motif of 5'-TTN-3', which is located 5' to the target sequence.

In some embodiments, the Cas12i polypeptide can be a Cas12i2 polypeptide. In other 20 embodiments, the Cas12i polypeptide can be a Cas12i4 polypeptide.

In some embodiments, the Cas12i polypeptide is a Cas12i2 polypeptide, which comprises an amino acid sequence at least 95% identical to SEQ ID NO: 922 and comprises one or more mutations relative to SEQ ID NO: 922. In some embodiments, the one or more mutations in the Cas12i2 polypeptide are at positions D581, G624, F626, P868, I926, V1030, E1035, and/or 25 S1046 of SEQ ID NO: 922. In some examples, the one or more mutations are amino acid substitutions, which optionally is D581R, G624R, F626R, P868T, I926R, V1030G, E1035R, S1046G, or a combination thereof.

In one example, the Cas12i2 polypeptide comprises mutations at positions D581, D911, I926, and V1030 (*e.g.*, amino acid substitutions of D581R, D911R, I926R, and V1030G). In 30 another example, the Cas12i2 polypeptide comprises mutations at positions D581, I926, and V1030 (*e.g.*, amino acid substitutions of D581R, I926R, and V1030G). In yet another example, the Cas12i2 polypeptide comprises mutations at positions D581, I926, V1030, and S1046 (*e.g.*,

amino acid substitutions of D581R, I926R, V1030G, and S1046G). In still another example, the Cas12i2 polypeptide comprises mutations at positions D581, G624, F626, I926, V1030, E1035, and S1046 (*e.g.*, amino acid substitutions of D581R, G624R, F626R, I926R, V1030G, E1035R, and S1046G). In another example, the Cas12i2 polypeptide comprises mutations at positions
 5 D581, G624, F626, P868, I926, V1030, E1035, and S1046 (*e.g.*, amino acid substitutions of D581R, G624R, F626R, P868T, I926R, V1030G, E1035R, and S1046G).

Exemplary Cas12i2 polypeptides for use in any of the gene editing systems disclosed herein may comprise the amino acid sequence of any one of SEQ ID NOs: 923-927. In one example, the exemplary Cas12i2 polypeptide for use in any of the gene editing systems disclosed
 10 herein comprises the amino acid sequence of SEQ ID NO: 924. In another example, the exemplary Cas12i2 polypeptide for use in any of the gene editing systems disclosed herein comprises the amino acid sequence of SEQ ID NO: 927.

In some embodiments, the gene editing system may comprise the first nucleic acid encoding the Cas12i polypeptide (*e.g.*, the Cas12i2 polypeptide as disclosed herein). In some
 15 instances, the first nucleic acid is located in a first vector (*e.g.*, a viral vector such as an adeno-associated viral vector or AAV vector). In some instances, the first nucleic acid is a messenger RNA (mRNA). In some instances, the nucleic acid encoding the Cas12i polypeptide (*e.g.*, the Cas12i2 polypeptide as disclosed herein) is codon-optimized.

In some embodiments, the target sequence may be within exon 1 or exon 2 of the HAO1
 20 gene. In some examples, the target sequence comprises 5'-CAAAGTCTATATATGACTAT-3' (SEQ ID NO: 1025), 5'-GGAAGTACTGATTTAGCATG-3' (SEQ ID NO: 1026), 5'-TAGATGGAAGCTGTATCCAA-3' (SEQ ID NO: 1046), 5'-CGGAGCATCCTTGGATACAG-3' (SEQ ID NO: 1047), or 5'-AGGACAGAGGGTCAGCATGC-3' (SEQ ID NO: 1052). In specific examples, the target sequence can be the nucleotide sequence of SEQ ID NO: 1047.

In some embodiments, the spacer sequence may be 20-30-nucleotide in length. In some
 25 examples, the spacer sequence is 20-nucleotide in length. In some examples, the spacer sequence comprises 5'-CAAAGUCUAUAUAUGACUAU-3' (SEQ ID NO: 1093); 5'-GGAAGUACUGAUUUAGCAUG-3' (SEQ ID NO: 1094); 5'-UAGAUGGAAGCUGUAUCCAA-3' (SEQ ID NO: 1095); 5'-CGGAGCAUCCUUGGAUACAG-3' (SEQ ID NO: 1096); or 5'-AGGACAGAGGGUCAGCAUGC-3' (SEQ ID NO: 1097). In specific examples, the spacer
 30 sequence may comprise SEQ ID NO: 1096.

In some embodiments, the RNA guide comprises the spacer and a direct repeat sequence. In some examples, the direct repeat sequence is 23-36-nucleotide in length. In one example, the direct repeat sequence is at least 90% identical to any one of SEQ ID NOs: 1-10 or a fragment thereof that is at least 23-nucleotide in length. In some specific examples, the direct repeat sequence is any one of SEQ ID NOs: 1-10, or a fragment thereof that is at least 23-nucleotide in length. By way of non-limiting example, the direct repeat sequence is 5'-AGAAAUCCGUCUUUCAUUGACGG-3' (SEQ ID NO: 10).

In specific examples, the RNA guide may comprise the nucleotide sequence of 5'-AGAAAUCCGUCUUUCAUUGACGGCAAAGUCUAUAUAUGACUAU-3' (SEQ ID NO: 967), 5'-AGAAAUCCGUCUUUCAUUGACGGGGAAGUACUGAUUUAGCAUG-3' (SEQ ID NO: 968), 5'-AGAAAUCCGUCUUUCAUUGACGGUAGAUGGAAGCUGUAUCCAA-3' (SEQ ID NO: 988), 5'-AGAAAUCCGUCUUUCAUUGACGGCGGAGCAUCCUUGGAUACAG-3' (SEQ ID NO: 989), or 5'-AGAAAUCCGUCUUUCAUUGACGGAGGACAGAGGGUCAGCAUGC-3' (SEQ ID NO: 994). In specific examples, the RNA guide may comprise SEQ ID NO: 989.

In some embodiments, the system may comprise the second nucleic acid encoding the RNA guide. In some examples, the nucleic acid encoding the RNA guide may be located in a viral vector. In some examples, the viral vector comprises the both the first nucleic acid encoding the Cas12i2 polypeptide and the second nucleic acid encoding the RNA guide.

In some embodiments, any of the systems described herein may comprise the first nucleic acid encoding the Cas12i2 polypeptide, which is located in a first vector, and the second nucleic acid encoding the RNA guide, which is located on a second vector. In some examples, the first and/or second vector is a viral vector. In some specific examples, the first and second vectors are the same vector. In other examples, the first and second vectors are different vectors.

In some embodiments, any of the systems described herein may comprise one or more lipid nanoparticles (LNPs), which encompass the Cas12i2 polypeptide or the first nucleic acid encoding the Cas12i2 polypeptide, the RNA guide or the second nucleic acid encoding the RNA guide, or both.

In some embodiments, the system described herein may comprise a LNP, which encompass the Cas12i2 polypeptide or the first nucleic acid encoding the Cas12i2 polypeptide, and a viral vector comprising the second nucleic acid encoding the RNA guide. In some examples, the viral vector is an AAV vector. In other embodiments, the system described herein may comprise a LNP, which encompass the RNA guide or the second nucleic acid encoding the

RNA guide, and a viral vector comprising the first nucleic acid encoding the Cas12i2 polypeptide. In some examples, the viral vector is an AAV vector.

In some aspects, the present disclosure also provides a pharmaceutical composition comprising any of the gene editing systems disclosed herein, or a kit comprising the components
5 of the gene editing system.

In other aspects, the present disclosure also features a method for editing a hydroxyacid oxidase 1 (HAO1) gene in a cell, the method comprising contacting a host cell with any of the systems disclosed herein to genetically edit the HAO1 gene in the host cell. In some examples, the host cell is cultured *in vitro*. In other examples, the contacting step is performed by
10 administering the system for editing the HAO1 gene to a subject comprising the host cell.

Also within the scope of the present disclosure is a cell comprising a disrupted a hydroxyacid oxidase 1 (HAO1) gene, which can be produced by contacting a host cell with the system disclosed herein genetically edit the HAO1 gene in the host cell.

Still in other aspects, the present disclosure provides a method for treating primary
15 hyperoxaluria (PH) in a subject. The method may comprise administering to a subject in need thereof any of the systems for editing a hydroxyacid oxidase 1 (HAO1) gene or any of the modified cells disclosed herein. In some embodiments, the subject may be a human patient having the PH. In some examples, the PH is PH1, PH2, or PH3. In a specific example, the PH is PH1.

Also provided herein is an RNA guide, comprising (i) a spacer sequence as disclosed
20 herein that is specific to a target sequence in a hydroxyacid oxidase 1 (HAO1) gene, wherein the target sequence is adjacent to a protospacer adjacent motif (PAM) comprising the motif of 5'-TTN-3', which is located 5' to the target sequence; and (ii) a direct repeat sequence.

In some embodiments, the spacer may be 20-30-nucleotide in length. In some examples,
25 the spacer is 20-nucleotide in length.

In some embodiments, the direct repeat sequence may be 23-36-nucleotide in length. In some examples, the direct repeat sequence is 23-nucleotide in length.

In some embodiments, the target sequence may be within exon 1 or exon 2 of the HAO1 gene. In some examples, the target sequence comprises 5'-CAAAGTCTATATATGACTAT-3'
30 (SEQ ID NO: 1025), 5'-GGAAGTACTGATTAGCATG-3' (SEQ ID NO: 1026), 5'-TAGATGGAAGCTGTATCCAA-3' (SEQ ID NO: 1046), 5'-CGGAGCATCCTTGGATACAG-

3' (SEQ ID NO: 1047), or 5'-AGGACAGAGGGTCAGCATGC-3' (SEQ ID NO: 1052). In specific examples, the target sequence may comprise SEQ ID NO: 1047.

In some embodiments, the spacer sequence may be set forth as 5'-CAAAGUCUAUAUAUGACUAU-3' (SEQ ID NO: 1093); 5'-GGAAGUACUGAUUUAGCAUG-3' (SEQ ID NO: 1094); 5'-UAGAUGGAAGCUGUAUCCAA-3' (SEQ ID NO: 1095); 5'-CGGAGCAUCCUUGGAUACAG-3' (SEQ ID NO: 1096); or 5'-AGGACAGAGGGUCAGCAUGC-3' (SEQ ID NO: 1097). In specific examples, the spacer sequence may comprise SEQ ID NO: 1096.

In some embodiments, the direct repeat sequence may be at least 90% identical to any one of SEQ ID NOs: 1-10 or a fragment thereof that is at least 23-nucleotide in length. In some examples, the direct repeat sequence is any one of SEQ ID NOs: 1-10, or a fragment thereof that is at least 23-nucleotide in length. By way of non-limiting example, the direct repeat sequence is 5'-AGAAAUCCGUCUUUCAUUGACGG-3' (SEQ ID NO: 10).

In some embodiments, the RNA guide may comprise the nucleotide sequence of 5'-AGAAAUCCGUCUUUCAUUGACGGCAAAGUCUAUAUAUGACUAU-3' (SEQ ID NO: 967), 5'-AGAAAUCCGUCUUUCAUUGACGGGGAAGUACUGAUUUAGCAUG-3' (SEQ ID NO: 968), 5'-AGAAAUCCGUCUUUCAUUGACGGUAGAUGGAAGCUGUAUCCAA-3' (SEQ ID NO: 988), 5'-AGAAAUCCGUCUUUCAUUGACGGCGGAGCAUCCUUGGAUACAG-3' (SEQ ID NO: 989), or 5'-AGAAAUCCGUCUUUCAUUGACGGAGGACAGAGGGUCAGCAUGC-3' (SEQ ID NO: 994). In specific examples, the RNA guide may comprise SEQ ID NO: 989.

Also provided herein are any of the gene editing systems disclosed herein, pharmaceutical compositions or kits comprising such, or genetically modified cells generated by the gene editing system for use in treating PH in a subject, as well as uses of the gene editing systems disclosed herein, pharmaceutical compositions or kits comprising such, or genetically modified cells generated by the gene editing system for manufacturing a medicament for treatment of PH in a subject.

The details of one or more embodiments of the invention are set forth in the description below. Other features or advantages of the present invention will be apparent from the following drawings and detailed description of several embodiments, and also from the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present disclosure, which can be better understood by

reference to the drawing in combination with the detailed description of specific embodiments presented herein.

FIG. 1 is a graph showing the ability of RNPs prepared with a Cas12i2 polypeptide and a crRNA to edit the HAO1 gene in HEK293 cells. The darker grey bars represent target sequences with perfect homology to both rhesus macaque (*Macaca mulatta*) and crab-eating macaque (*Macaca fascicularis*) sequences.

FIG. 2 is a graph showing the ability of RNPs prepared with a Cas12i2 polypeptide and a crRNA to edit the HAO1 gene in HepG2 cells.

FIG. 3 is a graph showing the ability of RNPs prepared with a Cas12i2 polypeptide and a crRNA to edit the HAO1 gene in primary hepatocytes.

FIG. 4 is a graph showing knockdown of HAO1 mRNA in primary human hepatocytes with a Cas12i2 polypeptide and an HAO1-targeting crRNA.

FIG. 5A is a graph showing % indels induced by an HAO1-targeting crRNA and the variant Cas12i2 polypeptide of SEQ ID NO: 924 or SEQ ID NO: 927 in HepG2 cells. **FIG. 5B** shows the size (left) and start position (right) of indels induced in HepG2 cells by the variant Cas12i2 of SEQ ID NO: 924 and the HAO1-targeting RNA guide of E1T3 (SEQ ID NO: 968).

FIG. 6 is a graph showing % indels induced by chemically modified HAO1-targeting crRNAs of SEQ ID NO: 1091 and SEQ ID NO: 1092 and the variant Cas12i2 mRNA of SEQ ID NO: 1089 or SEQ ID NO: 1090.

FIG. 7A shows plots depicting tagmentation-based tag integration site sequencing (TTISS) reads for variant Cas12i2 of SEQ ID NO: 924 and HAO1-targeting RNA guides E2T5 (SEQ ID NO: 989), E1T2 (SEQ ID NO: 967), E1T3 (SEQ ID NO: 968), and E2T10 (SEQ ID NO: 994). The black wedge and centered number represent the fraction of on-target TTISS reads. Each gray wedge represents a unique off-target site identified by TTISS. The size of each gray wedge represents the fraction of TTISS reads mapping to a given off-target. **FIG. 7B** shows plots depicting two replicates of TTISS reads for variant Cas12i2 of SEQ ID NO: 927 and HAO1-targeting RNA guides E2T5 (SEQ ID NO: 989), E1T2 (SEQ ID NO: 967), and E1T3 (SEQ ID NO: 968). The black wedge and centered number represent the fraction of on-target TTISS reads. Each gray wedge represents a unique off-target site identified by TTISS. The size of each gray wedge represents the fraction of TTISS reads mapping to a given off-target.

FIG. 8 is a Western Blot showing knockdown of HAO1 protein following electroporation of primary human hepatocytes with variant Cas12i2 of SEQ ID NO: 924 and RNA guide E2T5 (SEQ ID NO: 989).

DETAILED DESCRIPTION

The present disclosure relates to a system for genetic editing of a hydroxyacid oxidase 1 (HAO1) gene (*a.k.a.*, glycolate oxidase gene), which comprises (i) a Cas12i polypeptide or a first nucleic acid encoding the Cas12i polypeptide, and (ii) an RNA guide or a second nucleic acid encoding the RNA guide, wherein the RNA guide comprises a spacer sequence specific to a target sequence within an HAO1 gene, the target sequence being adjacent to a protospacer adjacent motif (PAM) comprising the motif of 5'-TTN-3', which is located 5' to the target sequence. Also provided in the present disclosure are a pharmaceutical composition or a kit comprising such system as well as uses thereof. Further disclosed herein are a method for editing a HAO1 gene in a cell, a cell so produced that comprises a disrupted a HAO1 gene, a method of treating primary hyperoxaluria (PH) in a subject, and an RNA guide that comprises (i) a spacer that is specific to a target sequence in a HAO1 gene, wherein the target sequence is adjacent to a protospacer adjacent motif (PAM) comprising the motif of 5'-TTN-3', which is located 5' to the target sequence; and (ii) a direct repeat sequence as well as uses thereof.

The Cas12i polypeptide for use in the gene editing system disclosed herein may be a Cas12i2 polypeptide, e.g., a wild-type Cas12i polypeptide or a variant thereof as those disclosed herein. In some examples, the Cas12i2 polypeptide comprises an amino acid sequence at least 95% identical to SEQ ID NO: 922 and comprises one or more mutations relative to SEQ ID NO: 922. In other examples, the Cas12i polypeptide may be a Cas12i4 polypeptide, which is also disclosed herein.

Definitions

The present disclosure will be described with respect to particular embodiments and with reference to certain Figures, but the disclosure is not limited thereto but only by the claims. Terms as set forth hereinafter are generally to be understood in their common sense unless indicated otherwise.

As used herein, the term “activity” refers to a biological activity. In some embodiments, activity includes enzymatic activity, *e.g.*, catalytic ability of a Cas12i polypeptide. For example, activity can include nuclease activity.

As used herein the term “HAO1” refers to “glycolate oxidase 1,” which is also known as “hydroxyacid oxidase.” HAO1 is a peroxisome protein expressed primarily in the liver and pancreas, and its activities include oxidation of glycolate and 2-hydroxy fatty acids. SEQ ID NO: 928 as set forth herein provides an example of an HAO1 gene sequence.

As used herein, the term “Cas12i polypeptide” (also referred to herein as Cas12i) refers to a polypeptide that binds to a target sequence on a target nucleic acid specified by an RNA guide, wherein the polypeptide has at least some amino acid sequence homology to a wild-type Cas12i polypeptide. In some embodiments, the Cas12i polypeptide comprises at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% sequence identity with any one of SEQ ID NOs: 1-5 and 11-18 of U.S. Patent No. 10,808,245, which is incorporated by reference for the subject matter and purpose referenced herein. In some embodiments, a Cas12i polypeptide comprises at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% sequence identity with any one of SEQ ID NOs: 8, 2, 11, and 9 of the present application. In some embodiments, a Cas12i polypeptide of the disclosure is a Cas12i2 polypeptide as described in WO/2021/202800, the relevant disclosures of which are incorporated by reference for the subject matter and purpose referenced herein. In some embodiments, the Cas12i polypeptide cleaves a target nucleic acid (*e.g.*, as a nick or a double strand break).

As used herein, the term “adjacent to” refers to a nucleotide or amino acid sequence in close proximity to another nucleotide or amino acid sequence. In some embodiments, a nucleotide sequence is adjacent to another nucleotide sequence if no nucleotides separate the two sequences (*i.e.*, immediately adjacent). In some embodiments, a nucleotide sequence is adjacent to another nucleotide sequence if a small number of nucleotides separate the two sequences (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides). In some embodiments, a first sequence is adjacent to a second sequence if the two sequences are separated by about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 nucleotides. In some embodiments, a

first sequence is adjacent to a second sequence if the two sequences are separated by up to 2 nucleotides, up to 5 nucleotides, up to 8 nucleotides, up to 10 nucleotides, up to 12 nucleotides, or up to 15 nucleotides. In some embodiments, a first sequence is adjacent to a second sequence if the two sequences are separated by 2-5 nucleotides, 4-6 nucleotides, 4-8 nucleotides, 4-10
5 nucleotides, 6-8 nucleotides, 6-10 nucleotides, 6-12 nucleotides, 8-10 nucleotides, 8-12 nucleotides, 10-12 nucleotides, 10-15 nucleotides, or 12-15 nucleotides.

As used herein, the term “complex” refers to a grouping of two or more molecules. In some embodiments, the complex comprises a polypeptide and a nucleic acid molecule interacting with (*e.g.*, binding to, coming into contact with, adhering to) one another. For
10 example, the term “complex” can refer to a grouping of an RNA guide and a polypeptide (*e.g.*, a Cas12i polypeptide). Alternatively, the term “complex” can refer to a grouping of an RNA guide, a polypeptide, and the complementary region of a target sequence. In another example, the term “complex” can refer to a grouping of an HAO1-targeting RNA guide and a Cas12i polypeptide.

As used herein, the term “protospacer adjacent motif” or “PAM” refers to a DNA
15 sequence adjacent to a target sequence (*e.g.*, an HAO1 target sequence) to which a complex comprising an RNA guide (*e.g.*, an HAO1-targeting RNA guide) and a Cas12i polypeptide binds. In a double-stranded DNA molecule, the strand containing the PAM motif is called the “PAM-strand” and the complementary strand is called the “non-PAM strand.” The RNA guide binds to a site in the non-PAM strand that is complementary to a target sequence disclosed
20 herein.

In some embodiments, the PAM strand is a coding (*e.g.*, sense) strand. In other embodiments, the PAM strand is a non-coding (*e.g.*, antisense strand). Since an RNA guide binds the non-PAM strand *via* base-pairing, the non-PAM strand is also known as the target strand, while the PAM strand is also known as the non-target strand.

As used herein, the term “target sequence” refers to a DNA fragment adjacent to a PAM
25 motif (on the PAM strand). The complementary region of the target sequence is on the non-PAM strand. A target sequence may be immediately adjacent to the PAM motif. Alternatively, the target sequence and the PAM may be separately by a small sequence segment (*e.g.*, up to 5 nucleotides, for example, up to 4, 3, 2, or 1 nucleotide). A target sequence may be located at the
30 3’ end of the PAM motif or at the 5’ end of the PAM motif, depending upon the CRISPR nuclease that recognizes the PAM motif, which is known in the art. For example, a target sequence is located at the 3’ end of a PAM motif for a Cas12i polypeptide (*e.g.*, a Cas12i2

polypeptide such as those disclosed herein). In some embodiments, the target sequence is a sequence within an HAO1 gene sequence, including, but not limited, to the sequence set forth in SEQ ID NO: 928.

As used herein, the term “adjacent to” refers to a nucleotide or amino acid sequence in close proximity to another nucleotide or amino acid sequence. In some embodiments, a nucleotide sequence is adjacent to another nucleotide sequence if no nucleotides separate the two sequences (*i.e.*, immediately adjacent). In some embodiments, a nucleotide sequence is adjacent to another nucleotide sequence if a small number of nucleotides separate the two sequences (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides). In some embodiments, a first sequence is adjacent to a second sequence if the two sequences are separated by about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 nucleotides. In some embodiments, a first sequence is adjacent to a second sequence if the two sequences are separated by up to 2 nucleotides, up to 5 nucleotides, up to 8 nucleotides, up to 10 nucleotides, up to 12 nucleotides, or up to 15 nucleotides. In some embodiments, a first sequence is adjacent to a second sequence if the two sequences are separated by 2-5 nucleotides, 4-6 nucleotides, 4-8 nucleotides, 4-10 nucleotides, 6-8 nucleotides, 6-10 nucleotides, 6-12 nucleotides, 8-10 nucleotides, 8-12 nucleotides, 10-12 nucleotides, 10-15 nucleotides, or 12-15 nucleotides.

As used herein, the term “spacer” or “spacer sequence” is a portion in an RNA guide that is the RNA equivalent of the target sequence (a DNA sequence). The spacer contains a sequence capable of binding to the non-PAM strand *via* base-pairing at the site complementary to the target sequence (in the PAM strand). Such a spacer is also known as specific to the target sequence. In some instances, the spacer may be at least 75% identical to the target sequence (*e.g.*, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99%), except for the RNA-DNA sequence difference. In some instances, the spacer may be 100% identical to the target sequence except for the RNA-DNA sequence difference.

As used herein, the term “RNA guide” or “RNA guide sequence” refers to any RNA molecule or a modified RNA molecule that facilitates the targeting of a polypeptide (*e.g.*, a Cas12i polypeptide) described herein to a target sequence (*e.g.*, a sequence of an HAO1 gene). For example, an RNA guide can be a molecule that is designed to be complementary to a specific nucleic acid sequence (a target sequence such as a target sequence with an HAO1 gene). An RNA guide may comprise a spacer sequence and a direct repeat (DR) sequence. In some instances, the RNA guide can be a modified RNA molecule comprising one or more

deoxyribonucleotides, for example, in a DNA-binding sequence contained in the RNA guide, which binds a sequence complementary to the target sequence. In some examples, the DNA-binding sequence may contain a DNA sequence or a DNA/RNA hybrid sequence. The terms CRISPR RNA (crRNA), pre-crRNA and mature crRNA are also used herein to refer to an RNA guide.

As used herein, the term “complementary” refers to a first polynucleotide (*e.g.*, a spacer sequence of an RNA guide) that has a certain level of complementarity to a second polynucleotide (*e.g.*, the complementary sequence of a target sequence) such that the first and second polynucleotides can form a double-stranded complex *via* base-pairing to permit an effector polypeptide that is complexed with the first polynucleotide to act on (*e.g.*, cleave) the second polynucleotide. In some embodiments, the first polynucleotide may be substantially complementary to the second polynucleotide, *i.e.*, having at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% complementarity to the second polynucleotide. In some embodiments, the first polynucleotide is completely complementary to the second polynucleotide, *i.e.*, having 100% complementarity to the second polynucleotide.

The “percent identity” (*a.k.a.*, sequence identity) of two nucleic acids or of two amino acid sequences is determined using the algorithm of Karlin and Altschul *Proc. Natl. Acad. Sci. USA* 87:2264-68, 1990, modified as in Karlin and Altschul *Proc. Natl. Acad. Sci. USA* 90:5873-77, 1993. Such an algorithm is incorporated into the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al. J. Mol. Biol.* 215:403-10, 1990. BLAST nucleotide searches can be performed with the NBLAST program, score=100, wordlength=12 to obtain nucleotide sequences homologous to the nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score=50, word length=3 to obtain amino acid sequences homologous to the protein molecules of the invention. Where gaps exist between two sequences, Gapped BLAST can be utilized as described in Altschul *et al., Nucleic Acids Res.* 25(17):3389-3402, 1997. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used.

As used herein, the term “edit” refers to one or more modifications introduced into a target nucleic acid, *e.g.*, within the HAO1 gene. The edit can be one or more substitutions, one or more insertions, one or more deletions, or a combination thereof. As used herein, the term “substitution” refers to a replacement of a nucleotide or nucleotides with a different nucleotide or

nucleotides, relative to a reference sequence. As used herein, the term “insertion” refers to a gain of a nucleotide or nucleotides in a nucleic acid sequence, relative to a reference sequence. As used herein, the term “deletion” refers to a loss of a nucleotide or nucleotides in a nucleic acid sequence, relative to a reference sequence.

5 No particular process is implied in how to make a sequence comprising a deletion. For instance, a sequence comprising a deletion can be synthesized directly from individual nucleotides. In other embodiments, a deletion is made by providing and then altering a reference sequence. The nucleic acid sequence can be in a genome of an organism. The nucleic acid sequence can be in a cell. The nucleic acid sequence can be a DNA sequence. The deletion can
10 be a frameshift mutation or a non-frameshift mutation. A deletion described herein refers to a deletion of up to several kilobases.

 As used herein, the terms “upstream” and “downstream” refer to relative positions within a single nucleic acid (*e.g.*, DNA) sequence in a nucleic acid molecule. “Upstream” and “downstream” relate to the 5’ to 3’ direction, respectively, in which RNA transcription occurs. A
15 first sequence is upstream of a second sequence when the 3’ end of the first sequence occurs before the 5’ end of the second sequence. A first sequence is downstream of a second sequence when the 5’ end of the first sequence occurs after the 3’ end of the second sequence. In some embodiments, the 5’-NTTN-3’ or 5’-TTN-3’ sequence is upstream of an indel described herein, and a Cas12i-induced indel is downstream of the 5’-NTTN-3’ or 5’-TTN-3’ sequence.

20

I. Gene Editing Systems

 In some aspects, the present disclosure provides gene editing systems comprising an RNA guide targeting an HAO1 gene. Such a gene editing system can be used to edit the HAO1 target gene, *e.g.*, to disrupt the HAO1 gene.

25 Hydroxyacid oxidase 1 (HAO1, also known as glycolate oxidase [GOX or GO]), converts glycolate into glyoxylate. It has been proposed that inhibition of HAO1 in individuals with PH1 would block formation of glyoxylate, and excess glycolate would be excreted through the urine. The idea of treating PH1 by inhibition of HAO1 is further supported that some individuals with abnormal splice variants of HAO1 are asymptomatic for glycolic aciduria, whereby there was
30 increased urinary glycolic acid excretion without apparent kidney pathology. Thus, inhibition of HAO1 expression would block production of glyoxylate, and in turn block production of its

metabolite, oxalate. Accordingly, the gene editing systems disclosed here, targeting the HAO1 gene, could be used to treat primary hyperoxaluria (PH) in a subject in need of the treatment.

In some embodiments, the RNA guide is comprised of a direct repeat component and a spacer component. In some embodiments, the RNA guide binds a Cas12i polypeptide. In some
5 embodiments, the spacer component is specific to an HAO1 target sequence, wherein the HAO1 target sequence is adjacent to a 5'-NTTN-3' or 5'-TTN-3' PAM sequence as described herein. In the case of a double-stranded target, the RNA guide binds to a first strand of the target (*i.e.*, the non-PAM strand) and a PAM sequence as described herein is present in the second, complementary strand (*i.e.*, the PAM strand).

10 In some embodiments, the present disclosure provides compositions comprising a complex, wherein the complex comprises an RNA guide targeting HAO1. In some embodiments, the present disclosure comprises a complex comprising an RNA guide and a Cas12i polypeptide. In some embodiments, the RNA guide and the Cas12i polypeptide bind to each other in a molar ratio of about 1:1. In some embodiments, a complex comprising an RNA guide and a Cas12i
15 polypeptide binds to the complementary region of a target sequence within an HAO1 gene. In some embodiments, a complex comprising an RNA guide targeting HAO1 and a Cas12i polypeptide binds to the complementary region of a target sequence within an HAO1 gene at a molar ratio of about 1:1. In some embodiments, the complex comprises enzymatic activity, such as nuclease activity, that can cleave the HAO1 target sequence and/or the complementary
20 sequence. The RNA guide, the Cas12i polypeptide, and the complementary region of the HAO1 target sequence, either alone or together, do not naturally occur. In some embodiments, the RNA guide in the complex comprises a direct repeat and/or a spacer sequence described herein. In some embodiments, the sequence of the RNA guide has at least 90% identity (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to a sequence of any one of SEQ
25 ID NOs: 967-1023. In some embodiments, the RNA guide has a sequence of any one of SEQ ID NOs: 967-1023.

In some embodiments, the present disclosure described herein comprises compositions comprising an RNA guide as described herein and/or an RNA encoding a Cas12i polypeptide as described herein. In some embodiments, the RNA guide and the RNA encoding a Cas12i
30 polypeptide are comprised together within the same composition. In some embodiments, the RNA guide and the RNA encoding a Cas12i polypeptide are comprised within separate compositions. In some embodiments, the RNA guide comprises a direct repeat and/or a spacer

sequence described herein. In some embodiments, the sequence of the RNA guide has at least 90% identity (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to a sequence of any one of SEQ ID NOs: 967-1023. In some embodiments, the RNA guide has a sequence of any one of SEQ ID NOs: 967-1023.

5 Use of the gene editing systems disclosed herein has advantages over those of other known nuclease systems. Cas12i polypeptides are smaller than other nucleases. For example, Cas12i2 is 1,054 amino acids in length, whereas *S. pyogenes* Cas9 (SpCas9) is 1,368 amino acids in length, *S. thermophilus* Cas9 (StCas9) is 1,128 amino acids in length, FxCpf1 is 1,300 amino acids in length, AsCpf1 is 1,307 amino acids in length, and LbCpf1 is 1,246 amino acids in
10 length. Cas12i RNA guides, which do not require a trans-activating CRISPR RNA (tracrRNA), are also smaller than Cas9 RNA guides. The smaller Cas12i polypeptide and RNA guide sizes are beneficial for delivery. Compositions comprising a Cas12i polypeptide also demonstrate decreased off-target activity compared to compositions comprising an SpCas9 polypeptide. *See* PCT/US2021/025257, which is incorporated by reference in its entirety. Furthermore, indels
15 induced by compositions comprising a Cas12i polypeptide differ from indels induced by compositions comprising an SpCas9 polypeptide. For example, SpCas9 polypeptides primarily induce insertions and deletions of 1 nucleotide in length. However, Cas12i polypeptides induce larger deletions, which can be beneficial in disrupting a larger portion of a gene such as HAO1.

 Also provided herein is a system for genetic editing of a hydroxyacid oxidase 1 (HAO1)
20 gene, which comprises (i) a Cas12i polypeptide (*e.g.*, a Cas12i2 polypeptide) or a first nucleic acid encoding the Cas12i polypeptide (*e.g.*, a Cas12i2 polypeptide comprises an amino acid sequence at least 95% identical to SEQ ID NO: 922, which may and comprises one or more mutations relative to SEQ ID NO: 922); and (ii) an RNA guide or a second nucleic acid encoding
25 the RNA guide, wherein the RNA guide comprises a spacer sequence specific to a target sequence within an HAO1 gene (*e.g.*, within exon 1 or exon 2 of the HAO1 gene), the target sequence being adjacent to a protospacer adjacent motif (PAM) comprising the motif of 5'-TTN-3' (5'-NTTN-3'), which is located 5' to the target sequence.

A. RNA Guides

30 In some embodiments, the gene editing system described herein comprises an RNA guide targeting a HAO1 gene, for example, targeting exon 1 or exon 2 of the HAO1 gene. In some

embodiments, the gene editing system described herein may comprise two or more (*e.g.*, 2, 3, 4, 5, 6, 7, 8, 9, or more) RNA guides targeting HAO1.

The RNA guide may direct the Cas12i polypeptide contained in the gene editing system as described herein to an HAO1 target sequence. Two or more RNA guides may direct two or
5 more separate Cas12i polypeptides (*e.g.*, Cas12i polypeptides having the same or different sequence) as described herein to two or more (*e.g.*, 2, 3, 4, 5, 6, 7, 8, 9, or more) HAO1 target sequences.

Those skilled in the art reading the below examples of particular kinds of RNA guides will understand that, in some embodiments, an RNA guide is HAO1 target-specific. That is, in
10 some embodiments, an RNA guide binds specifically to one or more HAO1 target sequences (*e.g.*, within a cell) and not to non-targeted sequences (*e.g.*, non-specific DNA or random sequences within the same cell).

In some embodiments, the RNA guide comprises a spacer sequence followed by a direct repeat sequence, referring to the sequences in the 5' to 3' direction. In some embodiments, the
15 RNA guide comprises a first direct repeat sequence followed by a spacer sequence and a second direct repeat sequence, referring to the sequences in the 5' to 3' direction. In some embodiments, the first and second direct repeats of such an RNA guide are identical. In some embodiments, the first and second direct repeats of such an RNA guide are different.

In some embodiments, the spacer sequence and the direct repeat sequence(s) of the RNA
20 guide are present within the same RNA molecule. In some embodiments, the spacer and direct repeat sequences are linked directly to one another. In some embodiments, a short linker is present between the spacer and direct repeat sequences, *e.g.*, an RNA linker of 1, 2, or 3 nucleotides in length. In some embodiments, the spacer sequence and the direct repeat sequence(s) of the RNA guide are present in separate molecules, which are joined to one another
25 by base pairing interactions.

Additional information regarding exemplary direct repeat and spacer components of RNA guides is provided as follows.

(i). Direct Repeat

30 In some embodiments, the RNA guide comprises a direct repeat sequence. In some embodiments, the direct repeat sequence of the RNA guide has a length of between 12-100, 13-

75, 14-50, or 15-40 nucleotides (*e.g.*, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, or 40 nucleotides).

In some embodiments, the direct repeat sequence is a sequence of **Table 1** or a portion of a sequence of **Table 1**. The direct repeat sequence can comprise nucleotide 1 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can comprise nucleotide 2 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can comprise nucleotide 3 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can comprise nucleotide 4 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can comprise nucleotide 5 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can comprise nucleotide 6 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can comprise nucleotide 7 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can comprise nucleotide 8 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can comprise nucleotide 9 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can comprise nucleotide 10 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can comprise nucleotide 11 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can comprise nucleotide 12 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can comprise nucleotide 13 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can comprise nucleotide 14 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8.

The direct repeat sequence can comprise nucleotide 1 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can comprise nucleotide 2 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can comprise nucleotide 3 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can comprise nucleotide 4 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can comprise nucleotide 5 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can comprise nucleotide 6 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can comprise nucleotide 7 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can comprise nucleotide 8 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can comprise nucleotide 9 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can comprise nucleotide 10 through nucleotide 34 of SEQ ID

NO: 9. The direct repeat sequence can comprise nucleotide 11 through nucleotide 34 of SEQ ID
 NO: 9. The direct repeat sequence can comprise nucleotide 12 through nucleotide 34 of SEQ ID
 NO: 9. In some embodiments, the direct repeat sequence is set forth in SEQ ID NO: 10. In some
 embodiments, the direct repeat sequence comprises a portion of the sequence set forth in SEQ ID
 5 NO: 10.

In some embodiments, the direct repeat sequence has at least 90% identity (*e.g.*, at least
 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity) to a sequence of **Table 1** or
 a portion of a sequence of **Table 1**. The direct repeat sequence can have at least 90% identity to a
 sequence comprising nucleotide 1 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5,
 10 6, 7, or 8. The direct repeat sequence can have at least 90% identity to a sequence comprising 2
 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat
 sequence can have at least 90% identity to a sequence comprising 3 through nucleotide 36 of any
 one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can have at least 90%
 identity to a sequence comprising 4 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4,
 15 5, 6, 7, or 8. The direct repeat sequence can have at least 90% identity to a sequence comprising
 5 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat
 sequence can have at least 90% identity to a sequence comprising 6 through nucleotide 36 of any
 one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can have at least 90%
 identity to a sequence comprising 7 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4,
 20 5, 6, 7, or 8. The direct repeat sequence can have at least 90% identity to a sequence comprising
 8 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat
 sequence can have at least 90% identity to a sequence comprising 9 through nucleotide 36 of any
 one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can have at least 90%
 identity to a sequence comprising 10 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3,
 25 4, 5, 6, 7, or 8. The direct repeat sequence can have at least 90% identity to a sequence
 comprising 11 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The
 direct repeat sequence can have at least 90% identity to a sequence comprising 12 through
 nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can
 have at least 90% identity to a sequence comprising 13 through nucleotide 36 of any one of SEQ
 30 ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8. The direct repeat sequence can have at least 90% identity to a
 sequence comprising 14 through nucleotide 36 of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or
 8. The direct repeat sequence can have at least 90% identity to a sequence comprising 1 through

nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can have at least 90% identity to a sequence comprising 2 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can have at least 90% identity to a sequence comprising 3 through nucleotide 34 of SEQ ID NO: 9.

5 The direct repeat sequence can have at least 90% identity to a sequence comprising 4 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can have at least 90% identity to a sequence comprising 5 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can have at least 90% identity to a sequence comprising 6 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can have at least 90% identity to a sequence comprising 7 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can have at least 90% identity to a sequence comprising 8 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can have at least 90% identity to a sequence comprising 9 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can have at least 90% identity to a sequence comprising 10 through nucleotide 34 of SEQ ID NO: 9. The direct repeat sequence can have at least 90% identity to a sequence comprising 11 through nucleotide 34 of SEQ ID NO: 9.

15 The direct repeat sequence can have at least 90% identity to a sequence comprising 12 through nucleotide 34 of SEQ ID NO: 9. In some embodiments, the direct repeat sequence has at least 90% identity (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity) to SEQ ID NO: 10. In some embodiments, the direct repeat sequence has at least 90% identity to a portion of the sequence set forth in SEQ ID NO: 10.

20 In some embodiments, compositions comprising a Cas12i2 polypeptide and an RNA guide comprising the direct repeat of SEQ ID NO: 10 and a spacer length of 20 nucleotides are capable of introducing indels into an HAO1 target sequence. *See, e.g.*, Example 1, where indels were measured at forty-four HAO1 target sequences following delivery of an RNA guide and a Cas12i2 polypeptide of SEQ ID NO: 924 to HEK293T cells by RNP; Example 2, where indels were measured at eleven HAO1 target sequences following delivery of an RNA guide and a Cas12i2 polypeptide of SEQ ID NO: 924 to HepG2 cells by RNP; and Example 3, where indels were measured at five HAO1 target sequences following delivery of an RNA guide and a Cas12i2 polypeptide of SEQ ID NO: 924 to primary hepatocytes by RNP.

30 In some embodiments, the direct repeat sequence is at least 90% identical to the reverse complement of any one of SEQ ID NOs: 1-10 (*see*, **Table 1**). In some embodiments, the direct repeat sequence is the reverse complement of any one of SEQ ID NOs: 1-10.

Table 1. Cas12i2 Direct Repeat Sequences

Sequence identifier	Direct Repeat Sequence
SEQ ID NO: 1	GUUGCAAAACCCAAGAAAUCCGUCUUUCAUUGACGG
SEQ ID NO: 2	AAUAGCGGCCCUAAGAAAUCCGUCUUUCAUUGACGG
SEQ ID NO: 3	AUUGGAACUGGCGAGAAAUCCGUCUUUCAUUGACGG
SEQ ID NO: 4	CCAGCAACACCUAAGAAAUCCGUCUUUCAUUGACGG
SEQ ID NO: 5	CGGCGCUCGAAUAGGAAAUCCGUCUUUCAUUGACGG
SEQ ID NO: 6	GUGGCAACACCUAAGAAAUCCGUCUUUCAUUGACGG
SEQ ID NO: 7	GUUGCAACACCUAAGAAAUCCGUCUUUCAUUGACGG
SEQ ID NO: 8	GUUGCAAUGCCUAAGAAAUCCGUCUUUCAUUGACGG
SEQ ID NO: 9	GCAACACCUAAGAAAUCCGUCUUUCAUUGACGGG
SEQ ID NO: 10	AGAAAUCCGUCUUUCAUUGACGG

In some embodiments, the direct repeat sequence is a sequence of **Table 2** or a portion of a sequence of **Table 2**. The direct repeat sequence can comprise nucleotide 1 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can comprise nucleotide 2 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can comprise nucleotide 3 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can comprise nucleotide 4 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can comprise nucleotide 5 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can comprise nucleotide 6 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can comprise nucleotide 7 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can comprise nucleotide 8 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can comprise nucleotide 9 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can comprise nucleotide 10 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can comprise nucleotide 11 through nucleotide 36

of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can comprise nucleotide 12 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can comprise nucleotide 13
 5 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can comprise nucleotide 14 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953.

In some embodiments, the direct repeat sequence has at least 95% identity (*e.g.*, at least
 10 95%, 96%, 97%, 98% or 99% identity) to a sequence of **Table 2** or a portion of a sequence of **Table 2**. The direct repeat sequence can have at least 95% identity to a sequence comprising nucleotide 1 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 95% identity to a sequence comprising 2 through nucleotide 36 of any one of SEQ
 15 ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 95% identity to a sequence comprising 3 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 95% identity to a sequence comprising 4 through nucleotide 36 of any one of SEQ ID NOs: 936, 937,
 20 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 95% identity to a sequence comprising 5 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 95% identity to a sequence comprising 6 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939,
 25 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 95% identity to a sequence comprising 7 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 95% identity to a sequence comprising 8 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941,
 30 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 95% identity to a sequence comprising 9 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or

953. The direct repeat sequence can have at least 95% identity to a sequence comprising 10 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 95% identity to a sequence comprising 11 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 95% identity to a sequence comprising 12 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 95% identity to a sequence comprising 13 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953.

In some embodiments, the direct repeat sequence has at least 90% identity (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity) to a sequence of **Table 2** or a portion of a sequence of **Table 2**. The direct repeat sequence can have at least 90% identity to a sequence comprising nucleotide 1 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 90% identity to a sequence comprising 2 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 90% identity to a sequence comprising 3 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 90% identity to a sequence comprising 4 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 90% identity to a sequence comprising 5 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 90% identity to a sequence comprising 6 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 90% identity to a sequence comprising 7 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 90% identity to a sequence comprising 8 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct

repeat sequence can have at least 90% identity to a sequence comprising 9 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 90% identity to a sequence comprising 10 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 90% identity to a sequence comprising 11 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 90% identity to a sequence comprising 12 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. The direct repeat sequence can have at least 90% identity to a sequence comprising 13 through nucleotide 36 of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953.

In some embodiments, the direct repeat sequence is at least 90% identical to the reverse complement of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. In some embodiments, the direct repeat sequence is at least 95% identical to the reverse complement of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953. In some embodiments, the direct repeat sequence is the reverse complement of any one of SEQ ID NOs: 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, or 953.

In some embodiments, the direct repeat sequence is at least 90% identical to SEQ ID NO: 954 or a portion of SEQ ID NO: 954. In some embodiments, the direct repeat sequence is at least 95% identical to SEQ ID NO: 954 or a portion of SEQ ID NO: 954. In some embodiments, the direct repeat sequence is 100% identical to SEQ ID NO: 954 or a portion of SEQ ID NO: 954.

25

Table 2. Cas12i4 Direct Repeat Sequences

Sequence identifier	Direct Repeat Sequence
SEQ ID NO: 936	UCUCAACGAUAGUCAGACAUGUGUCCUCAGUGACAC
SEQ ID NO: 937	UUUUAACAACACUCAGGCAUGUGUCCACAGUGACAC
SEQ ID NO: 938	UUGAACGGAUACUCAGACAUGUGUUCCAGUGACAC
SEQ ID NO: 939	UGCCCUCAAUAGUCAGAUUGUGUGUCCACAGUGACAC
SEQ ID NO: 940	UCUCAAUAGAUACUUAGAUACGUGUCCUCAGUGACAC
SEQ ID NO: 941	UCUCAAUAGAUACUCAGACAUGUGUCCACAGUGACAC
SEQ ID NO: 942	UCUCAAUAGAUACUAAGACAUGUGUCCUCAGUGACAC
SEQ ID NO: 943	UCUCAACUAUACUCAGACAUGUGUCCUCAGUGACAC
SEQ ID NO: 944	UCUCAACGAUACUCAGACAUGUGUCCUCAGUGACAC

SEQ ID NO: 945	UCUCAACGAUACUAAGAUAUGUGUCCUCAGCGACAC
SEQ ID NO: 946	UCUCAACGAUACUAAGAUAUGUGUCCCCAGUGACAC
SEQ ID NO: 947	UCUCAACGAUACUAAGAUAUGUGUCCACAGUGACAC
SEQ ID NO: 948	UCUCAACAAUACUCAGACAUGUGUCCCCAGUGACAC
SEQ ID NO: 949	UCUCAACAAUACUAAGGCAUGUGUCCCCAGUGACCC
SEQ ID NO: 950	UCUCAAAAGAUACUCAGACACGUGUCCCCAGUGACAC
SEQ ID NO: 951	UCUCAAAAAUACUCAGACAUGUGUCCUCAGUGACAC
SEQ ID NO: 952	GCGAAACAACAGUCAGACAUGUGUCCCCAGUGACAC
SEQ ID NO: 953	CCUCAACGAUAUUAAGACAUGUGUCCGCAGUGACAC
SEQ ID NO: 954	AGACAUGUGUCCUCAGUGACAC

In some embodiments, the direct repeat sequence is a sequence of **Table 3** or a portion of a sequence of **Table 3**. In some embodiments, the direct repeat sequence has at least 95% identity (*e.g.*, at least 95%, 96%, 97%, 98% or 99% identity) to a sequence of **Table 3** or a portion of a sequence of **Table 3**. In some embodiments, the direct repeat sequence has at least 90% identity (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity) to a sequence of **Table 3** or a portion of a sequence of **Table 3**. In some embodiments, the direct repeat sequence is at least 90% identical to the reverse complement of any one of SEQ ID NOs: 959-961. In some embodiments, the direct repeat sequence is at least 95% identical to the reverse complement of any one of SEQ ID NOs: 959-961. In some embodiments, the direct repeat sequence is the reverse complement of any one of SEQ ID NOs: 959-961.

Table 3. Cas12i1 Direct Repeat Sequences

Sequence identifier	Direct Repeat Sequence
SEQ ID NO: 959	GUUGGAAUGACUAAUUUUUGUGCCCACCGUUGGCAC
SEQ ID NO: 960	AAUUUUUGUGCCCAUCGUUGGCAC
SEQ ID NO: 961	AUUUUUGUGCCCAUCGUUGGCAC

In some embodiments, the direct repeat sequence is a sequence of **Table 4** or a portion of a sequence of **Table 4**. In some embodiments, the direct repeat sequence has at least 95% identity (*e.g.*, at least 95%, 96%, 97%, 98% or 99% identity) to a sequence of **Table 4** or a portion of a sequence of **Table 4**. In some embodiments, the direct repeat sequence has at least 90% identity (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity) to a sequence of **Table 4** or a portion of a sequence of **Table 4**. In some embodiments, the direct repeat sequence is at least 90% identical to the reverse complement of any one of SEQ ID NOs: 962-964. In some embodiments, the direct repeat sequence is at least 95% identical to the reverse complement of any one of SEQ ID NOs: 962-964. In some embodiments, the direct repeat sequence is the reverse complement of any one of SEQ ID NOs: 962-964.

Table 4. Cas12i3 Direct Repeat Sequences

Sequence identifier	Direct Repeat Sequence
SEQ ID NO: 962	CUAGCAAUGACCUGAAUAGUGUGUCCUUAGUUGACAU
SEQ ID NO: 963	CCUACAAUACCUAAGAAAUCCGUCCUAAGUUGACGG
SEQ ID NO: 964	AUAGUGUGUCCUUAGUUGACAU

In some embodiments, a direct repeat sequence described herein comprises an uracil (U).

- 5 In some embodiments, a direct repeat sequence described herein comprises a thymine (T). In some embodiments, a direct repeat sequence according to **Tables 1-4** comprises a sequence comprising a thymine in one or more places indicated as uracil in **Tables 1-4**.

(ii). Spacer Sequence

- 10 In some embodiments, the RNA guide comprises a DNA targeting or spacer sequence. In some embodiments, the spacer sequence of the RNA guide has a length of between 12-100, 13-75, 14-50, or 15-30 nucleotides (*e.g.*, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides) and is complementary to a non-PAM strand sequence. In some embodiments, the spacer sequence is designed to be complementary to a specific DNA strand, *e.g.*, of a genomic locus.

- 15 In some embodiments, the RNA guide spacer sequence is substantially identical to a complementary strand of a target sequence. In some embodiments, the RNA guide comprises a sequence (*e.g.*, a spacer sequence) having at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or at least about 99.5% sequence identity to a complementary strand of a reference nucleic acid sequence, *e.g.*, a target sequence. The percent identity between two such nucleic acids can be determined manually by inspection of the two optimally aligned nucleic acid sequences or by using software programs or algorithms (*e.g.*, BLAST, ALIGN, CLUSTAL) using standard parameters.

- 25 In some embodiments, the RNA guide comprises a spacer sequence that has a length of between 12-100, 13-75, 14-50, or 15-30 nucleotides (*e.g.*, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides) and at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% complementary to a region on the non-PAM strand that is complementary to the target sequence. In some embodiments, the RNA guide comprises a

sequence at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% complementary to a target DNA sequence. In some embodiments, the RNA guide comprises a sequence at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% complementary to a target genomic sequence. In some embodiments, the RNA guide comprises a sequence, *e.g.*, RNA sequence, that is a length of up to 50 and at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% complementary to a region on the non-PAM strand that is complementary to the target sequence. In some embodiments, the RNA guide comprises a sequence at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% complementary to a target DNA sequence. In some embodiments, the RNA guide comprises a sequence at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% complementary to a target genomic sequence.

In some embodiments, the spacer sequence is a sequence of **Table 5** or a portion of a sequence of **Table 5**. It should be understood that an indication of SEQ ID NOs: 466-920 should be considered as equivalent to a listing of SEQ ID NOs: 466-920, with each of the intervening numbers present in the listing, *i.e.*, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781,

782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 5 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, and 920.

The spacer sequence can comprise nucleotide 1 through nucleotide 16 of any one of SEQ
10 ID NOs: 466-920. The spacer sequence can comprise nucleotide 1 through nucleotide 17 of any
one of SEQ ID NOs: 466-920. The spacer sequence can comprise nucleotide 1 through
nucleotide 18 of any one of SEQ ID NOs: 466-920. The spacer sequence can comprise
nucleotide 1 through nucleotide 19 of any one of SEQ ID NOs: 466-920. The spacer sequence
can comprise nucleotide 1 through nucleotide 20 of any one of SEQ ID NOs: 466-920. The
15 spacer sequence can comprise nucleotide 1 through nucleotide 21 of any one of SEQ ID NOs:
466-920. The spacer sequence can comprise nucleotide 1 through nucleotide 22 of any one of
SEQ ID NOs: 466-920. The spacer sequence can comprise nucleotide 1 through nucleotide 23 of
any one of SEQ ID NOs: 466-920. The spacer sequence can comprise nucleotide 1 through
nucleotide 24 of any one of SEQ ID NOs: 466-920. The spacer sequence can comprise
20 nucleotide 1 through nucleotide 25 of any one of SEQ ID NOs: 466-920. The spacer sequence
can comprise nucleotide 1 through nucleotide 26 of any one of SEQ ID NOs: 466-920. The
spacer sequence can comprise nucleotide 1 through nucleotide 27 of any one of SEQ ID NOs:
466-920. The spacer sequence can comprise nucleotide 1 through nucleotide 28 of any one of
SEQ ID NOs: 466-920. The spacer sequence can comprise nucleotide 1 through nucleotide 29 of
25 any one of SEQ ID NOs: 466-920. The spacer sequence can comprise nucleotide 1 through
nucleotide 30 of any one of SEQ ID NOs: 466-920.

In some embodiments, the spacer sequence has at least 90% identity (*e.g.*, at least 90%,
91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity) to a sequence of **Table 5** or a
portion of a sequence of **Table 5**. The spacer sequence can have at least 90% identity to a
30 sequence comprising nucleotide 1 through nucleotide 16 of any one of SEQ ID NOs: 466-920.
The spacer sequence can have at least 90% identity to a sequence comprising nucleotide 1
through nucleotide 17 of any one of SEQ ID NOs: 466-920. The spacer sequence can have at

least 90% identity to a sequence comprising nucleotide 1 through nucleotide 18 of any one of SEQ ID NOs: 466-920. The spacer sequence can have at least 90% identity to a sequence comprising nucleotide 1 through nucleotide 19 of any one of SEQ ID NOs: 466-920. The spacer sequence can have at least 90% identity to a sequence comprising nucleotide 1 through nucleotide 20 of any one of SEQ ID NOs: 466-920. The spacer sequence can have at least 90% identity to a sequence comprising nucleotide 1 through nucleotide 21 of any one of SEQ ID NOs: 466-920. The spacer sequence can have at least 90% identity to a sequence comprising nucleotide 1 through nucleotide 22 of any one of SEQ ID NOs: 466-920. The spacer sequence can have at least 90% identity to a sequence comprising nucleotide 1 through nucleotide 23 of any one of SEQ ID NOs: 466-920. The spacer sequence can have at least 90% identity to a sequence comprising nucleotide 1 through nucleotide 24 of any one of SEQ ID NOs: 466-920.

The spacer sequence can have at least 90% identity to a sequence comprising nucleotide 1 through nucleotide 25 of any one of SEQ ID NOs: 466-920. The spacer sequence can have at least 90% identity to a sequence comprising nucleotide 1 through nucleotide 26 of any one of SEQ ID NOs: 466-920. The spacer sequence can have at least 90% identity to a sequence comprising nucleotide 1 through nucleotide 27 of any one of SEQ ID NOs: 466-920. The spacer sequence can have at least 90% identity to a sequence comprising nucleotide 1 through nucleotide 28 of any one of SEQ ID NOs: 466-920. The spacer sequence can have at least 90% identity to a sequence comprising nucleotide 1 through nucleotide 29 of any one of SEQ ID NOs: 466-920. The spacer sequence can have at least 90% identity to a sequence comprising nucleotide 1 through nucleotide 30 of any one of 466-920.

Table 5. Target and Spacer Sequences

HAO1	Strand	PAM*	SEQ ID NO	Target Sequence	SEQ ID NO	Spacer Sequence
HAO1_exon1	+	CTTA	11	CCTGGAAAAATGCTGCAA TATTATCAGCCAA	466	CCUGGAAAAAUGCUGCAAUA UUAUCAGCCAA
HAO1_exon1	+	ATTT	12	TCTTACCTGGAAAATGCTG TGCAATATTATCA	467	UCUUACCUGGAAAAAUGCUG CAAUAUUAUCA
HAO1_exon1	+	TTTT	13	CTTACCTGGAAAATGCTG GCAATATTATCAG	468	CUUACCUGGAAAAAUGCUGC AAUAUUAUCAG
HAO1_exon1	+	TTTC	14	TTACCTGGAAAATGCTG CAATATTATCAGC	469	UUACCUGGAAAAAUGCUGCA AUUAUUAUCAGC
HAO1_exon1	+	ATTA	15	TCAGCCAAAGTTTCTTC ATCATTTGCCCCA	470	UCAGCCAAAGUUUCUUCAU CAUUUGCCCCA
HAO1_exon1	+	GTTT	16	CTTCATCATTTGCCCCA GACCTGTAATAGT	471	CUUCAUCAUUUGCCCCAGA CCUGUAAUAGU
HAO1_exon1	+	TTTC	17	TTCATCATTTGCCCCAG ACCTGTAATAGTC	472	UUCAUCAUUUGCCCCAGAC CUGUAAUAGUC

HA01_exon1	+	CTTC	18	ATCATTGCCCCAGACC TGTAATAGTCATA	473	AUCAUUUGCCCCAGACCUG UAAUAGUCAUA
HA01_exon1	+	ATTT	19	GCCCCAGACCTGTAATA GTCATATATAGAC	474	GCCCCAGACCUGUAAUAGU CAUUAUAGAC
HA01_exon1	+	TTTG	20	CCCCAGACCTGTAATAG TCATATATAGACT	475	CCCCAGACCUGUAAUAGUC AUUAUAGACU
HA01_exon1	+	TTTT	21	AAAAAATAAATTTTCTT ACCTGGAAAATGC	476	AAAAAUAAAAUUUCUAC CUGGAAAAUGC
HA01_exon1	+	CTTT	22	GGAAGTACTGATTTAGC ATGTTGTTTCATAA	477	GGAAGUACUGAUUUAGCAU GUUGUUCAUAA
HA01_exon1	+	ATTT	23	AGCATGTTGTTTCATAAT CATTGATACAAAT	478	AGCAUGUUGUUCAUAAUCA UUGAUACAAAU
HA01_exon1	+	TTTA	24	GCATGTTGTTTCATAATC ATTGATACAAATT	479	GCAUGUUGUUCAUAAUCAU UGAUACAAAUU
HA01_exon1	+	GTTG	25	TTCATAATCATTGATAC AAATTAGCCGGGG	480	UUCAUUAUCAUUGAUACAA AUUAGCCGGGG
HA01_exon1	+	GTTT	26	ATAATCATTGATACAAA TTAGCCGGGGGAG	481	AUAUACAUUGAUACAAAUU AGCCGGGGGAG
HA01_exon1	+	ATTG	27	ATACAAATTAGCCGGGG GAGCATTTTCACA	482	AUACAAAUUAGCCGGGGGA GCAUUUUCACA
HA01_exon1	+	ATTA	28	GCCGGGGGAGCATTTTC ACAGGTTATTGCT	483	GCCGGGGGAGCAUUUUCAC AGGUUAUUGCU
HA01_exon1	+	ATTT	29	TCACAGGTTATTGCTAT CCCAGATGGAGTT	484	UCACAGGUUAUUGCUAUCC CAGAUUGGAGUU
HA01_exon1	+	TTTT	30	CACAGGTTATTGCTATC CCAGATGGAGTTC	485	CACAGGUUAUUGCUAUCCC AGAUGGAGUUC
HA01_exon1	+	TTTC	31	ACAGGTTATTGCTATCC CAGATGGAGTTTCG	486	ACAGGUUAUUGCUAUCCCA GAUGGAGUUCG
HA01_exon1	+	TTTG	32	GAAGTACTGATTTAGCA TGTTGTTTCATAAT	487	GAAGUACUGAUUUAGCAUG UUGUUCAUAAU
HA01_exon1	+	ATTT	33	TAAAAATAAATTTTCT TACCTGGAAAATG	488	UAAAAAUAAAAUUUCUUA CCUGGAAAAUG
HA01_exon1	+	TTTA	34	AAAAATAAATTTTCTTA CCTGGAAAATGCT	489	AAAAUAAAAUUUCUUAACC UGGAAAAUGCU
HA01_exon1	+	TTTT	35	AAAACATGATTTTAAAA AATAAATTTTCTT	490	AAAACAUGAUUUUAAAAAA UAAAUUUUCUU
HA01_exon1	-	TTTG	36	TATCAATGATTATGAAC AACATGCTAAATC	491	UAUCAUUGAUUAUGACAA CAUGC UAAAUUC
HA01_exon1	-	ATTA	37	TGAACAACATGCTAAAT CAGTACTTCCAAA	492	UGAACAACAUUGC UAAAUCA GUACUCCAAA
HA01_exon1	-	CTTC	38	CAAAGTCTATATATGAC TATTACAGGTCTG	493	CAAAGUCUAUAUUGACUA UUACAGGUCUG
HA01_exon1	-	ATTA	39	CAGGTCTGGGGCAAATG ATGAAGAAACTTT	494	CAGGUCUGGGGCAAUGAU GAAGAAACUUU
HA01_exon1	-	CTTT	40	GGCTGATAATATTGCAG CATTTTCCAGGTA	495	GGCUGAUAAUUAUUGCAGCA UUUUCAGGUA
HA01_exon1	-	TTTG	41	GCTGATAATATTGCAGC ATTTTCCAGGTAA	496	GCUGAUAAUUAUUGCAGCAU UUUCCAGGUAA
HA01_exon1	-	ATTG	42	CAGCATTTTCCAGGTAA GAAAATTTATTTT	497	CAGCAUUUUCAGGUAAAGA AAAUUUUUUUU
HA01_exon1	-	ATTT	43	TCCAGGTAAGAAAATTT ATTTTTTAAATC	498	UCCAGGUAAAGAAAUUUAU UUUUUAAAAUC
HA01_exon1	-	TTTT	44	CCAGGTAAGAAAATTTA TTTTTTAAATCA	499	CCAGGUAAAGAAAUUUAUU UUUUAAAAUCA
HA01_exon1	+	TTTA	45	AAACATGATTTTAAAAA ATAAATTTTCTTA	500	AAACAUGAUUUUAAAAAAU AAAUUUUCUUA

HA01_exon1	-	ATTT	46	ATTTTTTAAATCATGT TTTAAATTACAC	501	AUUUUUUAAAAUCAUGUUU UAAAAUUACAC
HA01_exon1	-	TTTC	47	CAGGTAAGAAAATTTAT TTTTTAAATCAT	502	CAGGUAAGAAAAUUUUAUUU UUUAAAAUCAU
HA01_exon1	-	ATTT	48	TTTAAATCATGTTTAA AAATTACACAAAG	503	UUUAAAAUCAUGUUUUAAA AUUACACAAAG
HA01_exon1	-	TTTT	49	TTAAATCATGTTTAA AATTACACAAAGA	504	UUAAAAUCAUGUUUUAAAA UUACACAAAGA
HA01_exon1	-	TTTT	50	TAAATCATGTTTTAAA ATTACACAAAGAC	505	UAAAAUCAUGUUUUAAAAU UACACAAAGAC
HA01_exon1	-	TTTT	51	AAAATCATGTTTTAAA TTACACAAAGACC	506	AAAAUCAUGUUUUAAAAUU ACACAAAGACC
HA01_exon1	-	TTTA	52	AAATCATGTTTTAAAA TACACAAAGACCG	507	AAAUCAUGUUUUAAAAUUA CACAAAGACCG
HA01_exon1	+	CTTT	53	GTGTAATTTTAAACAT GATTTTAAAAAAT	508	GUGUAAUUUUAAAAACAUGA UUUUAAAAAAU
HA01_exon1	+	TTTG	54	TGTAATTTTAAACATG ATTTTAAAAAATA	509	UGUAAUUUUAAAAACAUGAU UUUAAAAAAUA
HA01_exon1	+	ATTT	55	TAAACATGATTTTAAA AAATAAATTTTCT	510	UAAACAUGAUUUUUAAAAA AUAAAUUUUCU
HA01_exon1	-	TTTA	56	TTTTTAAATCATGTT TTAAATTACACA	511	UUUUUUAAAAUCAUGUUUU AAAAUUACACA
HA01_exon1	-	ATTT	57	GTATCAATGATTATGAA CAACATGCTAAAT	512	GUUCAAUUAUGAACA ACAUGCUAAAU
HA01_exon1	+	GTTA	58	TTGCTATCCAGATGGA GTTCTGTT	513	UUGCUAUCCCAGAUUGGAGU UCGUU
HA01_exon1	+	ATTG	59	CTATCCAGATGGAGTT CGTT	514	CUAUCCCAGAUUGGAGUUCG UU
HA01_exon2	-	TTTA	60	TTTTTTAATTCTAGATG GAAGCTGTATCCA	515	UUUUUUAAUUCUAGAUGGA AGCUGUAUCCA
HA01_exon2	-	TTTT	61	ATTTTATTTTTTAATTC TAGATGGAAGCTG	516	AUUUUAAUUUUUAAUUCUA GAUGGAAGCUG
HA01_exon2	-	TTTT	62	ATTTTTAATTCTAGAT GGAAGCTGTATCC	517	AUUUUUAAUUCUAGAUGG AAGCUGUAUCC
HA01_exon2	-	ATTT	63	TATTTTTTAATTCTAGA TGGAAGCTGTATC	518	UAUUUUUAAUUCUAGAUG GAAGCUGUAUC
HA01_exon2	-	TTTA	64	TTTTTATTTTTAATTCT AGATGGAAGCTGT	519	UUUUAAUUUUUAAUUCUAG AUGGAAGCUGU
HA01_exon2	+	ATTA	65	AAAAATAAAATAAAATA AAAGGCTTTAGAG	520	AAAAUAAAAUAAAAUAAA AGGCUUUAGAG
HA01_exon2	-	TTTT	66	ATTTTATTTTATTTTTT AATTCTAGATGGA	521	AUUUUAAUUUUUAAUUCUAA UUCUAGAUGGA
HA01_exon2	-	CTTT	67	TATTTTATTTTATTTTT TAATTCTAGATGG	522	UAUUUUAAUUUUUAAUUCUAA AUUCUAGAUGG
HA01_exon2	-	ATTC	68	TGAAACTCTAAAGCCTT TTATTTTATTTTA	523	UGAAACUCUAAAGCCUUUU AUUUUAAUUUA
HA01_exon2	-	ATTT	69	TTTAATTCAGATGGAA GCTGTATCCAAGG	524	UUUAAUUCUAGAUGGAAGC UGUAUCCAAGG
HA01_exon2	-	TTTA	70	TTTTATTTTATTTTTTA ATTCTAGATGGAA	525	UUUUAAUUUUUAAUUCUAAU UCUAGAUGGAA
HA01_exon2	-	TTTT	71	TTAATTCAGATGGAAG CTGTATCCAAGGA	526	UUAAUUCUAGAUGGAAGCU GUAUCCAAGGA
HA01_exon2	-	ATTT	72	TATTTTATTTTAAATT CTAGATGGAAGCT	527	UAUUUUAAUUUUUAAUUCU AGAUGGAAGCU
HA01_exon2	+	CTTC	73	CATCTAGAATTAAAAAA TAAATAAAATAA	528	CAUCUAGAAUUAAAAAAUA AAAUAAAAUAA

HA01_exon2	–	TTTT	74	TAATTCTAGATGGAAGC TGTATCCAAGGAT	529	UAAUUCUAGAUGGAAGCUG UAUCCAAGGAU
HA01_exon2	–	TTTT	75	AATTCTAGATGGAAGCT GTATCCAAGGATG	530	AAUUCUAGAUGGAAGCUGU AUCCAAGGAUG
HA01_exon2	–	TTTA	76	ATTCTAGATGGAAGCTG TATCCAAGGATGC	531	AUUCUAGAUGGAAGCUGUA UCCAAGGAUGC
HA01_exon2	–	ATTC	77	TAGATGGAAGCTGTATC CAAGGATGCTCCG	532	UAGAUGGAAGCUGUAUCCA AGGAUGCUCCG
HA01_exon2	–	GTTG	78	CTGAAACAGATCTGTCTG ACTTCTGTTTTAG	533	CUGAAACAGAUUCUGUCGAC UUCUGUUUUAG
HA01_exon2	–	GTTT	79	TAGGACAGAGGGTCAGC ATGCCAATATGTG	534	UAGGACAGAGGGUCAGCAU GCCAAUAUGUG
HA01_exon2	–	TTTT	80	AGGACAGAGGGTCAGCA TGCCAATATGTGT	535	AGGACAGAGGGUCAGCAUG CCAAUAUGUGU
HA01_exon2	–	TTTA	81	GGACAGAGGGTCAGCAT GCCAATATGTGTG	536	GGACAGAGGGUCAGCAUGC CAAUAUGUGUG
HA01_exon2	–	CTTG	82	CCACTGTGAGAGGTAGG AGGAAGATTGTCA	537	CCACUGUGAGAGGUAGGAG GAAGAUUGUCA
HA01_exon2	–	CTTC	83	TGTTTTAGGACAGAGGG TCAGCATGCCAAT	538	UGUUUUAGGACAGAGGGUC AGCAUGCCAAU
HA01_exon2	+	GTTA	84	GCCTCCTTCTGTCCCTG TGGTGACAATCTT	539	GCCUCCUUCUGUCCUGUG GUGACAAUCUU
HA01_exon2	–	ATTG	85	TCACCACAGGGACAGAA GGAGGCTAACGTT	540	UCACCACAGGGACAGAAGG AGGCUAACGUU
HA01_exon2	+	ATTC	86	CGGAGCATCCTTGGATA CAGCTTCCATCTA	541	CGGAGCAUCCUUGGAUACA GCUUCCAUCUA
HA01_exon2	+	TTTC	87	AGCAACATTCCGGAGCA TCCTTGGATACAG	542	AGCAACAUUCCGGAGCAUC CUUGGAUACAG
HA01_exon2	+	GTTT	88	CAGCAACATTCCGGAGC ATCCTTGGATACA	543	CAGCAACAUUCCGGAGCAU CCUUGGAUACA
HA01_exon2	+	CTTG	89	GATACAGCTTCCATCTA GAATTAATAAATA	544	GAUACAGCUUCCAUCUAGA AUUAAAAAUA
HA01_exon2	+	CTTC	90	CTCCTACCTCTCACAGT GGCAAGCTCGCCG	545	CUCCUACCUCUCACAGUGG CAAGCUCGCCG
HA01_exon2	+	CTTC	91	TGTCCCTGTGGTGACAA TCTTCTCCTACC	546	UGUCCUGUGGUGACAAUC UUCUCCUACC
HA01_exon2	+	ATTG	92	GCATGCTGACCCTCTGT CCTAAACAGAAG	547	GCAUGCUGACCCUCUGUCC UAAAACAGAAG
HA01_exon3	–	CTTA	93	CCTGGGCAACCGTCTGG ATGATGTGCGTAA	548	CCUGGGCAACCGUCUGGAU GAUGUGCGUAA
HA01_exon3	+	TTTG	94	AATCTGTTACGCACATC ATCCAGACGGTTG	549	AAUCUGUUACGCACAUCAU CCAGACGGUUG
HA01_exon3	+	GTTT	95	GAATCTGTTACGCACAT CATCCAGACGGTT	550	GAAUCUGUUACGCACAUCA UCCAGACGGUU
HA01_exon3	+	GTTG	96	TGGCGGCAGTTTGAATC TGTTACGCACATC	551	UGGCGGCAGUUUGAAUCUG UUACGCACAUC
HA01_exon3	+	GTTA	97	CCTGAGTTGTGGCGGCA GTTTGAATCTGTT	552	CCUGAGUUUGGCGGCAGU UUGAAUCUGUU
HA01_exon3	+	TTTC	98	GCCTCAGCTCGGGGCCC ACATGATCATGGT	553	GCCUCAGCUCGGGGCCCCAC AUGAUCUAGGU
HA01_exon3	+	CTTT	99	CGCCTCAGCTCGGGGCC CACATGATCATGG	554	CGCCUCAGCUCGGGGCCCCA CAUGAUCUAGG
HA01_exon3	–	ATTC	100	AAACTGCCGCCACAAC CAGGTAACCATGA	555	AAACUGCCGCCACAACUCA GGUAACCAUGA
HA01_exon3	–	TTTG	101	TGACAGTGGACACACCT TACCTGGGCAACC	556	UGACAGUGGACACACCUUA CCUGGGCAACC

HA01_exon3	–	CTTG	102	ATCATCCCTTTCTTTC TCAGCCTGTCAGT	557	AUCAUCCCUUUCUUUCUC AGCCUGUCAGU
HA01_exon3	–	GTTG	103	GCTGCAACTGTATATCT ACAAGGACCGAGA	558	GCUGCAACUGUAUAUCUAC AAGGACCGAGA
HA01_exon3	–	ATTG	104	AAGAAGTGGCGGAAGCT GGTCCTGAGGCAC	559	AAGAAGUGGCGGAAGCUGG UCCUGAGGCAC
HA01_exon3	–	GTTT	105	CTGGGCCACCTCCTCAA TTGAAGAAGTGGC	560	CUGGGCCACCUCUCAAUU GAAGAAGUGGC
HA01_exon3	–	GTTG	106	AGTTCCTGGGCCACCTC CTCAATTGAAGAA	561	AGUUCUGGGGCCACCUCCU CAAUUGAAGAA
HA01_exon3	–	TTTC	107	TCAGCCTGTCAGTCCCT GGGAACGGGCATG	562	UCAGCCUGUCAGUCCUGG GAACGGGCAUG
HA01_exon3	–	CTTT	108	CTCAGCCTGTCAGTCCC TGGGAACGGGCAT	563	CUCAGCCUGUCAGUCCUG GGAACGGGCAU
HA01_exon3	–	TTTC	109	TTTCTCAGCCTGTCAGT CCCTGGGAACGGG	564	UUUCUCAGCCUGUCAGUCC CUGGGAACGGG
HA01_exon3	–	CTTT	110	CTTTCTCAGCCTGTCAG TCCCTGGGAACGG	565	CUUUCUCAGCCUGUCAGUC CCUGGGAACGG
HA01_exon3	+	GTTA	111	CGCACATCATCCAGACG GTTGCCCAGGTAA	566	CGCACAUCAUCCAGACGGU UGCCCAGGUAA
HA01_exon3	–	ATTT	112	GTGACAGTGGACACACC TTACCTGGGCAAC	567	GUGACAGUGGACACACCUU ACCUGGGCAAC
HA01_exon3	+	GTTG	113	CCCAGGTAAGGTGTGTC CACTGTACAAAT	568	CCCAGGUAAGGUGUGUCCA CUGUCACAAAU
HA01_exon3	–	CTTC	114	GTTGGCTGCAACTGTAT ATCTACAAGGACC	569	GUUGGCUGCAACUGUAUUA CUACAAGGACC
HA01_exon3	+	CTTC	115	TCTGCCTGCCGCACTAG CTTCTTGGTGACT	570	UCUGCCUGCCGACUAGCU UCUUGGUGACU
HA01_exon3	+	CTTG	116	TAGCCCATCTTCTCTGC CTGCCGCACTAGC	571	UAGCCCAUCUUCUCUGCCU GCCGCACUAGC
HA01_exon3	+	GTTT	117	CCAGGGACTGACAGGCT GAGAAAGAAAGGG	572	CCAGGGACUGACAGGCUGA GAAAGAAAGGG
HA01_exon3	+	ATTG	118	AGGAGGTGGCCAGGAA CTCAACATCATGC	573	AGGAGGUGCCAGGAACU CAACAUCAUAGC
HA01_exon3	+	CTTC	119	TTCAATTGAGGAGGTGG CCCAGGAACCTAA	574	UUCAAUUGAGGAGGUGGCC CAGGAACUCAA
HA01_exon3	+	CTTC	120	CGCCACTTCTTCAATTG AGGAGGTGGCCCA	575	CGCCACUUCUCAAUUGAG GAGGUGGCCCA
HA01_exon3	+	CTTC	121	AATTGAGGAGGTGGCCC AGGAACCTCAACAT	576	AAUUGAGGAGGUGGCCAG GAACUCAACAU
HA01_exon3	+	CTTG	122	TAGATATACAGTTGCAG CCAACGAAGTGCC	577	UAGAUUAACAGUUGCAGCC AACGAAGUGCC
HA01_exon3	+	CTTC	123	TCGGTCCTTGTAGATAT ACAGTTGCAGCCA	578	UCGGUCCUUGUAGAUUAC AGUUGCAGCCA
HA01_exon3	+	CTTG	124	GTGACTTCTCGGTCCTT GTAGATATACAGT	579	GUGACUUCUCGGUCCUUGU AGAUUAACAGU
HA01_exon3	+	CTTC	125	TTGGTGACTTCTCGGTC CTTGTAGATATAC	580	UUGGUGACUUCUCGGUCCU UGUAGAUUAC
HA01_exon3	+	GTTG	126	CAGCCAACGAAGTGCCT CAGGACCAGCTTC	581	CAGCCAACGAAGUGCCUCA GGACCAGCUUC
HA01_exon4	–	ATTT	127	CTAATTTGGCAAATTC TCATTTTATGCAT	582	CUAAUUUGGCAAAUUUCUC AUUUUAUGCAU
HA01_exon4	+	TTTC	128	ATCCTAAAATAAGAAAT GCATAAAATGAGA	583	AUCCUAAAAUAAGAAUUGC AUAAAAUGAGA
HA01_exon4	+	ATTC	129	AAGTAGAGAAATAAACG AACCTCTCAAAAT	584	AAGUAGAGAAAUAACGAA CCUCUCAAAAU

HA01_exon4	–	TTTC	130	TCTACTTGAATTCATAC TGACTTTGTGATC	585	UCUACUUGAAUUCAUACUG ACUUUGUGAUC
HA01_exon4	–	TTTC	131	TAATTTGGCAAATTTCT CATTTTATGCATT	586	UAAUUUGGCAAAUUCUCA UUUUUAGCAUU
HA01_exon4	–	ATTT	132	TATGCATTTCTTATTTT AGGATGAAAAATT	587	UAUGCAUUUCUUAUUUAG GAUGAAAAAUU
HA01_exon4	–	TTTG	133	GCAAATTTCTCATTTTA TGCATTTCTTATT	588	GCAAAUUUCUCAUUUUUAG CAUUUCUUAUU
HA01_exon4	–	ATTT	134	CTCATTTTATGCATTTTC TTATTTTAGGATG	589	CUCAUUUUUAGCAUUUCUU AUUUUAGGAUG
HA01_exon4	–	TTTC	135	TCATTTTATGCATTTCT TATTTTAGGATGA	590	UCAUUUUUAGCAUUUCUUA UUUUAGGAUGA
HA01_exon4	+	TTTT	136	CATCCTAAAATAAGAAA TGCATAAAATGAG	591	CAUCCUAAAAUAAGAAAUG CAUAAAAUGAG
HA01_exon4	–	ATTT	137	GGCAAATTTCTCATTTT ATGCATTTCTTAT	592	GGCAAAUUUCUCAUUUUUAG GCAUUUCUUAU
HA01_exon4	+	TTTT	138	TCATCCTAAAATAAGAA ATGCATAAAATGA	593	UCAUCCUAAAAUAAGAAAUG GCAUAAAAUGA
HA01_exon4	+	ATTT	139	TCCTCAGGAGAAAAATGA TAAAGTACTGGTT	594	UCCUCAGGAGAAAAUGAUA AAGUACUGGUU
HA01_exon4	+	TTTC	140	AAAATTTTTCATCCTAA AATAAGAAATGCA	595	AAAAUUUUUCAUCCUAAAA UAAGAAAUGCA
HA01_exon4	+	GTTT	141	CAAAATTTTTCATCCTA AAATAAGAAATGC	596	CAAAUUUUUCAUCCUAAAA AUAAGAAAUGC
HA01_exon4	+	TTTC	142	CTCAGGAGAAAAATGATA AAGTACTGGTTTC	597	CUCAGGAGAAAAUGAUAUA GUACUGGUUUC
HA01_exon4	+	TTTT	143	CCTCAGGAGAAAAATGAT AAAGTACTGGTTT	598	CCUCAGGAGAAAAUGAUAUA AGUACUGGUUU
HA01_exon4	–	TTTT	144	ATGCATTTCTTATTTTA GGATGAAAAATTT	599	AUGCAUUUCUUAUUUUAGG AUGAAAAAUUU
HA01_exon4	+	TTTA	145	GCCACATATGCAGCAAG TCCACTGTCGTCT	600	GCCACAUUAGCAGCAAGUC CACUGUCGUCU
HA01_exon4	+	CTTT	146	AGCCACATATGCAGCAA GTCCACTGTCGTC	601	AGCCACAUUAGCAGCAAGU CCACUGUCCGUC
HA01_exon4	+	ATTG	147	CTTTAGCCACATATGCA GCAAGTCCACTGT	602	CUUUAGCCACAUUAGCAGC AAGUCCACUGU
HA01_exon4	+	CTTC	148	CCAGCTGATAGATGGGT CTATTGCTTTAGC	603	CCAGCUGAUAGAUGGGUCU AUUGCUUUAGC
HA01_exon4	+	TTTG	149	ATATCTTCCCAGCTGAT AGATGGGTCTATT	604	AUAUCUCCCAGCUGAUAG AUGGGUCUUAU
HA01_exon4	+	ATTT	150	GATATCTTCCCAGCTGA TAGATGGGTCTAT	605	GAUAUCUCCCAGCUGAUA GAUGGGUCUAU
HA01_exon4	+	CTTC	151	TCAGCCATTTGATATCT TCCCAGCTGATAG	606	UCAGCCAUUUUGAUUUCUUC CCAGCUGAUAG
HA01_exon4	+	ATTG	152	GCAATGATGTCAGTCTT CTCAGCCATTGGA	607	GCAAUGAUGUCAGUCUUCU CAGCCAUUUGA
HA01_exon4	+	ATTT	153	TTCATCCTAAAATAAGA AATGCATAAAATG	608	UUCAUCCUAAAAUAAGAAA UGCAUAAAAUG
HA01_exon4	–	TTTA	154	TGCATTTCTTATTTTAG GATGAAAAATTTT	609	UGCAUUUCUUAUUUUAGGA UGAAAAAUUUU
HA01_exon4	–	TTTA	155	GGATGAAAAATTTTGAA ACCAGTACTTTAT	610	GGAUGAAAAUUUUGAAAC CAGUACUUUAU
HA01_exon4	–	TTTC	156	TTATTTTAGGATGAAAA ATTTTGAAACCAG	611	UUUUUUUAGGAUGAAAAAU UUUGAAACCAG
HA01_exon4	–	ATTG	157	CCAATTGTTGCAAAGGG CATTTTGAGAGGT	612	CCAAUUGUUGCAAAGGGCA UUUUAGAGAGU

HA01_exon4	-	ATTG	158	TTGCAAAGGGCATTTTG AGAGGTTCTGTTTA	613	UUGCAAAGGGCAUUUUGAG AGGUUCGUUUA
HA01_exon4	+	CTTT	159	GCAACAATTGGCAATGA TGTCAGTCTTCTC	614	GCAACAAUUGGCAAUGAUG UCAGUCUUCUC
HA01_exon4	-	GTTG	160	CAAAGGGCATTTTGAGA GGTTCGTTTATTT	615	CAAAGGGCAUUUUGAGAGG UUCGUUUUUUU
HA01_exon4	-	ATTT	161	TGAGAGGTTCTGTTTATT TCTCTACTTGAAT	616	UGAGAGGUUCGUUUUUUUC UCUACUUGAAU
HA01_exon4	-	TTTT	162	GAGAGGTTCTGTTTATTT CTCTACTTGAATT	617	GAGAGGUUCGUUUUUUUCU CUACUUGAAUU
HA01_exon4	-	TTTG	163	AGAGGTTCTGTTTATTTCT TCTACTTGAATTC	618	AGAGGUUCGUUUUUUUCUC UACUUGAAUUC
HA01_exon4	-	GTTC	164	GTTTATTTCTCTACTTG AATTCATCTGAC	619	GUUUUUUUCUCUACUUGAA UUCUACUGAC
HA01_exon4	-	GTTT	165	ATTTCTCTACTTGAATT CATACTGACTTTG	620	AUUUCUCUACUUGAAUUCU UACUGACUUUG
HA01_exon4	-	TTTA	166	TTTCTCTACTTGAATTC ATACTGACTTTGT	621	UUUCUCUACUUGAAUUCU ACUGACUUUGU
HA01_exon4	-	ATTT	167	CTCTACTTGAATTCATA CTGACTTTGTGAT	622	CUCUACUUGAAUUCUACU GACUUUGUGAU
HA01_exon4	-	CTTG	168	CTGCATATGTGGCTAAA GCAATAGACCCAT	623	CUGCAUAUGUGGCUAAAAGC AAUAGACCCAU
HA01_exon4	-	ATTT	169	CTTATTTTAGGATGAAA AATTTTGAAACCA	624	CUUAAUUUAGGAUGAAAAA UUUUGAAACCA
HA01_exon4	-	TTTG	170	GAGACGACAGTGGACTT GCTGCATATGTGG	625	GAGACGACAGUGGACUUGC UGCAUAUGUGG
HA01_exon4	-	ATTT	171	TGGAGACGACAGTGGAC TTGCTGCATATGT	626	UGGAGACGACAGUGGACUU GCUGCAUAUGU
HA01_exon4	-	TTTC	172	TCCTGAGGAAAATTTTG GAGACGACAGTGG	627	UCCUGAGGAAAUUUUGGA GACGACAGUGG
HA01_exon4	-	TTTT	173	CTCCTGAGGAAAATTTT GGAGACGACAGTG	628	CUCCUGAGGAAAUUUUGG AGACGACAGUG
HA01_exon4	-	ATTT	174	TCTCTGAGGAAAATTT TGGAGACGACAGT	629	UCUCCUGAGGAAAUUUUG GAGACGACAGU
HA01_exon4	-	TTTA	175	TCATTTTCTCCTGAGGA AAATTTTGGAGAC	630	UCAUUUUCUCCUGAGGAAA AUUUUGGAGAC
HA01_exon4	-	CTTT	176	ATCATTTTCTCCTGAGG AAAATTTTGGAGA	631	AUCAUUUUCUCCUGAGGAA AAUUUUGGAGA
HA01_exon4	-	TTTG	177	AAACCACTACTTTTATCA TTTTCTCCTGAGG	632	AAACCACTACTTTTATCA UUCUCCUGAGG
HA01_exon4	-	TTTT	178	GAAACCACTACTTTATC ATTTTCTCCTGAG	633	GAAACCACTACTTTATC UUUCUCCUGAG
HA01_exon4	-	ATTT	179	TGAAACCACTACTTTAT CATTTTCTCCTGA	634	UGAAACCACTACTTTAT UUUUCUCCUGA
HA01_exon4	-	TTTT	180	AGGATGAAAAATTTTGA AACCAGTACTTTA	635	AGGAUGAAAAUUUUGAAA CCAGUACUUUA
HA01_exon4	-	ATTT	181	TAGGATGAAAAATTTTG AAACCACTACTTT	636	UAGGAUGAAAAUUUUGAA ACCAGUACUUU
HA01_exon4	-	CTTA	182	TTTTAGGATGAAAAATT TTGAAACCACTAC	637	UUUUAGGAUGAAAAUUUU GAAACCACTAC
HA01_exon4	-	TTTT	183	GGAGACGACAGTGGACT TGCTGCATATGTG	638	GGAGACGACAGUGGACUUG CUGCAUAUGUG
HA01_exon4	+	TTTG	184	CAACAATTGGCAATGAT GTCAGTCTTCTCA	639	CAACAATTGGCAATGAT CAGUCUUCUCA
HA01_exon4	-	CTTG	185	AATTCATACTGACTTTG TGATCCTTTGTG	640	AAUUCUACUACUUGUG AUCCUUUGUG

HA01_exon4	-	ATTC	186	ATACTGACTTTGTGATC CTTTGTG	641	AUACUGACUUUGUGAUCCU UUGUG
HA01_exon5	-	GTTA	187	AGTTACAGTTTCCCTAA GGTGCTGTGTTTA	642	AGUUACAGUUUCCCUAAGG UGCUUGUUUUA
HA01_exon5	+	ATTC	188	AAGCCATGTTTAAACAGC CTCCCTGGCATCA	643	AAGCCAUUUUAACAGCCU CCCUGGCAUCA
HA01_exon5	+	TTTA	189	ACAGCCTCCCTGGCATC ATCACCTGGAGAG	644	ACAGCCUCCUGGCAUCAU CACCUGGAGAG
HA01_exon5	+	GTTT	190	AACAGCCTCCCTGGCAT CATCACCTGGAGA	645	AACAGCCUCCUGGCAUCA UCACCUGGAGA
HA01_exon5	+	ATTC	191	GACACCAAGATCCCAT CAAGCCATGTTTA	646	GACACCAAGAUCCCAUUA AGCCAUGUUUA
HA01_exon5	+	GTTG	192	TCGAGCCCCATGATTTCG ACACCAAGATCCC	647	UCGAGCCCCAUGAUUCGAC ACCAAGAUC
HA01_exon5	+	CTTA	193	GCGTCTGCCAAAACCTCA CAGTGGCTGGCAC	648	GCGUCUGCCAAAACUCACA GUGGCUGGCAC
HA01_exon5	-	TTTG	194	GCAGACGCTAAGATTTC CTTTTGGAGTTCC	649	GCAGACGCUAAGAUUUCCU UUUGGAGUCC
HA01_exon5	-	GTTT	195	TGGCAGACGCTAAGATT TCCTTTTGGAGTT	650	UGGCAGACGCUAAGAUUUC CUUUUGGAGUU
HA01_exon5	-	CTTG	196	GTGTCCAATCATGGGGC TCGACAACTCGAT	651	GUGUCGAAUCAUGGGGUC GACAACUCGAU
HA01_exon5	-	TTTT	197	GGCAGACGCTAAGATT CCTTTTGGAGTTCC	652	GGCAGACGCUAAGAUUUC UUUUGGAGUUC
HA01_exon5	-	GTTA	198	AACATGGCTTGAATGGG ATCTTGGTGTCTGA	653	AACAUGGCUUGAUGGGGAU CUUGGUGUGCA
HA01_exon5	-	TTTA	199	CTCTCTCCAGGTGATGA TGCCAGGGAGGCT	654	CUCUCUCCAGGUGAUGAUG CCAGGGAGGCU
HA01_exon5	-	TTTT	200	ACTCTCTCCAGGTGATG ATGCCAGGGAGGC	655	ACUCUCUCCAGGUGAUGAU GCCAGGGAGGC
HA01_exon5	-	GTTT	201	TACTCTCTCCAGGTGAT GATGCCAGGGAGG	656	UACUCUCUCCAGGUGAUGA UGCCAGGGAGG
HA01_exon5	-	CTTG	202	TTTTACTCTCTCCAGGT GATGATGCCAGGG	657	UUUUACUCUCUCCAGGUGA UGAUGCCAGGG
HA01_exon5	-	TTTC	203	CCTAAGGTGCTTGTGTTT ACTCTCTCCAGGT	658	CCUAAGGUGCUUGUUUUAC UCUCUCCAGGU
HA01_exon5	-	GTTT	204	CCCTAAGGTGCTTGTGTTT TACTCTCTCCAGG	659	CCCUAAGGUGCUUGUUUUUA CUCUCUCCAGG
HA01_exon5	-	GTTA	205	CAGTTTCCCTAAGGTGC TTGTTTTACTCTC	660	CAGUUUCCCUAAGGUGCUU GUUUUACUCUC
HA01_exon5	-	CTTG	206	AATGGGATCTTGGTGTC GAATCATGGGGCT	661	AAUGGGAUCUUGGUGUGCA AUAUGGGGCU
HA01_exon5	-	ATTT	207	CCTTTTGGAGTTCCCAT TTCCATC	662	CCUUUUGGAGUCCCAUUU CCAUC
HA01_exon5	-	TTTC	208	CTTTTGGAGTTCCCAT TCCATC	663	CUUUUGGAGUCCCAUUUC CAUC
HA01_exon5	+	CTTA	209	GGGAAACTGTAACCTAA CAGGCAG	664	GGGAAACUGUAACUUAACA GGCAG
HA01_exon6	-	TTTA	210	CAACTTCTTTTCTTTT ATGATCTTTAAGT	665	CAACUUUCUUUCUUUUUAU GAUCUUUAAGU
HA01_exon6	-	ATTC	211	CGGTTGGCCATGGCTCT GAGTGGTAAGACT	666	CGGUUGGCCAUGGCUCUGA GUGGUAAGACU
HA01_exon6	-	GTTG	212	GCCATGGCTCTGAGTGG TAAGACTCATTCT	667	GCCAUGGCUCUGAGUGGUA AGACUCAUUCU
HA01_exon6	-	ATTC	213	TTGTTTACAACCTTTCTT TTCTTTTATGATC	668	UUGUUUACAACUUUCUUUU CUUUUAUGAUC

HA01_exon6	-	CTTG	214	TTTACAACCTTTCTTTTC TTTTATGATCTTT	669	UUUACAACUUUCUUUUUUUU UUAUGAUCUUU
HA01_exon6	-	GTTT	215	ACAACCTTTCTTTTCTTT TATGATCTTTAAG	670	ACAACUUUCUUUUUUUUUA UGAUCUUUAAG
HA01_exon6	+	CTTA	216	AAGATCATAAAAGAAAA GAAAGTTGTAAAC	671	AAGAUCAUAAAAGAAAAGA AAGUUGUAAAC
HA01_exon6	+	GTTG	217	TCTATTTTATATATTCA TTTCTTTGTCCAG	672	UCUAUUUUUAUAUUCAUU UCUUUGUCCAG
HA01_exon6	+	CTTA	218	CCACTCAGAGCCATGGC CAACCGGAATTCT	673	CCACUCAGAGCCAUGGCCA ACCGGAUUUCU
HA01_exon6	+	ATTC	219	TTCCTTTAGTATCTCGA GGACATCTTGAAC	674	UUCUUUUAGUAUCUCGAGG ACAUCUUGAAC
HA01_exon6	+	CTTC	220	CTTTAGTATCTCGAGGA CATCTTGAACACC	675	CUUUAGUAUCUCGAGGACA UCUUGAACACC
HA01_exon6	+	CTTT	221	AGTATCTCGAGGACATC TTGAACACCTTTTC	676	AGUAUCUCGAGGACAUCUU GAACACCUUUC
HA01_exon6	+	TTTA	222	GTATCTCGAGGACATCT TGAACACCTTTCT	677	GUAUCUCGAGGACAUCUUG AACACCUUUCU
HA01_exon6	-	GTTT	223	AAGATGTCCTCGAGATA CTAAAGGAAGAAT	678	AAGAUGUCCUCGAGAUACU AAAGGAAGAAU
HA01_exon6	+	GTTG	224	TAAACAAGAATGAGTCT TACCACCTCAGAGC	679	UAAACAAGAAUGAGUCUUA CCACUCAGAGC
HA01_exon6	-	GTTA	225	GGGGGAGAAAGGTGTTC AAGATGTCCTCGA	680	GGGGGAGAAAGGUGUUCAA GAUGUCCUCGA
HA01_exon6	-	CTTT	226	CCAGGTAACCTGGACAAA GAAATGAATATAT	681	CCAGGUAAACUGGACAAAGA AAUGAAUAUAU
HA01_exon6	-	TTTC	227	ACTTGCTTAGGGGGAGA AAGGTGTTCAAGA	682	ACUUGGUUAGGGGGAGAAA GGUGUUCAGA
HA01_exon6	-	TTTT	228	CACTTGCTTAGGGGGAG AAAGGTGTTCAAG	683	CACUUGGUUAGGGGGAGAA AGGUGUUCAG
HA01_exon6	-	CTTT	229	TCACTTGCTTAGGGGGA GAAAGGTGTTCAA	684	UCACUUGGUUAGGGGGAGA AAGGUGUUCAA
HA01_exon6	-	GTTT	230	TGAATCACTCTGTATCT TTTCACTTGTTTA	685	UGAAUCACUCUGUAUCUUU UCACUUGGUUA
HA01_exon6	-	TTTA	231	GTTCTGAATCACTCTGT ATCTTTTCACTTG	686	GUUCUGAAUCACUCUGUAU CUUUUCACUUG
HA01_exon6	-	ATTT	232	AGTTCTGAATCACTCTG TATCTTTTCACTT	687	AGUUCUGAAUCACUCUGUA UCUUUUCACUU
HA01_exon6	-	CTTG	233	ACAGTAAACAAATGAA TAAACAAGTCAG	688	ACAGUAAAACAAUGAAUA AAACAAGUCAG
HA01_exon6	-	TTTC	234	CAGGTAACCTGGACAAAG AAATGAATATATA	689	CAGGUAACUGGACAAAGAA AUGAAUAUAUA
HA01_exon6	+	CTTG	235	AACACCTTTCTCCCCCT AACCAAGTGAAAA	690	AACACCUUUCUCCCCCUAA CCAAGUGAAAA
HA01_exon6	-	CTTA	236	GCTTTCCAGGTAACCTGG ACAAAGAAATGAA	691	GCUUCCAGGUAACUGGAC AAAGAAUUGAA
HA01_exon6	-	TTTG	237	GGGCTTAGCTTTCCAGG TAACTGGACAAAG	692	GGGCUUAGCUUUCAGGUA ACUGGACAAAG
HA01_exon6	-	GTTT	238	GGGGCTTAGCTTTCCAG GTAACCTGGACAAA	693	GGGGCUUAGCUUUCAGGU AACUGGACAAA
HA01_exon6	-	TTTG	239	TGGGGAGACCAATCGTT TGGGGCTTAGCTT	694	UGGGGAGACCAAUUCGUUUG GGGCUUAGCUU
HA01_exon6	-	GTTT	240	GTGGGGAGACCAATCGT TTGGGGCTTAGCT	695	GUGGGGAGACCAAUUCGUUU GGGGCUUAGCU
HA01_exon6	-	CTTG	241	GTTAGGGGGAGAAAGGT GTTCAAGATGTCC	696	GUUAGGGGGAGAAAGGUGU UCAAGAUGUCC

HA01_exon6	+	CTTT	242	CTCCCCCTAACCAAGTG AAAAGATACAGAG	697	CUCCCCCUAACCAAGUGAA AAGAUACAGAG
HA01_exon6	+	GTTT	243	TACTGTCAAGTTGTCTA TTTTATATATTCA	698	UACUGUCAAGUUGUCUAUU UUUAUUAUUA
HA01_exon6	+	ATTC	244	AGAACTAAATCAGTCTG ACTTGTTTTATTTC	699	AGAACUAAAUCAGUCUGAC UUGUUUUUAUUC
HA01_exon6	+	G TTC	245	AATAATGTGACTCTATT AACACTGAATTGT	700	AAUAAUGUGACUCUAUUAA CACUGAAUUGU
HA01_exon6	+	T TTC	246	TGGCAGAACATCAATCT GGGGAAGAAAAAG	701	UGGCAGAACAUCAAUCUGG GGAAAGAAAAAG
HA01_exon6	+	ATTT	247	CTGGCAGAACATCAATC TGGGGAAGAAAA	702	CUGGCAGAACAUCAAUCUG GGGGAAGAAAA
HA01_exon6	+	CTTC	248	CACAGCCTCCACAATTT CTGGCAGAACATC	703	CACAGCCUCCACAUAUUUCU GGCAGAACAU
HA01_exon6	+	CTTC	249	CCTTCCACAGCCTCCAC AATTTCTGGCAGA	704	CCUCCACAGCCUCCACAA UUUCUGGCAGA
HA01_exon6	+	CTTC	250	CACCTTCCCTTCCACAG CCTCCACAATTTTC	705	CACCUUCCCUUCCACAGCC UCCACAUAUUUC
HA01_exon6	+	T TTC	251	CGCACACCCCCGTCCAG GAAGACTTCCACC	706	CGCACACCCCCGUCCAGGA AGACUCCACC
HA01_exon6	+	CTTT	252	CCGCACACCCCCGTCCA GGAAGACTTCCAC	707	CCGCACACCCCCGUCCAGG AAGACUCCACC
HA01_exon6	+	T TTC	253	AGAACATCAGTGCCTTT CCGCACACCCCCG	708	AGAACAUCAGUGCCUUUCC GCACACCCCCG
HA01_exon6	+	CTTT	254	CAGAACATCAGTGCCTT TCCGCACACCCCC	709	CAGAACAUCAUGCCUUUUC CGCACACCCCC
HA01_exon6	+	CTTG	255	GCGCCAAGAGCCAGAGC TTTCAGAACATCA	710	GCGCCAAGAGCCAGAGCUU UCAGAACAUCA
HA01_exon6	+	ATTG	256	GTCTCCCCACAAACACA GCCTTGGCGCCAA	711	GUCUCCCCACAAACACAGC CUUGGCGCCAA
HA01_exon6	+	GTTA	257	CCTGGAAGCTAAGCCC CAAACGATTGGTC	712	CCUGGAAAGCUAAGCCCCA AACGAUUGGUC
HA01_exon6	+	TTTG	258	TCCAGTTACCTGGAAAG CTAAGCCCCAAAC	713	UCCAGUUACCUGGAAAGCU AAGCCCCAAAC
HA01_exon6	+	CTTT	259	GTCCAGTTACCTGGAAA GCTAAGCCCCAAA	714	GUCCAGUUACCUGGAAAGC UAAGCCCCAAA
HA01_exon6	+	T TTC	260	TTTGTCCAGTTACCTGG AAAGCTAAGCCCC	715	UUUGUCCAGUUACCUGGAA AGCUAAGCCCC
HA01_exon6	+	ATTT	261	CTTTGTCCAGTTACCTG GAAAGCTAAGCCC	716	CUUUGUCCAGUUACCUGGA AAGCUAAGCCC
HA01_exon6	+	CTTG	262	TTTTATTTCATTGTTTT ACTGTCAAGTTGT	717	UUUUUAUUAUUUUUUUAC UGUCAAGUUGU
HA01_exon6	+	GTTT	263	TATTCATTTGTTTTACT GTCAAGTTGTCTA	718	UAUUCAUUUUUUUUUACUGU CAAGUUGUCUA
HA01_exon6	+	TTTT	264	ATTCATTTGTTTTACTG TCAAGTTGTCTAT	719	AUUCAUUUUUUUUUACUGUC AAGUUGUCUAU
HA01_exon6	+	T TTA	265	TTCATTGTGTTTTACTGT CAAGTTGTCTATT	720	UUCAUUUUUUUUUACUGUCA AGUUGUCUAUU
HA01_exon6	+	ATTC	266	ATTTGTTTTACTGTCAA GTTGTCTATTTTA	721	AUUUGUUUUUACUGUCAAGU UGUCUAUUUUUA
HA01_exon6	+	ATTT	267	GTTTTACTGTCAAGTTG TCTATTTTATATA	722	GUUUUACUGUCAAGUUGUC UAUUUUUAUUA
HA01_exon6	+	T TTC	268	TCCCCCTAACCAAGTGA AAAGATACAGAGT	723	UCCCCCUAACCAAGUGAAA AGAUAACAGAGU
HA01_exon6	+	TTTG	269	TTTTACTGTCAAGTTGT CTATTTTATATAT	724	UUUUACUGUCAAGUUGUCU AUUUUAUUAU

HA01_exon6	+	TTTT	270	ACTGTCAAGTTGTCTAT TTTATATATTCAT	725	ACUGUCAAGUUGUCUAUUU UAUAUAUUCAU
HA01_exon6	+	TTTA	271	CTGTCAAGTTGTCTATT TTATATATTCATT	726	CUGUCAAGUUGUCUAUUUU AUAUAUUCAUU
HA01_exon6	+	ATTT	272	TATATATTCATTTCTTT GTCCAGTTACCTG	727	UAUAUAUUCAUUUCUUUGU CCAGUUACCUG
HA01_exon6	+	TTTT	273	ATATATTCATTTCTTTG TCCAGTTACCTGG	728	AUAUAUUCAUUUCUUUGUC CAGUUACCUGG
HA01_exon6	+	TTTA	274	TATATTCATTTCTTTGT CCAGTTACCTGGA	729	UAUAUUCAUUUCUUUGUCC AGUUACCUGGA
HA01_exon6	+	ATTC	275	ATTTCTTTGTCCAGTTA CCTGGAAAGCTAA	730	AUUUCUUUGUCCAGUUACC UGGAAAGCUAA
HA01_exon6	-	CTTG	276	GCGCCAAGGCTGTGTTT GTGGGAGACCAA	731	GCGCCAAGGCUGUGUUUGU GGGGAGACCAA
HA01_exon6	-	G TTC	277	TGAAAGCTCTGGCTCTT GGCGCCAAGGCTG	732	UGAAAGCUCUGGCUCUUGG CGCCAAGGCUG
HA01_exon6	-	ATTG	278	TGGAGGCTGTGGAAGGG AAGGTGGAAGTCT	733	UGGAGGCUGUGGAAGGGAA GGUGGAAGUCU
HA01_exon6	-	ATTA	279	TTGAACTTTTCTTTCCC CAGATTGATGTTT	734	UGAACUUUUUUUUCCCCA GAUUGAUGUUC
HA01_exon6	-	G TTC	280	TGCCAGAAATTGTGGAG GCTGTGGAAGGGA	735	UGCCAGAAAUUGUGGAGGC UGUGGAAGGGA
HA01_exon6	-	ATTG	281	ATGTTCTGCCAGAAATT GTGGAGGCTGTGG	736	AUGUUCUGCCAGAAAUUGU GGAGGCUGUGG
HA01_exon6	-	TTTC	282	CCCAGATTGATGTTCTG CCAGAAATTGTGG	737	CCCAGAUUGAUGUUCUGCC AGAAAUUGUGG
HA01_exon6	-	CTTT	283	CCCAGATTGATGTTCT GCCAGAAATTGTG	738	CCCAGAUUGAUGUUCUGC CAGAAAUUGUG
HA01_exon6	-	TTTC	284	TTTCCCCAGATTGATGT TCTGCCAGAAATT	739	UUUCCCCAGAUUGAUGUUC UGCCAGAAAUU
HA01_exon6	-	TTTT	285	CTTTCCCCAGATTGATG TTCTGCCAGAAAT	740	CUUCCCCAGAUUGAUGUU CUGCCAGAAAU
HA01_exon6	-	CTTT	286	TCTTTCCCCAGATTGAT GTTCTGCCAGAAA	741	UCUUCCCCAGAUUGAUGU UCUGCCAGAAA
HA01_exon6	-	ATTG	287	AACTTTTCTTTCCCCAG ATTGATGTTCTGC	742	AACUUUUUUUUCCCCAGAU UGAUGUUCUGC
HA01_exon6	-	G TTA	288	ATAGAGTCACATTATTG AACTTTTCTTTCC	743	AUAGAGUCACAUUAUUGAA CUUUUCUUUCC
HA01_exon6	-	ATTC	289	AGTGTTAATAGAGTCAC ATTATTGAACTTT	744	AGUGUUAUAGAGUCACAU UAUUGAACUUU
HA01_exon6	-	CTTC	290	CTGGACGGGGGTGTGCG GAAAGGCACTGAT	745	CUGGACGGGGGUGUGCGGA AAGGCACUGAU
HA01_exon6	-	CTTT	291	CTTTTCTTTTATGATCT TTAAGT	746	CUUUUCUUUUUGAUCUUU AAGU
HA01_exon6	-	TTTC	292	TTTTCTTTTATGATCTT TAAGT	747	UUUUCUUUUUGAUCUUUA AGU
HA01_exon7	-	ATTT	293	TTTCAGGGTGCCAGAAAT GTGAAAGTCATCG	748	UUUCAGGGUGCCAGAAUGU GAAAGUCAUCG
HA01_exon7	-	ATTA	294	TTTTTTCAGGGTGCCAG AATGTGAAAGTCA	749	UUUUUUCAGGGUGCCAGAA UGUGAAAGUCA
HA01_exon7	-	ATTG	295	TAAGCTCAGGTTCAAAG TGTTGGTAATGCC	750	UAAGCUCAGGUUCAAAGUG UUGGUAAUGCC
HA01_exon7	-	G TTC	296	ATATTAAATGTATGCAT TATTTTTCAGGG	751	AUAUUAAAUGUAUGCAUUA UUUUUUCAGGG
HA01_exon7	-	ATTC	297	AGTTCATATTAAATGTA TGCATTATTTTTT	752	AGUUCAUAUUAAAUGUAUG CAUUUUUUUUU

HA01_exon7	–	TTTT	298	TTCAGGGTGCCAGAATG TGAAAGTCATCGA	753	UUCAGGGUGCCAGAAUGUG AAAGUCAUCGA
HA01_exon7	–	ATTA	299	AATGTATGCATTATTTT TTCAGGGTGCCAG	754	AAUGUAUGCAUUUUUUUU CAGGGUGCCAG
HA01_exon7	–	TTTT	300	TCAGGGTGCCAGAATGT GAAAGTCATCGAC	755	UCAGGGUGCCAGAAUGUGA AAGUCAUCGAC
HA01_exon7	–	TTTG	301	GCCGTTTCCAAGATCTG ACAGTGCACAATA	756	GCCGUUCCAAGAUUCUGAC AGUGCACAAUA
HA01_exon7	–	TTTC	302	AGGGTGCCAGAATGTGA AAGTCATCGACAA	757	AGGGUGCCAGAAUGUGAAA GUCAUCGACAA
HA01_exon7	–	ATTG	303	GTGAGGAAAAATCCTTT GGCCGTTTCCAAG	758	GUGAGGAAAAUCCUUUGG CCGUUCCAAG
HA01_exon7	–	CTTT	304	GGCCGTTTCCAAGATCT GACAGTGCACAAT	759	GGCCGUUCCAAGAUUCUGA CAGUGCACAAU
HA01_exon7	–	ATTG	305	CATTCACTTCATATTAA ATGTATGCATTAT	760	CAUUCAGUUAUUAUUAAU GUAUGCAUUUAU
HA01_exon7	–	GTTT	306	CCAAGATCTGACAGTGC ACAATATTTTCCC	761	CCAAGAUCUGACAGUGCAC AAUUAUUUCCC
HA01_exon7	–	TTTC	307	CAAGATCTGACAGTGCA CAATATTTTCCA	762	CAAGAUCUGACAGUGCACA AUUUUUUCCCA
HA01_exon7	–	TTTT	308	CAGGGTGCCAGAATGTG AAAGTCATCGACA	763	CAGGGUGCCAGAAUGUGAA AGUCAUCGACA
HA01_exon7	–	ATTA	309	TTGCATTCACTTCATAT TAAATGTATGCAT	764	UUGCAUUCAGUUAUUAUA AAUGUAUGCAU
HA01_exon7	–	ATTG	310	GAGGTAGCAAACACTAA GGTGAAGATAA	765	GAGGUAGCAAACACUAAGG UGAAAAGAUAA
HA01_exon7	–	GTTT	311	AGACAACGTCATCCCT GGCAGGCTAAAGT	766	AGACAACGUAUCCCCUGG CAGGCUAAGU
HA01_exon7	–	CTTA	312	AATTGTAAGCTCAGGTT CAAAGTGTGGTA	767	AAUUGUAAGCUCAGGUUCA AAGUGUUGGUA
HA01_exon7	–	GTTT	313	TTAAATGTGAAGCTCAG GTTCAAAGTGTG	768	UUAAAUGUAAGCUCAGGU UCAAGUGUUG
HA01_exon7	–	TTTA	314	AAACAGTGGTCTTTAA TTGTAAGCTCAGG	769	AAACAGUGGUUCUUAAAAU GUAAGCUCAGG
HA01_exon7	–	CTTT	315	AAAACAGTGGTCTTTAA ATTGTAAGCTCAG	770	AAAACAGUGGUUCUUAAAAU UGUAAGCUCAG
HA01_exon7	–	TTTA	316	CATGCTTTTAAACAGT GGTCTTTAAATTG	771	CAUGUCUUUAAAACAGUGG UUCUUAAAUG
HA01_exon7	–	GTTT	317	ACATGCTCTTTAAACAG TGGTCTTTAAATT	772	ACAUGUCUUUAAAACAGUG GUUCUUAAAUA
HA01_exon7	–	ATTC	318	TGTTTACATGTCTTTAA AACAGTGGTCTTT	773	UGUUUACAUGUCUUUAAAA CAGUGGUUCUU
HA01_exon7	–	ATTA	319	ACCTGTATTCTGTTTAC ATGCTCTTTAAAC	774	ACCUGUAUUCUGUUUACAU GUCUUUAAAAC
HA01_exon7	–	TTTA	320	TTAACCTGTATTCTGTT TACATGCTTTTAA	775	UUAACCUGUAUUCUGUUUA CAUGUCUUUAA
HA01_exon7	–	GTTT	321	ATTAACCTGTATTCTGT TTACATGCTTTTAA	776	AUUAACCUGUAUUCUGUUU ACAUGUCUUUA
HA01_exon7	–	ATTG	322	TTTATTAACCTGTATTCT TGTTTACATGTCT	777	UUUUAUAACCUGUAUUCUG UUUACAUGUCU
HA01_exon7	–	ATTT	323	TCCCATCTGTATTATTT TTTTTCAGCATGT	778	UCCCAUCUGUAUUUUUUUU UUUCAGCAUGU
HA01_exon7	–	TTTA	324	GTAAAATTGGAGGTAGC AAACACTAAGGTG	779	GUAAAAUUGGAGGUAGCAA ACACUAAGGUG
HA01_exon7	–	CTTT	325	AGTAAAATTGGAGGTAG CAAACACTAAGGT	780	AGUAAAAUUGGAGGUAGCA AACACUAAGGU

HA01_exon7	–	TTTA	326	GACAACGTCATCCCCTG GCAGGCTAAAGTG	781	GACAACGUCAUCCCCUGGC AGGCUAAAAGUG
HA01_exon7	–	ATTA	327	TTATTGCATTTCAGTTCA TATTAATGTATG	782	UUAUUGCAUUCAGUUCAUA UUAAAUGUAUG
HA01_exon7	–	TTTT	328	CCCATCTGTATTATTTT TTTTCAGCATGTA	783	CCCAUCUGUAUUUUUUUU UUCAGCAUGUA
HA01_exon7	–	TTTT	329	TTCAGCATGTATTACTT GACAAAGAGACAC	784	UUCAGCAUGUAUUACUUGA CAAAGAGACAC
HA01_exon7	–	ATTA	330	TTTTTTTTTCAGCATGTA TTACTTGACAAAG	785	UUUUUUUUCAGCAUGUAUU ACUUGACAAAG
HA01_exon7	–	TTTC	331	ATTGCTTTTGACTTTTC AATGGGTGTCCTA	786	AUUGCUUUUGACUUUUCAA UGGGUGUCCUA
HA01_exon7	–	ATTG	332	CTTTTGACTTTTCAATG GGTGCTCAGGAA	787	CUUUUGACUUUUCAAUGGG UGUCCUAGGAA
HA01_exon7	–	CTTT	333	TGACTTTTCAATGGGTG TCCTAGGAACCTT	788	UGACUUUUCAAUGGGUGUC CUAGGAACCUU
HA01_exon7	–	TTTT	334	GACTTTTCAATGGGTGT CCTAGGAACCTTT	789	GACUUUUCAAUGGGUGUCC UAGGAACCUUU
HA01_exon7	–	TTTG	335	ACTTTTCAATGGGTGTC CTAGGAACCTTTT	790	ACUUUUCAAUGGGUGUCCU AGGAACCUUUU
HA01_exon7	–	CTTT	336	TCAATGGGTGTCCTAGG AACCTTTTAGAAA	791	UCAAUGGGUGUCCUAGGAA CCUUUUAGAAA
HA01_exon7	–	TTTT	337	CAATGGGTGTCCTAGGA ACCTTTTAGAAAAG	792	CAAUGGGUGUCCUAGGAAC CUUUUAGAAAAG
HA01_exon7	–	TTTC	338	AATGGGTGTCCTAGGAA CCTTTTAGAAAAGA	793	AAUGGGUGUCCUAGGAACC UUUUAGAAAAGA
HA01_exon7	–	CTTT	339	TAGAAAGAAATGGACTT TCATCCTGGAAAT	794	UAGAAAGAAUUGGACUUUC AUCCUGGAAAU
HA01_exon7	–	TTTT	340	AGAAAGAAATGGACTTT CATCCTGGAAATA	795	AGAAAGAAUUGGACUUUCA UCCUGGAAUA
HA01_exon7	–	TTTA	341	GAAAGAAATGGACTTTC ATCCTGGAAATAT	796	GAAAGAAUUGGACUUUCAU CCUGGAAUAU
HA01_exon7	–	CTTT	342	CATCCTGGAAATATATT AACTGTAAAAAG	797	CAUCCUGGAAUAUAUUAA CUGUAAAAAG
HA01_exon7	–	TTTC	343	ATCCTGGAAATATATTA ACTGTAAAAAGA	798	AUCCUGGAAUAUAUUAAAC UGUAAAAAGA
HA01_exon7	–	ATTA	344	ACTGTAAAAAGAAAAC ATTGAAAATGTGT	799	ACUGUAAAAAGAAAACAU UGAAAUGUGU
HA01_exon7	–	GTTA	345	AAAAGAAAACATTGAAA ATGTGTTTAGACA	800	AAAAGAAAACAUUGAAAAU GUGUUUAGACA
HA01_exon7	–	ATTT	346	CATTGCTTTTGACTTTT CAATGGGTGTCCT	801	CAUUGCUUUUGACUUUUCA AUGGGUGUCCU
HA01_exon7	–	TTTC	347	CCATCTGTATTATTTTT TTTCAGCATGTAT	802	CCAUCUGUAUUUUUUUUU UCAGCAUGUAU
HA01_exon7	–	TTTA	348	TTTCATTGCTTTTGACT TTTCAATGGGTGT	803	UUUCAUUGCUUUUGACUUU UCAUUGGGUGU
HA01_exon7	–	CTTT	349	TATTTATTGCTTTTGA CTTTTCAATGGGT	804	UAUUUCAUUGCUUUUGACU UUUCAUUGGGU
HA01_exon7	–	ATTG	350	AAAATGTGTTTAGACAA CGTCATCCCCTGG	805	AAAUGUGUUUAGACAACG UCAUCCCCUGG
HA01_exon7	–	ATTT	351	TTTTTCAGCATGTATTA CTTGACAAAGAGA	806	UUUUUCAGCAUGUAUUACU UGACAAAGAGA
HA01_exon7	–	TTTT	352	TTTTCAGCATGTATTAC TTGACAAAGAGAC	807	UUUUCAGCAUGUAUUACUU GACAAAGAGAC
HA01_exon7	–	TTTT	353	TTTCAGCATGTATTACT TGACAAAGAGACA	808	UUUCAGCAUGUAUUACUUG ACAAAGAGACA

HA01_exon7	-	TTTT	354	TCAGCATGTATTACTTG ACAAAGAGACACT	809	UCAGCAUGUAUUACUUGAC AAAGAGACACU
HA01_exon7	-	TTTT	355	CAGCATGTATTACTTGA CAAAGAGACACTG	810	CAGCAUGUAUUACUUGACA AAGAGACACUG
HA01_exon7	-	TTTC	356	AGCATGTATTACTTGAC AAAGAGACACTGT	811	AGCAUGUAUUACUUGACAA AGAGACACUGU
HA01_exon7	-	ATTA	357	CTTGACAAAGAGACACT GTGCAGAGGGTGA	812	CUUGACAAAGAGACACUGU GCAGAGGGUGA
HA01_exon7	-	CTTG	358	ACAAAGAGACACTGTGC AGAGGGTGACCAC	813	ACAAAGAGACACUGUGCAG AGGGUGACCAC
HA01_exon7	-	ATTC	359	CCCACCTCAATACAAAG GGTGTCTCTTT	814	CCCACUUCAAUACAAAGGG UGUCGUUCUUU
HA01_exon7	-	CTTC	360	AATACAAAGGGTGTCTGT TCTTTTCCAACAA	815	AAUACAAAGGGUGUCGUUC UUUUCCAACAA
HA01_exon7	-	GTTT	361	TTTTTCAACAAAATAGC AATCCCTTTTATT	816	UUUUCCAACAAAUAAGCAA UCCCUUUUAUU
HA01_exon7	-	CTTT	362	TCCAACAAAATAGCAAT CCCTTTTATTTCA	817	UCCAACAAAUAAGCAAUCC CUUUUAUUUCA
HA01_exon7	-	TTTT	363	CCAACAAAATAGCAATC CCTTTTATTTTAT	818	CCAACAAAUAAGCAAUCCC UUUUUAUUUCAU
HA01_exon7	-	TTTC	364	CAACAAAATAGCAATCC CTTTTATTTTATT	819	CAACAAAUAAGCAAUCCCU UUUAUUUCAUU
HA01_exon7	-	TTTT	365	ATTTTCATTGCTTTTGAC TTTTCAATGGGTG	820	AUUUCAUUGCUUUUGACUU UUCAAUGGGUG
HA01_exon7	-	GTTT	366	AAAGTGTGGTAATGCC TGATTCACAACCT	821	AAAGUGUUGGUAUAGCCUG AUUCACAACUU
HA01_exon7	+	ATTT	367	CTCTCTAAGAAGTAACA TACATCTAAAAC	822	CUCUCUAAGAAGUAACAUA CAUCCUAAAAC
HA01_exon7	-	ATTC	368	ACAACTTTGAGAAGGTA GCACCTGGAGAGAA	823	ACAACUUUGAGAAGGUAGC ACUGGAGAGAA
HA01_exon7	+	TTTC	369	ACCTTAGTGTGTTGCTAC CTCCAATTTTACT	824	ACCUUAGUGUUUGCUACCU CCAAUUUUACU
HA01_exon7	+	CTTA	370	GTGTTTGCTACCTCCAA TTTTACTAAAGGA	825	GUGUUUGCUACCUCCAAUU UUACUAAAAGGA
HA01_exon7	+	GTTT	371	GCTACCTCCAATTTTAC TAAAGGATACAGC	826	GCUACCUCCAUUUUUACUA AAGGAUACAGC
HA01_exon7	+	TTTG	372	CTACCTCCAATTTTACT AAAGGATACAGCA	827	CUACCUCCAUUUUUACUAA AGGAUACAGCA
HA01_exon7	+	ATTT	373	TACTAAAGGATACAGCA CTTTAGCCTGCCA	828	UACUAAAGGAUACAGCACU UUAGCCUGCCA
HA01_exon7	+	TTTT	374	ACTAAAGGATACAGCAC TTTAGCCTGCCAG	829	ACUAAAGGAUACAGCACUU UAGCCUGCCAG
HA01_exon7	+	TTTA	375	CTAAAGGATACAGCACT TTAGCCTGCCAGG	830	CUAAAGGAUACAGCACUUU AGCCUGCCAGG
HA01_exon7	+	CTTT	376	AGCCTGCCAGGGGATGA CGTTGTCTAAACA	831	AGCCUGCCAGGGGAUGACG UUGUCUAAACA
HA01_exon7	+	TTTA	377	GCCTGCCAGGGGATGAC GTTGTCTAAACAC	832	GCCUGCCAGGGGAUGACGU UGUCUAAACAC
HA01_exon7	+	TTTT	378	CACCTTAGTGTGTTGCTA CCTCCAATTTTAC	833	CACCUUAGUGUUUGCUACC UCCAAUUUUAC
HA01_exon7	+	GTTG	379	TCTAAACACATTTTCAA TGTTTTCTTTTTA	834	UCUAAACACAUUUUCAUG UUUUCUUUUUA
HA01_exon7	+	TTTT	380	CAATGTTTTCTTTTTAA CAGTTAATATATT	835	CAAGUUUUUCUUUUUAACA GUUAAUAUAUU
HA01_exon7	+	TTTC	381	AATGTTTTCTTTTTAAC AGTTAATATATTT	836	AAUGUUUUUCUUUUUAACAG UUAAUAUAUUU

HA01_exon7	+	GTTT	382	TCTTTTAAACAGTTAAT ATATTCCAGGAT	837	UCUUUUUAAACAGUUAUAU AUUCCAGGAU
HA01_exon7	+	TTTT	383	CTTTTAAACAGTTAATA TATTTCAGGATG	838	CUUUUUAAACAGUUAUAU UUUCCAGGAUG
HA01_exon7	+	TTTC	384	TTTTTAAACAGTTAATAT ATTTCAGGATGA	839	UUUUUAAACAGUUAUAU UCCAGGAUGA
HA01_exon7	+	CTTT	385	TTAACAGTTAATATATT TCCAGGATGAAAG	840	UUAACAGUUAUAUAUUUC CAGGAUGAAAG
HA01_exon7	+	TTTT	386	TAACAGTTAATATATTT CCAGGATGAAAGT	841	UAACAGUUAUAUAUUUCC AGGAUGAAAGU
HA01_exon7	+	TTTT	387	AACAGTTAATATATTTTC CAGGATGAAAGTC	842	AACAGUUAUAUAUUUCCA GGAUGAAAGUC
HA01_exon7	+	TTTA	388	ACAGTTAATATATTTCC AGGATGAAAGTCC	843	ACAGUUAUAUAUUUCCAG GAUGAAAGUCC
HA01_exon7	+	ATTT	389	TCAATGTTTTCTTTTAA ACAGTTAATATAT	844	UCAUGUUUUUUUUUAAAC AGUUAUAUAU
HA01_exon7	+	GTTA	390	ATATATTTCCAGGATGA AAGTCCATTTCTT	845	AUAUAUUUCCAGGAUGAAA GUCCAUAUUUU
HA01_exon7	+	CTTT	391	TCACCTTAGTGTGTTGCT ACCTCCAATTTTA	846	UCACCUUAGUGUUUGCUAC CUCCAUAUUUU
HA01_exon7	+	GTTA	392	ATAAACAATGAGATCAT TATCTTTTCACCT	847	AUAAACAAGAGAUAUAUA UCUUUUCACCU
HA01_exon7	+	TTTC	393	TCTCTAAGAAGTAACAT ACATCCTAAACA	848	UCUCUAAGAAGUAACAUAC AUCCUAAAACA
HA01_exon7	+	ATTT	394	GGATATATTCAGACACT AAAGATGTGATTG	849	GGAUUAUUCAGACACUAA AGAUGUGAUUG
HA01_exon7	+	TTTG	395	GATATATTCAGACACTA AAGATGTGATTGG	850	GAUAUAUUCAGACACUAAA GAUGUGAUUGG
HA01_exon7	+	ATTC	396	AGACACTAAAGATGTGA TTGGAATCTACA	851	AGACACUAAAGAUGUGAUU GGAAUUCUACA
HA01_exon7	+	ATTG	397	GAAATCTACATTCAAAG AAGTATCACCAT	852	GAAAUUCUACAUAUCAAAG GUAUCACCAAU
HA01_exon7	+	ATTC	398	AAAGAAGTATCACCAT TACCGCACCCAT	853	AAAGAAGUAUCACCAUUA CCGCCACCCA
HA01_exon7	+	ATTA	399	CCGCCACCCATTCCAAT TCTCTCCAGTGCT	854	CCGCCACCAUUCACCAUUC UCUCCAGUGCU
HA01_exon7	+	ATTC	400	CAATTCTCTCCAGTGCT ACCTTCTCAAAGT	855	CAAUUCUCCAGUGCUAC CUUCUCAAGU
HA01_exon7	+	ATTC	401	TCTCCAGTGCTACCTTC TCAAAGTTGTGAA	856	UCUCCAGUGCUACCUUCUC AAAGUUGUGAA
HA01_exon7	+	ATTA	402	TCTTTTCACCTTAGTGT TTGCTACCTCCAA	857	UCUUUUCACCUUAGUGUUU GCUACCUCAA
HA01_exon7	+	CTTC	403	TCAAAGTTGTGAATCAG GCATTACCAACAC	858	UCAAGUUGUGAAUCAGGC AUUACCAACAC
HA01_exon7	+	ATTA	404	CCAACACTTTGAACCTG AGCTTACAATTTA	859	CCAACACUUUGAACCUAG CUUACAAUUUA
HA01_exon7	+	CTTT	405	GAACCTGAGCTTACAAT TTAAGAACCACCTG	860	GAACCUGAGCUUACAAUUU AAGAACCACUG
HA01_exon7	+	TTTG	406	AACCTGAGCTTACAATT TAAGAACCACCTGT	861	AACCUGAGCUUACAAUUUA AGAACCACUGU
HA01_exon7	+	CTTA	407	CAATTTAAGAACCACCTG TTTTAAAGACATG	862	CAUUUUAAGAACCACUGUU UUAAAGACAUG
HA01_exon7	+	ATTT	408	AAGAACCACCTGTTTTAA AGACATGTAAACA	863	AAGAACCACUGUUUUAAAG ACAUGUAAACA
HA01_exon7	+	TTTA	409	AGAACCACCTGTTTTAAA GACATGTAAACAG	864	AGAACCACUGUUUUAAAGA CAUGUAAACAG

HA01_exon7	+	GTTT	410	TAAAGACATGTAAACAG AATACAGGTTAAT	865	UAAAGACAUGUAAACAGAA UACAGGUUAAU
HA01_exon7	+	TTTT	411	AAAGACATGTAAACAGA ATACAGGTTAATA	866	AAAGACAUGUAAACAGAAU ACAGGUUAAUA
HA01_exon7	+	TTTA	412	AAGACATGTAAACAGAA TACAGGTTAATAA	867	AAGACAUGUAAACAGAAUA CAGGUUAAUUA
HA01_exon7	+	GTTG	413	TGAATCAGGCATTACCA ACACTTTGAACCT	868	UGAAUCAGGCAUUACCAAC ACUUUGAACCU
HA01_exon7	-	GTTG	414	GTAATGCCTGATTACACA ACTTTGAGAAGGT	869	GUA AUGCCUGAUUCACAAC UUUGAGAAGGU
HA01_exon7	+	ATTT	415	CCAGGATGAAAGTCCAT TTCTTTCTAAAAAG	870	CCAGGAUGAAAGUCCAUUU CUUUCUAAAAAG
HA01_exon7	+	CTTT	416	ATTTCTCTCTAAGAAGT AACATACATCCTA	871	AUUUCUCUCUAAGAAGUAA CAUACAUCCUA
HA01_exon7	+	TTTC	417	ACATTCTGGCACCCCTGA AAAAATAATGCAT	872	ACAUUCUGGCACCCUGAAA AAAUAAUGCAU
HA01_exon7	+	ATTC	418	TGGCACCTGAAAAAAT AATGCATACATTT	873	UGGCACCCUGAAAAAUAA UGCAUACAUUU
HA01_exon7	+	TTTA	419	TTTCTCTCTAAGAAGTA ACATACATCCTAA	874	UUUCUCUCUAAGAAGUAAC AUACAUCCUAA
HA01_exon7	+	CTTC	420	CCAAAAATGCTTTATTT CTCTCTAAGAAGT	875	CCAAAAAUGCUUUUUUUUCU CUCUAAGAAGU
HA01_exon7	-	CTTC	421	TTAGAGAGAAATAAAGC ATTTTGGGAAGA	876	UUAGAGAGAAAUAAGCAU UUUUGGGAAGA
HA01_exon7	-	GTTA	422	CTTCTTAGAGAGAAATA AAGCATTTTGGG	877	CUUCUUAGAGAGAAUAAA GCAUUUUUGGG
HA01_exon7	-	TTTA	423	GGATGTATGTTACTTCT TAGAGAGAAATAA	878	GGAUGUAUGUUACUUCUUA GAGAGAAUAA
HA01_exon7	-	TTTT	424	AGGATGTATGTTACTTC TTAGAGAGAAATA	879	AGGAUGUAUGUUACUUCUU AGAGAGAAUUA
HA01_exon7	-	GTTT	425	TAGGATGTATGTTACTT CTTAGAGAGAAAT	880	UAGGAUGUAUGUUACUUCU UAGAGAGAAAU
HA01_exon7	+	CTTT	426	CACATTCTGGCACCCCTG AAAAATAATGCA	881	CACAUCUGGCACCCUGAA AAAAUAAUGCA
HA01_exon7	-	TTTA	427	GTGTCTGAATATATCCA AATGTTTtaggAT	882	GUGUCUGAAUAUAUCCAAA UGUUUUAGGAU
HA01_exon7	-	TTTC	428	CAATCACATCTTTAGTG TCTGAATATATCC	883	CAAUCACAUUUUAGUGUC UGAAUAUAUCC
HA01_exon7	-	ATTT	429	CCAATCACATCTTTAGT GTCTGAATATATC	884	CCAUCACAUUUUAGUGU CUGAAUAUAUC
HA01_exon7	-	TTTG	430	AATGTAGATTTCCAATC ACATCTTTAGTGT	885	AAUGUAGAUUUCCAUCAC AUCUUUAGUGU
HA01_exon7	-	CTTT	431	GAATGTAGATTTCCAAT CACATCTTTAGTG	886	GAAUGUAGAUUUCCAUCA CAUCUUUAGUG
HA01_exon7	-	CTTC	432	TTTGAATGTAGATTTC AATCACATCTTTA	887	UUUGAAUGUAGAUUUCCAA UCACAUCUUUA
HA01_exon7	-	ATTG	433	GTGATACTTCTTTGAAT GTAGATTTCCAAT	888	GUGAUACUUCUUUGAAUGU AGAUUUCCAAU
HA01_exon7	-	ATTG	434	GAATGGGTGGCGGTAAT TGGTGATACTTCT	889	GAAUGGGUGGCGGUAAUUG GUGAUACUUCU
HA01_exon7	-	TTTG	435	AGAAGGTAGCACTGGAG AGAATTGGAATGG	890	AGAAGGUAGCACUGGAGAG AAUUGGAAUGG
HA01_exon7	-	CTTT	436	GAGAAGGTAGCACTGGA GAGAATTGGAATG	891	GAGAAGGUAGCACUGGAGA GAAUUGGAAUG
HA01_exon7	-	CTTT	437	AGTGTCTGAATATATCC AAATGTTTtaggA	892	AGUGUCUGAAUAUAUCCAA AUGUUUUAGGA

HA01_exon7	+	TTTC	438	CAGGATGAAAGTCCATT TCTTTCTAAAAGG	893	CAGGAUGAAAGUCCAUUUC UUUCUAAAAGG
HA01_exon7	+	CTTG	439	TCGATGACTTTTCACATT CTGGCACCCCTGAA	894	UCGAUGACUUUCACAUUCU GGCACCCUGAA
HA01_exon7	+	TTTT	440	CCTCACCAATGTCTTGT CGATGACTTTTCAC	895	CCUCACCAAUGUCUUGUCG AUGACUUUCAC
HA01_exon7	+	TTTC	441	TTTCTAAAAGGTTTCCTA GGACACCCATTGA	896	UUUCUAAAAGGUUCCUAGG ACACCCAUUGA
HA01_exon7	+	CTTT	442	CTAAAAGGTTTCCTAGGA CACCCATTGAAAA	897	CUAAAAGGUUCCUAGGACA CCCAUUGAAAA
HA01_exon7	+	TTTC	443	TAAAAGGTTTCCTAGGAC ACCCATTGAAAAAG	898	UAAAAGGUUCCUAGGACAC CCAUUGAAAAAG
HA01_exon7	+	GTTT	444	CTAGGACACCCATTGAA AAGTCAAAGCAA	899	CUAGGACACCCAUUGAAAA GUCAAAGCAA
HA01_exon7	+	ATTG	445	AAAAGTCAAAGCAATG AAATAAAAGGGAT	900	AAAAGUCAAAAAGCAUUGAA AUAAAAGGGAU
HA01_exon7	+	ATTG	446	CTATTTTGTGTGGAAAAG AACGACACCCCTTT	901	CUAUUUUGUUGGAAAAGAA CGACACCCUUU
HA01_exon7	+	ATTT	447	TGTTGGAAAAGAACGAC ACCCCTTGTATTG	902	UGUUGGAAAAGAACGACAC CCUUUGUAUUG
HA01_exon7	+	TTTT	448	GTTGGAAAAGAACGACA CCCTTTGTATTGA	903	GUUGGAAAAGAACGACACC CUUUGUAUUGA
HA01_exon7	+	TTTG	449	TTGGAAAAGAACGACAC CCTTTGTATTGAA	904	UUGGAAAAGAACGACACCC UUUGUAUUGAA
HA01_exon7	+	TTTC	450	CTCACCAATGTCTTGTG GATGACTTTTACA	905	CUCACCAAUGUCUUGUCGA UGACUUUCACA
HA01_exon7	+	GTTG	451	GAAAAGAACGACACCCCT TTGTATTGAAGTG	906	GAAAAGAACGACACCCUUU GUUUGAAGUG
HA01_exon7	+	TTTG	452	TATTGAAGTGGGGAATT ACAGACTGTGGTC	907	UAUUGAAGUGGGGAUUAC AGACUGUGGUC
HA01_exon7	+	ATTG	453	AAGTGGGGAATTACAGA CTGTGGTCACCCCT	908	AAGUGGGGAUUACAGACU GUGGUCACCCU
HA01_exon7	+	ATTA	454	CAGACTGTGGTCACCCCT CTGCACAGTGTCT	909	CAGACUGGGUCACCCUCU GCACAGUGUCU
HA01_exon7	+	CTTT	455	GTCAAGTAATACATGCT GAAAAAAATAAT	910	GUCAAGUAAUACAUGCUGA AAAAAAAUAAU
HA01_exon7	+	TTTG	456	TCAAGTAATACATGCTG AAAAAAATAATA	911	UCAAGUAAUACAUGCUGAA AAAAAAUAAUA
HA01_exon7	+	ATTG	457	TGCACGTGCAGATCTTG GAAACGGCCAAAG	912	UGCACUGUCAGAUCUUGGA AACGGCCAAAG
HA01_exon7	+	CTTG	458	GAAACGGCCAAAGGATT TTTCCTACCAAT	913	GAAACGGCCAAAGGAUUUU UCCUCACCAAU
HA01_exon7	+	ATTT	459	TTCTCACCAATGTCTT GTCGATGACTTTT	914	UCCUCACCAAUGUCUUGU CGAUGACUUUC
HA01_exon7	+	TTTT	460	TCCTCACCAATGTCTTG TCGATGACTTTCA	915	UCCUCACCAAUGUCUUGUC GAUGACUUUCA
HA01_exon7	+	CTTT	461	GTATTGAAGTGGGGAAT TACAGACTGTGGT	916	GUUUGAAGUGGGGAUUUA CAGACUGUGGU
HA01_exon7	+	ATTT	462	CTTTCTAAAAGGTTTCCT AGGACACCCATTG	917	CUUUCUAAAAGGUUCCUAG GACACCCAUUG
HA01_exon7	-	CTTA	463	GAGAGAAATAAAGCATT TTTGGGAAGAA	918	GAGAGAAAUAAAGCAUUUU UGGGAAGAA
HA01_exon7	+	ATTT	464	AATATGAAGTGAATGCA ATAATAATCA	919	AAUUGAACUGAAUGCAAU AAUAAUCA
HA01_exon7	+	TTTA	465	ATATGAAGTGAATGCAA TAATAATCA	920	AUAUGAACUGAAUGCAAUA AUAUUA

* The 5'-TTN-3' 3-nucleotide PAM motif is in boldface.

The present disclosure includes all combinations of the direct repeats and spacers listed above, consistent with the disclosure herein.

5 In some embodiments, a spacer sequence described herein comprises an uracil (U). In some embodiments, a spacer sequence described herein comprises a thymine (T). In some embodiments, a spacer sequence according to **Table 5** comprises a sequence comprising a thymine in one or more places indicated as uracil in **Table 5**.

10 (iii). Exemplary RNA Guides

The present disclosure provides RNA guides that comprise any and all combinations of the direct repeats and spacers described herein (*e.g.*, as set forth in **Table 5**, above). In some embodiments, the sequence of an RNA guide has at least 90% identity (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to a sequence of any one of SEQ ID
15 NOs: 967-1023. In some embodiments, an RNA guide has a sequence of any one of SEQ ID NOs: 967-1023.

In some embodiments, exemplary RNA guides provided herein may comprise a spacer sequence of any one of SEQ ID NOs: 1093-1097. In one example, the RNA guide may comprise a spacer of SEQ ID NO: 1096.

20 Any of the exemplary RNA guides disclosed herein may comprise a direct sequence of any one of SEQ ID NOs: 1-10 or a fragment thereof that is at least 23-nucleotide in length. In one example, the direct sequence may comprise SEQ ID NO: 10.

In specific examples, the RNA guides provided herein may comprise the nucleotide sequence of SEQ ID NOs: 967, 968, 988, 989, or 994. In one example, the RNA guide provided
25 herein comprise the nucleotide sequence of SEQ ID NO: 989.

(iv). Modifications

The RNA guide may include one or more covalent modifications with respect to a reference sequence, in particular the parent polyribonucleotide, which are included within the
30 scope of the present disclosure.

Exemplary modifications can include any modification to the sugar, the nucleobase, the internucleoside linkage (*e.g.*, to a linking phosphate/to a phosphodiester linkage/to the

phosphodiester backbone), and any combination thereof. Some of the exemplary modifications provided herein are described in detail below.

The RNA guide may include any useful modification, such as to the sugar, the nucleobase, or the internucleoside linkage (*e.g.*, to a linking phosphate/to a phosphodiester linkage/to the phosphodiester backbone). One or more atoms of a pyrimidine nucleobase may be replaced or substituted with optionally substituted amino, optionally substituted thiol, optionally substituted alkyl (*e.g.*, methyl or ethyl), or halo (*e.g.*, chloro or fluoro). In certain embodiments, modifications (*e.g.*, one or more modifications) are present in each of the sugar and the internucleoside linkage. Modifications may be modifications of ribonucleic acids (RNAs) to deoxyribonucleic acids (DNAs), threose nucleic acids (TNAs), glycol nucleic acids (GNAs), peptide nucleic acids (PNAs), locked nucleic acids (LNAs) or hybrids thereof. Additional modifications are described herein.

In some embodiments, the modification may include a chemical or cellular induced modification. For example, some nonlimiting examples of intracellular RNA modifications are described by Lewis and Pan in “RNA modifications and structures cooperate to RNA guide-protein interactions” from *Nat Reviews Mol Cell Biol*, 2017, 18:202-210.

Different sugar modifications, nucleotide modifications, and/or internucleoside linkages (*e.g.*, backbone structures) may exist at various positions in the sequence. One of ordinary skill in the art will appreciate that the nucleotide analogs or other modification(s) may be located at any position(s) of the sequence, such that the function of the sequence is not substantially decreased. The sequence may include from about 1% to about 100% modified nucleotides (either in relation to overall nucleotide content, or in relation to one or more types of nucleotide, *i.e.*, any one or more of A, G, U or C) or any intervening percentage (*e.g.*, from 1% to 20%>, from 1% to 25%, from 1% to 50%, from 1% to 60%, from 1% to 70%, from 1% to 80%, from 1% to 90%, from 1% to 95%, from 10% to 20%, from 10% to 25%, from 10% to 50%, from 10% to 60%, from 10% to 70%, from 10% to 80%, from 10% to 90%, from 10% to 95%, from 10% to 100%, from 20% to 25%, from 20% to 50%, from 20% to 60%, from 20% to 70%, from 20% to 80%, from 20% to 90%, from 20% to 95%, from 20% to 100%, from 50% to 60%, from 50% to 70%, from 50% to 80%, from 50% to 90%, from 50% to 95%, from 50% to 100%, from 70% to 80%, from 70% to 90%, from 70% to 95%, from 70% to 100%, from 80% to 90%, from 80% to 95%, from 80% to 100%, from 90% to 95%, from 90% to 100%, and from 95% to 100%).

In some embodiments, sugar modifications (*e.g.*, at the 2' position or 4' position) or replacement of the sugar at one or more ribonucleotides of the sequence may, as well as backbone modifications, include modification or replacement of the phosphodiester linkages. Specific examples of a sequence include, but are not limited to, sequences including modified
5 backbones or no natural internucleoside linkages such as internucleoside modifications, including modification or replacement of the phosphodiester linkages. Sequences having modified backbones include, among others, those that do not have a phosphorus atom in the backbone. For the purposes of this application, and as sometimes referenced in the art, modified RNAs that do not have a phosphorus atom in their internucleoside backbone can also be
10 considered to be oligonucleosides. In particular embodiments, a sequence will include ribonucleotides with a phosphorus atom in its internucleoside backbone.

Modified sequence backbones may include, for example, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates such as 3'-alkylene phosphonates and chiral phosphonates,
15 phosphinates, phosphoramidates such as 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, and boranophosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity wherein the adjacent pairs of nucleoside units are linked 3'-5' to 5'-3' or 2'-5' to 5'-2'. Various salts, mixed salts and free acid forms are
20 also included. In some embodiments, the sequence may be negatively or positively charged.

The modified nucleotides, which may be incorporated into the sequence, can be modified on the internucleoside linkage (*e.g.*, phosphate backbone). Herein, in the context of the polynucleotide backbone, the phrases "phosphate" and "phosphodiester" are used interchangeably. Backbone phosphate groups can be modified by replacing one or more of the
25 oxygen atoms with a different substituent. Further, the modified nucleosides and nucleotides can include the wholesale replacement of an unmodified phosphate moiety with another internucleoside linkage as described herein. Examples of modified phosphate groups include, but are not limited to, phosphorothioate, phosphoroselenates, boranophosphates, boranophosphate esters, hydrogen phosphonates, phosphoramidates, phosphorodiamidates, alkyl or aryl
30 phosphonates, and phosphotriesters. Phosphorodithioates have both non-linking oxygens replaced by sulfur. The phosphate linker can also be modified by the replacement of a linking

oxygen with nitrogen (bridged phosphoramidates), sulfur (bridged phosphorothioates), and carbon (bridged methylene-phosphonates).

The α -thio substituted phosphate moiety is provided to confer stability to RNA and DNA polymers through the unnatural phosphorothioate backbone linkages. Phosphorothioate DNA and RNA have increased nuclease resistance and subsequently a longer half-life in a cellular environment.

In specific embodiments, a modified nucleoside includes an alpha-thio-nucleoside (*e.g.*, 5'-*O*-(1-thiophosphate)-adenosine, 5'-*O*-(1-thiophosphate)-cytidine (a-thio-cytidine), 5'-*O*-(1-thiophosphate)-guanosine, 5'-*O*-(1-thiophosphate)-uridine, or 5'-*O*-(1-thiophosphate)-pseudouridine).

Other internucleoside linkages that may be employed according to the present disclosure, including internucleoside linkages which do not contain a phosphorous atom, are described herein.

In some embodiments, the sequence may include one or more cytotoxic nucleosides. For example, cytotoxic nucleosides may be incorporated into sequence, such as bifunctional modification. Cytotoxic nucleoside may include, but are not limited to, adenosine arabinoside, 5-azacytidine, 4'-thio-aracytidine, cyclopentenylcytosine, cladribine, clofarabine, cytarabine, cytosine arabinoside, 1-(2-C-cyano-2-deoxy-beta-D-arabino-pentofuranosyl)-cytosine, decitabine, 5-fluorouracil, fludarabine, floxuridine, gemcitabine, a combination of tegafur and uracil, tegafur ((RS)-5-fluoro-1-(tetrahydrofuran-2-yl)pyrimidine-2,4(1H,3H)-dione), troxacitabine, tezacitabine, 2'-deoxy-2'-methylidenecytidine (DMDC), and 6-mercaptapurine. Additional examples include fludarabine phosphate, N4-behenoyl-1-beta-D-arabinofuranosylcytosine, N4-octadecyl-1-beta-D-arabinofuranosylcytosine, N4-palmitoyl-1-(2-C-cyano-2-deoxy-beta-D-arabino-pentofuranosyl) cytosine, and P-4055 (cytarabine 5'-elaidic acid ester).

In some embodiments, the sequence includes one or more post-transcriptional modifications (*e.g.*, capping, cleavage, polyadenylation, splicing, poly-A sequence, methylation, acylation, phosphorylation, methylation of lysine and arginine residues, acetylation, and nitrosylation of thiol groups and tyrosine residues, etc). The one or more post-transcriptional modifications can be any post-transcriptional modification, such as any of the more than one hundred different nucleoside modifications that have been identified in RNA (Rozenski, J, Crain, P, and McCloskey, J. (1999). The RNA Modification Database: 1999 update. Nucl Acids Res 27:

196-197) In some embodiments, the first isolated nucleic acid comprises messenger RNA (mRNA). In some embodiments, the mRNA comprises at least one nucleoside selected from the group consisting of pyridin-4-one ribonucleoside, 5-aza-uridine, 2-thio-5-aza-uridine, 2-thiouridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine, dihydropseudouridine, 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxyuridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, and 4-methoxy-2-thio-pseudouridine. In some embodiments, the mRNA comprises at least one nucleoside selected from the group consisting of 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, 4-methoxy-pseudoisocytidine, and 4-methoxy-1-methyl-pseudoisocytidine. In some embodiments, the mRNA comprises at least one nucleoside selected from the group consisting of 2-aminopurine, 2,6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl) adenosine, N6-glycinylnylcarbamoyl-adenosine, N6-threonylnylcarbamoyl-adenosine, 2-methylthio-N6-threonylcarbamoyl-adenosine, N6,N6-dimethyladenosine, 7-methyladenine, 2-methylthio-adenine, and 2-methoxy-adenine. In some embodiments, mRNA comprises at least one nucleoside selected from the group consisting of inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-

guanosine, 7-methyl-8-oxo-guanosine, 1-methyl-6-thio-guanosine, N2-methyl-6-thio-guanosine, and N2,N2-dimethyl-6-thio-guanosine.

The sequence may or may not be uniformly modified along the entire length of the molecule. For example, one or more or all types of nucleotides (*e.g.*, naturally-occurring
5 nucleotides, purine or pyrimidine, or any one or more or all of A, G, U, C, I, pU) may or may not be uniformly modified in the sequence, or in a given predetermined sequence region thereof. In some embodiments, the sequence includes a pseudouridine. In some embodiments, the sequence includes an inosine, which may aid in the immune system characterizing the sequence as endogenous versus viral RNAs. The incorporation of inosine may also mediate improved RNA
10 stability/reduced degradation. See for example, Yu, Z. et al. (2015) RNA editing by ADAR1 marks dsRNA as “self”. Cell Res. 25, 1283-1284, which is incorporated by reference in its entirety.

In some embodiments, one or more of the nucleotides of an RNA guide comprises a 2’-*O*-methyl phosphorothioate modification. In some embodiments, each of the first three
15 nucleotides of the RNA guide comprises a 2’-*O*-methyl phosphorothioate modification. In some embodiments, each of the last four nucleotides of the RNA guide comprises a 2’-*O*-methyl phosphorothioate modification. In some embodiments, each of the first to last, second to last, and third to last nucleotides of the RNA guide comprises a 2’-*O*-methyl phosphorothioate modification, and wherein the last nucleotide of the RNA guide is unmodified. In some
20 embodiments, each of the first three nucleotides of the RNA guide comprises a 2’-*O*-methyl phosphorothioate modification, and each of the first to last, second to last, and third to last nucleotides of the RNA guide comprises a 2’-*O*-methyl phosphorothioate modification.

When a gene editing system disclosed herein comprises nucleic acids encoding the Cas12i polypeptide disclosed herein, *e.g.*, mRNA molecules, such nucleic acid molecules may
25 contain any of the modifications disclosed herein, where applicable.

B. Cas12i Polypeptides

In some embodiments, the composition or system of the present disclosure includes a Cas12i polypeptide as described in WO/2019/178427, the relevant disclosures of which are
30 incorporated by reference for the subject matter and purpose referenced herein.

In some embodiments, the genetic editing system disclosed herein includes a Cas12i2 polypeptide described herein (*e.g.*, a polypeptide comprising SEQ ID NO: 922 and/or encoded

by SEQ ID NO: 921). In some embodiments, the Cas12i2 polypeptide comprises at least one RuvC domain.

A nucleic acid sequence encoding the Cas12i2 polypeptide described herein may be substantially identical to a reference nucleic acid sequence, *e.g.*, SEQ ID NO: 921. In some
5 embodiments, the Cas12i2 polypeptide is encoded by a nucleic acid comprising a sequence having least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or at least about 99.5% sequence identity to the reference
10 nucleic acid sequence, *e.g.*, SEQ ID NO: 921. The percent identity between two such nucleic acids can be determined manually by inspection of the two optimally aligned nucleic acid sequences or by using software programs or algorithms (*e.g.*, BLAST, ALIGN, CLUSTAL) using standard parameters. One indication that two nucleic acid sequences are substantially identical is that the nucleic acid molecules hybridize to the complementary sequence of the other
15 under stringent conditions of temperature and ionic strength (*e.g.*, within a range of medium to high stringency). *See, e.g.*, Tijssen, “Hybridization with Nucleic Acid Probes. Part I. Theory and Nucleic Acid Preparation” (Laboratory Techniques in Biochemistry and Molecular Biology, Vol 24).

In some embodiments, the Cas12i2 polypeptide is encoded by a nucleic acid sequence
20 having at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or more sequence identity, but not 100% sequence identity, to a reference nucleic acid sequence, *e.g.*, SEQ ID NO: 921.

25 In some embodiments, the Cas12i2 polypeptide of the present disclosure comprises a polypeptide sequence having at least 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 922.

In some embodiments, the present disclosure describes a Cas12i2 polypeptide having a specified degree of amino acid sequence identity to one or more reference polypeptides, *e.g.*, at
30 least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or even at least 99%, but not 100%, sequence identity to the amino acid sequence of

SEQ ID NO: 922. Homology or identity can be determined by amino acid sequence alignment, *e.g.*, using a program such as BLAST, ALIGN, or CLUSTAL, as described herein.

Also provided is a Cas12i2 polypeptide of the present disclosure having enzymatic activity, *e.g.*, nuclease or endonuclease activity, and comprising an amino acid sequence which
5 differs from the amino acid sequences of SEQ ID NO: 922 by 50, 40, 35, 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, or 0 amino acid residue(s), when aligned using any of the previously described alignment methods.

In some examples, the Cas12i2 polypeptide may contain one or more mutations relative to SEQ ID NO: 922, for example, at position D581, G624, F626, P868, I926, V1030, E1035,
10 S1046, or any combination thereof. In some instances, the one or more mutations are amino acid substitutions, for example, D581R, G624R, F626R, P868T, I926R, V1030G, E1035R, S1046G, or a combination thereof.

In some examples, the Cas12i2 polypeptide contains mutations at positions D581, D911, I926, and V1030. Such a Cas12i2 polypeptide may contain amino acid substitutions of D581R,
15 D911R, I926R, and V1030G (*e.g.*, SEQ ID NO: 923). In some examples, the Cas12i2 polypeptide contains mutations at positions D581, I926, and V1030. Such a Cas12i2 polypeptide may contain amino acid substitutions of D581R, I926R, and V1030G (*e.g.*, SEQ ID NO: 924). In some examples, the Cas12i2 polypeptide may contain mutations at positions D581, I926, V1030, and S1046. Such a Cas12i2 polypeptide may contain amino acid substitutions of D581R,
20 I926R, V1030G, and S1046G (*e.g.*, SEQ ID NO: 925). In some examples, the Cas12i2 polypeptide may contain mutations at positions D581, G624, F626, I926, V1030, E1035, and S1046. Such a Cas12i2 polypeptide may contain amino acid substitutions of D581R, G624R, F626R, I926R, V1030G, E1035R, and S1046G (*e.g.*, SEQ ID NO: 926). In some examples, the Cas12i2 polypeptide may contain mutations at positions D581, G624, F626, P868, I926, V1030,
25 E1035, and S1046. Such a Cas12i2 polypeptide may contain amino acid substitutions of D581R, G624R, F626R, P868T, I926R, V1030G, E1035R, and S1046G (*e.g.*, SEQ ID NO: 927).

In some embodiments, the Cas12i2 polypeptide of the present disclosure comprises a polypeptide sequence having at least 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 923, SEQ ID NO: 924,
30 SEQ ID NO: 925, SEQ ID NO: 926, or SEQ ID NO: 927. In some embodiments, a Cas12i2 polypeptide having at least 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 923, SEQ ID NO: 924, SEQ ID

NO: 925, SEQ ID NO: 926, or SEQ ID NO: 927 maintains the amino acid changes (or at least 1, 2, 3 etc. of these changes) that differentiate the polypeptide from its respective parent/reference sequence.

In some embodiments, the present disclosure describes a Cas12i2 polypeptide having a specified degree of amino acid sequence identity to one or more reference polypeptides, *e.g.*, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or even at least 99%, but not 100%, sequence identity to the amino acid sequence of SEQ ID NO: 923, SEQ ID NO: 924, SEQ ID NO: 925, SEQ ID NO: 926, or SEQ ID NO: 927.

Homology or identity can be determined by amino acid sequence alignment, *e.g.*, using a program such as BLAST, ALIGN, or CLUSTAL, as described herein.

Also provided is a Cas12i2 polypeptide of the present disclosure having enzymatic activity, *e.g.*, nuclease or endonuclease activity, and comprising an amino acid sequence which differs from the amino acid sequences of SEQ ID NO: 923, SEQ ID NO: 924, SEQ ID NO: 925, SEQ ID NO: 926, or SEQ ID NO: 927 by 50, 40, 35, 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, or 0 amino acid residue(s), when aligned using any of the previously described alignment methods.

In some embodiments, the composition of the present disclosure includes a Cas12i4 polypeptide described herein (*e.g.*, a polypeptide comprising SEQ ID NO: 956 and/or encoded by SEQ ID NO: 955). In some embodiments, the Cas12i4 polypeptide comprises at least one RuvC domain.

A nucleic acid sequence encoding the Cas12i4 polypeptide described herein may be substantially identical to a reference nucleic acid sequence, *e.g.*, SEQ ID NO: 955. In some embodiments, the Cas12i4 polypeptide is encoded by a nucleic acid comprising a sequence having least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or at least about 99.5% sequence identity to the reference nucleic acid sequence, *e.g.*, SEQ ID NO: 955. The percent identity between two such nucleic acids can be determined manually by inspection of the two optimally aligned nucleic acid sequences or by using software programs or algorithms (*e.g.*, BLAST, ALIGN, CLUSTAL) using standard parameters. One indication that two nucleic acid sequences are substantially

identical is that the nucleic acid molecules hybridize to the complementary sequence of the other under stringent conditions of temperature and ionic strength (*e.g.*, within a range of medium to high stringency).

In some embodiments, the Cas12i4 polypeptide is encoded by a nucleic acid sequence having at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or more sequence identity, but not 100% sequence identity, to a reference nucleic acid sequence, *e.g.*, SEQ ID NO: 955.

In some embodiments, the Cas12i4 polypeptide of the present disclosure comprises a polypeptide sequence having at least 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 956.

In some embodiments, the present disclosure describes a Cas12i4 polypeptide having a specified degree of amino acid sequence identity to one or more reference polypeptides, *e.g.*, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or even at least 99%, but not 100%, sequence identity to the amino acid sequence of SEQ ID NO: 956. Homology or identity can be determined by amino acid sequence alignment, *e.g.*, using a program such as BLAST, ALIGN, or CLUSTAL, as described herein.

Also provided is a Cas12i4 polypeptide of the present disclosure having enzymatic activity, *e.g.*, nuclease or endonuclease activity, and comprising an amino acid sequence which differs from the amino acid sequences of SEQ ID NO: 956 by 50, 40, 35, 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, or 0 amino acid residue(s), when aligned using any of the previously described alignment methods.

In some embodiments, the Cas12i4 polypeptide comprises a polypeptide having a sequence of SEQ ID NO: 957 or SEQ ID NO: 958.

In some embodiments, the Cas12i4 polypeptide of the present disclosure comprises a polypeptide sequence having at least 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 957 or SEQ ID NO: 958. In some embodiments, a Cas12i4 polypeptide having at least 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ

ID NO: 957 or SEQ ID NO: 958 maintains the amino acid changes (or at least 1, 2, 3 etc. of these changes) that differentiate it from its respective parent/reference sequence.

In some embodiments, the present disclosure describes a Cas12i4 polypeptide having a specified degree of amino acid sequence identity to one or more reference polypeptides, *e.g.*, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or even at least 99%, but not 100%, sequence identity to the amino acid sequence of SEQ ID NO: 957 or SEQ ID NO: 958. Homology or identity can be determined by amino acid sequence alignment, *e.g.*, using a program such as BLAST, ALIGN, or CLUSTAL, as described herein.

Also provided is a Cas12i4 polypeptide of the present disclosure having enzymatic activity, *e.g.*, nuclease or endonuclease activity, and comprising an amino acid sequence which differs from the amino acid sequences of SEQ ID NO: 957 or SEQ ID NO: 958 by 50, 40, 35, 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, or 0 amino acid residue(s), when aligned using any of the previously described alignment methods.

In some embodiments, the composition of the present disclosure includes a Cas12i1 polypeptide described herein (*e.g.*, a polypeptide comprising SEQ ID NO: 965). In some embodiments, the Cas12i4 polypeptide comprises at least one RuvC domain.

In some embodiments, the Cas12i1 polypeptide of the present disclosure comprises a polypeptide sequence having at least 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 965.

In some embodiments, the present disclosure describes a Cas12i1 polypeptide having a specified degree of amino acid sequence identity to one or more reference polypeptides, *e.g.*, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or even at least 99%, but not 100%, sequence identity to the amino acid sequence of SEQ ID NO: 965. Homology or identity can be determined by amino acid sequence alignment, *e.g.*, using a program such as BLAST, ALIGN, or CLUSTAL, as described herein.

Also provided is a Cas12i1 polypeptide of the present disclosure having enzymatic activity, *e.g.*, nuclease or endonuclease activity, and comprising an amino acid sequence which differs from the amino acid sequences of SEQ ID NO: 965 by 50, 40, 35, 30, 25, 20, 19, 18, 17,

16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, or 0 amino acid residue(s), when aligned using any of the previously described alignment methods.

In some embodiments, the composition of the present disclosure includes a Cas12i3 polypeptide described herein (*e.g.*, a polypeptide comprising SEQ ID NO: 966). In some
5 embodiments, the Cas12i4 polypeptide comprises at least one RuvC domain.

In some embodiments, the Cas12i3 polypeptide of the present disclosure comprises a polypeptide sequence having at least 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 966.

In some embodiments, the present disclosure describes a Cas12i3 polypeptide having a
10 specified degree of amino acid sequence identity to one or more reference polypeptides, *e.g.*, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or even at least 99%, but not 100%, sequence identity to the amino acid sequence of SEQ ID NO: 966. Homology or identity can be determined by amino acid sequence alignment,
15 *e.g.*, using a program such as BLAST, ALIGN, or CLUSTAL, as described herein.

Also provided is a Cas12i3 polypeptide of the present disclosure having enzymatic activity, *e.g.*, nuclease or endonuclease activity, and comprising an amino acid sequence which differs from the amino acid sequences of SEQ ID NO: 966 by 50, 40, 35, 30, 25, 20, 19, 18, 17,
16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, or 0 amino acid residue(s), when aligned using
20 any of the previously described alignment methods.

Although the changes described herein may be one or more amino acid changes, changes to the Cas12i polypeptide may also be of a substantive nature, such as fusion of polypeptides as amino- and/or carboxyl-terminal extensions. For example, the Cas12i polypeptide may contain additional peptides, *e.g.*, one or more peptides. Examples of additional peptides may include
25 epitope peptides for labelling, such as a polyhistidine tag (His-tag), Myc, and FLAG. In some embodiments, the Cas12i polypeptide described herein can be fused to a detectable moiety such as a fluorescent protein (*e.g.*, green fluorescent protein (GFP) or yellow fluorescent protein (YFP)).

In some embodiments, the Cas12i polypeptide comprises at least one (*e.g.*, two, three,
30 four, five, six, or more) nuclear localization signal (NLS). In some embodiments, the Cas12i polypeptide comprises at least one (*e.g.*, two, three, four, five, six, or more) nuclear export signal

(NES). In some embodiments, the Cas12i polypeptide comprises at least one (*e.g.*, two, three, four, five, six, or more) NLS and at least one (*e.g.*, two, three, four, five, six, or more) NES.

In some embodiments, the Cas12i polypeptide described herein can be self-inactivating. See, Epstein et al., “Engineering a Self-Inactivating CRISPR System for AAV Vectors,” *Mol. Ther.*, 24 (2016): S50, which is incorporated by reference in its entirety.

In some embodiments, the nucleotide sequence encoding the Cas12i polypeptide described herein can be codon-optimized for use in a particular host cell or organism. For example, the nucleic acid can be codon-optimized for any non-human eukaryote including mice, rats, rabbits, dogs, livestock, or non-human primates. Codon usage tables are readily available, for example, at the “Codon Usage Database” available at www.kazusa.or.jp/codon/ and these tables can be adapted in a number of ways. See Nakamura *et al. Nucl. Acids Res.* 28:292 (2000), which is incorporated herein by reference in its entirety. Computer algorithms for codon optimizing a particular sequence for expression in a particular host cell are also available, such as Gene Forge (Aptagen; Jacobus, PA). In some examples, the nucleic acid encoding the Cas12i polypeptides such as Cas12i2 polypeptides as disclosed herein can be an mRNA molecule, which can be codon optimized.

Exemplary Cas12i polypeptide sequences and corresponding nucleotide sequences are listed in **Table 6**.

Table 6. Cas12i and HAO1 Sequences

SEQ ID NO	Sequence	Description
921	ATGAGCAGCGCGATCAAAAGCTACAAGAGCGTTCTGCGTCCGAACGAGCGTAAGAA CCAACTGCTGAAAAGCACCATTTCAGTGCCTGGAAGACGGTAGCGGTTCTTTTCA AGATGCTGCAAGGCCTGTTTGGTGGCATCACCCGGAGATTGTTGTTTTCAGCACC GAACAGGAGAAACAGCAACAGGATATCGCGCTGTGGTGCGCGGTTAACTGGTTCCG TCCGGTGAGCCAAGACAGCCTGACCCACACCATTGCGAGCGATAACCTGGTGGAGA AGTTTGAGGAATACTATGGTGGCACC GCGAGCGACGCGATCAAACAGTACTTCAGC GCGAGCATTGGCGAAAGCTACTATTGGAACGACTGCCGTCAACAGTACTATGATCT GTGCCGTGAGCTGGGTGTTGAGGTGAGCGACCTGACCCATGATCTGGAGATCCTGT GCCGTGAAAAGTGCCTGGCGGTTGCGACCGAGAGCAACCAGAACACAGCATCATT AGCGTTCTGTTTGGCACCGGCGAAAAAGAGGACCGTAGCGTGAAACTGCGTATCAC CAAGAAAATTCTGGAGGCGATCAGCAACCTGAAAGAAATCCCGAAGAACGTTGCGC CGATTCAAGAGATCATTCTGAACGTGGCGAAAGCGACCAAGGAAACCTTCCGTCAG GTGTATGCGGGTAACCTGGGTGCGCCGAGCACCTGGAGAAATTTATCGCGAAGGA CGGCCAAAAGAGTTGATCTGAAGAACTGCAGACCGACCTGAAGAAAGTTATTC GTGGTAAAAGCAAGGAGCGTGATTGGTGCTGCCAGGAAGAGCTGCGTAGCTACGTG GAGCAAAACACCATCCAGTATGACCTGTGGGCGTGGGGCGAAATGTTCAACAAAGC GCACACCGCGCTGAAATCAAGAGACCCGTAACCTACAACCTTTCGGAAGCAACGTC TGGAACAGTTCAAAGAGATTAGAGCCTGAACACCTGCTGGTTGTGAAGAAGCTG	Nucleotide sequence encoding parent Cas12i2

	AACGACTTTTTCGATAGCGAATTTTTCAGCGGCGAGGAAACCTACACCATCTGCGT TCACCATCTGGGTGGCAAGGACCTGAGCAAACCTGTATAAGGCGTGGGAGGATGATC CGGCGGACCCGGAAAACGCGATTGTGGTTCTGTGCGACGATCTGAAAAACAACCTTT AAGAAAGAGCCGATCCGTAAACATTCTGCGTTACATCTTCACCATTCGTCAAGAATG CAGCGCGCAGGACATCCTGGCGGCGGCGAAGTACAACCAACAGCTGGATCGTTATA AAAGCCAAAAGGCGAACCCGAGCGTTCTGGGTAAACAGGGCTTTACCTGGACCAAC GCGGTGATCCTGCGGAGAAAGGCGCAGCGTAACGACCGTCCGAACAGCCTGGATCT GCGTATTGGCTGTACCTGAAACTGCGTCACCCGGACGCTCGTTGGAAGAAACACC ATATCCCGTTCTACGATACCCGTTTCTTCCAAGAAATTTATGCGGCGGGCAACAGC CCGGTTGACACCTGCCAGTTTCTGTAACCCGCGTTTCGGTTATCACCTGCCGAAACT GACCGATCAGACCCGATCCGTGTTAACAAGAAACATGTGAAAGCGGCGAAGACCG AGGCGCGTATTCGTCTGGCGATCCAACAGGGCACCTGCCGGTGAGCAACCTGAAG ATCACCGAAATTAGCGCGACCATCAACAGCAAAGGTCAAGTGCCTATTCCGGTTAA GTTTGACGTGGTCTGCAAAAAGGCACCCCTGCAGATCGGTGACCGTTTCTGCGGCT ACGATCAAAACCAGACCGCGAGCCACGCGTATAGCCTGTGGGAAAGTGGTTAAAGAG GGTCAATACCATAAAGAGCTGGGCTGCTTTGTTTCGTTTCATCAGCAGCGGTGACAT CGTGAGCATTACCGAGAACCGTGGAACCAATTTGATCAGCTGAGCTATGAAGGTC TGGCGTACCCGCAATATGCGGACTGGCGTAAGAAAGCGAGCAAGTTCGTGAGCCCTG TGGCAGATCACCAGAAAAACAAGAAAAAGGAAATCGTGACCGTTGAAGCGAAAGA GAAGTTTGACGCGATCTGCAAGTACCAGCCGCGTCTGTATAAATTCAACAAGGAGT ACGCGTATCTGCTGCGTGATATTGTTCTGTGGCAAAGCCTGGTGGAAGTGAACAG ATTCTGCAAGAGATCTTTCGTTTCATTGAACAGGACTGCGGTGTTACCCGCTCGGG CAGCCTGAGCCTGAGCACCCCTGGAACCGTGAAAGCGGTAAAGGGTATCATTACA GCTATTTTAGCACCGCGCTGAACGCGAGCAAGAACAACCCGATCAGCGACGAACAG CGTAAAGAGTTTGTATCCGGAACCTGTCGCGCTGCTGGAAAAGCTGGAGCTGATTCTG TACCCGTAAAAAGAAACAAAAAGTGAACGTATCGGAACAGCCTGATTACAGACCT GCCTGGAGAACAACATCAAGTTCATTCTGTGGTGAAGGCGACCTGAGCACCAAC AACGCGACCAAGAAAAAGGCGAACAGCCGTAGCATGGATTGGTTGGCGCGTGGTGT TTTTAACAAAATCCGTCAACTGGCGCCGATGCACAACATTACCCTGTTTCGGTTGCG GCAGCCTGTACACCAGCCACCAGGACCCGCTGGTGATCGTAACCCGGATAAAGCG ATGAAGTGCCGTTGGCGGCGATCCCGGTTAAGGACATTGGCGATTGGGTGCTGCG TAAGCTGAGCCAAAACCTGCGTGCAGAAAAACATCGGCACCCGCGAGTACTATCACC AAGGTGTTAAAGAGTTCCTGAGCCATTATGAAGTGCAGGACCTGGAGGAAGAGCTG CTGAAGTGGCGTAGCGATCGTAAAGCAACATTCCGTGCTGGGTGCTGCAGAACCG TCTGCGCGGAGAAGCTGGGCAACAAAGAAGCGGTGGTTTACATCCCGGTTCTGGTG GCCGTATTTATTTTTCGACCCACAAGGTGGCGACCCGGTGCCTGAGCATCGTTTTC GACCAAAAAACAAGTGTGGGTTTCAACGCGGATCATGTTGCGGCGGCGAACATCGC GCTGACCGTGAAGGGTATTGGCGAACAAAGCAGCGACGAAGAGAACCCGGATGGTA GCCGTATCAAACTGCAGCTGACCAGC	
922	MSSAIKSYKSVLRPNERNKQLLKSTIQCLEDSAFFFKMLQGLFGGITPEIVRFST EQEKQQQDIALWCANWFRPVSQDSLTHTIASDNLVEKFEEYYGGTASDAIKQYFS ASIGESYYWDCRQQYYDLCRELGVEVSDLTHDLEILCREKCLAVATESNQNNISII SVLFGTGEKEDRSVKLRITKKILEAISNLKEIPKNVAPIQEIILNVAKATKETFRQ VYAGNLGAPSTLEKFIKDGQKEFDLKKLQTDLKKVIRGKSKERDWCCQEELRSYV EQNTIQYDLWAWGEMFNKAHTALKIKSTRNYNFAQRLEQFKEIQSLNNLLVVKKL NDFFDSEFFSGEETYITCVHHLGGKDLKLYKAWEDDPADPENAIIVLCLDDLKNNF KKEPIRNILRYIFTIRQECQAQDILAAAKYNQQLDRYKSQKANPSVLGNQGFWTWN AVILPEKAQRNDRPNSLDLRIWLYLKLRLHPDGRWKKHHIPFYDTRFFQEIYAAGNS PVDTCQFRTPRFGYHLPKLTDQTAIRVNKKHVKAATEARIRLAIQQGTLPSVNLK ITEISATINSKGQVRIPVKFDVGRQKGTQLQIGDRFCGYDQNTASHAYSLWEVVKE GQYHKELGCFVRFISSGDIVSITENRGNQFDQLSYEGLAYPQYADWRKKASKFVSL WQITKKNKKKEIVTVEAKEKFDAICKYQPRLYKFNKEYAYLLRDIVRGKSLVELQQ IRQEIFRFIEQDCGVTRLGSLSLSTLETVKAVKGIIYSYFSTALNASKNNPISDEQ RKEFDELFALLEKLELIRTRKKKQKVERIANSLIQTCLENNIKFIRGEGLSTTN NATKKKANRSMDWLARGVFNKIRQLAPMHNITLFGCGSLYTSHQDPLVHRNPDKA MKCRWAAIPVKDIGDWVLRKLSQNLRAKNIGTGEYYHQGVKEFLSHYELQDLEEL LKWRSRDKSNIPCVWLQNRLAEKLGNEAVVYIPVRGGRIYFATHKVATGAVSIVF DQKQVWVCNADHVAAANIALTVKGIGEQSSDEENPDGSRIKLQLTS	Parent Cas12i2 amino acid sequence

923	MSSAIKSYKS VLRPNERKNQ LLKSTIQCLE DGSAFFFKML QGLFGGITPE IVRFSTEQEK QQQDIALWCA VNWFRPVSD SLTHTIASDN LVEKFEEYYG GTASDAIKQY FSASIGESYY WNDRCRQYYD LCRELGVEVS DLTHDLEILC REKCLAVATE SNQNNISIISV LFGTGEKEDR SVKLRITKKI LEAISNLKEI PKNVAPIQEI ILNVAKATKE TFRQVYAGNL GAPSTLEKFI AKDGQKEFDL KKLQTDLKKV IRGKSKERDW CCQEELRSYV EQNTIQYDLW AWGEMFNKAH TALKIKSTRN YNFAKQRLAQ FKEIQSLNNL LVVKKLNDFD DSEFFSGEET YTICVHHLGG KDLSKLYKAW EDDPADPENA IVVLCDDLKN NFKKEPIRNI LRYIFTIRQE CSAQDILAAA KYNQQLDRYK SQKANPSVLG NQGFTWNAV ILPEKAQRND RPNSLDLRIW LYLKLRHPDG RWKKHHIPFY DTRFFQEIYA AGNSPVDTCQ FRTPRFGYHL PKLTDQTAIR VNKKHVKAAC TEARIRLAIQ QGTLPVSNLK ITEISATINS KGQVRIPVKF RVGRQKGTLO IGDRFCGYDQ NQTASHAYSL WEVVKEGQYH KELGCFVRFI SSGDIVSITE NRGNQFDQLS YEGLAYPQYA DWRKKASKFV SLWQITKKNK KKEIVTVEAK EKFDAICKYQ PRLYKFNKEY AYLLRDIVRG KSLVELQQIR QEIFRFIEQD CGVTRLGSL LSTLETVKAV KGIIYSYFST ALNASKNNPI SDEQRKEFDP ELFALLEKLE LIRTRKKKQK VERIANSLIQ TCLENNIKFI RGEGLDSTTN NATKKKANSR SMDWLARGVF NKIRQLAPMH NITLFGCGSL YTSHQDPLVH RNPDKAMKCR WAAIPVKDIG RWVLRKLSQN LRAKNRGTGE YYHQGVKEFL SHYELQDLEE ELLKWRSDRK SNIPCWVLQN RLAEKLGKNE AVVYIPVRGG RIYFATHKVA TGAVSIVFDQ KQVWVCNADH VAAANIALTG KGIGEQQSSDE ENPDGSRILK QLTS	Variant Cas12i2 of SEQ ID NO: 3 of PCT/US2021/ 025257
924	MSSAIKSYKS VLRPNERKNQ LLKSTIQCLE DGSAFFFKML QGLFGGITPE IVRFSTEQEK QQQDIALWCA VNWFRPVSD SLTHTIASDN LVEKFEEYYG GTASDAIKQY FSASIGESYY WNDRCRQYYD LCRELGVEVS DLTHDLEILC REKCLAVATE SNQNNISIISV LFGTGEKEDR SVKLRITKKI LEAISNLKEI PKNVAPIQEI ILNVAKATKE TFRQVYAGNL GAPSTLEKFI AKDGQKEFDL KKLQTDLKKV IRGKSKERDW CCQEELRSYV EQNTIQYDLW AWGEMFNKAH TALKIKSTRN YNFAKQRLAQ FKEIQSLNNL LVVKKLNDFD DSEFFSGEET YTICVHHLGG KDLSKLYKAW EDDPADPENA IVVLCDDLKN NFKKEPIRNI LRYIFTIRQE CSAQDILAAA KYNQQLDRYK SQKANPSVLG NQGFTWNAV ILPEKAQRND RPNSLDLRIW LYLKLRHPDG RWKKHHIPFY DTRFFQEIYA AGNSPVDTCQ FRTPRFGYHL PKLTDQTAIR VNKKHVKAAC TEARIRLAIQ QGTLPVSNLK ITEISATINS KGQVRIPVKF RVGRQKGTLO IGDRFCGYDQ NQTASHAYSL WEVVKEGQYH KELGCFVRFI SSGDIVSITE NRGNQFDQLS YEGLAYPQYA DWRKKASKFV SLWQITKKNK KKEIVTVEAK EKFDAICKYQ PRLYKFNKEY AYLLRDIVRG KSLVELQQIR QEIFRFIEQD CGVTRLGSL LSTLETVKAV KGIIYSYFST ALNASKNNPI SDEQRKEFDP ELFALLEKLE LIRTRKKKQK VERIANSLIQ TCLENNIKFI RGEGLDSTTN NATKKKANSR SMDWLARGVF NKIRQLAPMH NITLFGCGSL YTSHQDPLVH RNPDKAMKCR WAAIPVKDIG DWVLRKLSQN LRAKNRGTGE YYHQGVKEFL SHYELQDLEE ELLKWRSDRK SNIPCWVLQN RLAEKLGKNE AVVYIPVRGG RIYFATHKVA TGAVSIVFDQ KQVWVCNADH VAAANIALTG KGIGEQQSSDE ENPDGSRILK QLTS	Variant Cas12i2 of SEQ ID NO: 4 of PCT/US2021/ 025257
925	MSSAIKSYKS VLRPNERKNQ LLKSTIQCLE DGSAFFFKML QGLFGGITPE IVRFSTEQEK QQQDIALWCA VNWFRPVSD SLTHTIASDN LVEKFEEYYG GTASDAIKQY FSASIGESYY WNDRCRQYYD LCRELGVEVS DLTHDLEILC REKCLAVATE SNQNNISIISV LFGTGEKEDR SVKLRITKKI LEAISNLKEI PKNVAPIQEI ILNVAKATKE TFRQVYAGNL GAPSTLEKFI AKDGQKEFDL KKLQTDLKKV IRGKSKERDW CCQEELRSYV EQNTIQYDLW AWGEMFNKAH TALKIKSTRN YNFAKQRLAQ FKEIQSLNNL LVVKKLNDFD DSEFFSGEET YTICVHHLGG KDLSKLYKAW EDDPADPENA IVVLCDDLKN NFKKEPIRNI LRYIFTIRQE CSAQDILAAA KYNQQLDRYK SQKANPSVLG NQGFTWNAV ILPEKAQRND RPNSLDLRIW LYLKLRHPDG RWKKHHIPFY DTRFFQEIYA AGNSPVDTCQ FRTPRFGYHL PKLTDQTAIR VNKKHVKAAC TEARIRLAIQ QGTLPVSNLK ITEISATINS KGQVRIPVKF RVGRQKGTLO IGDRFCGYDQ NQTASHAYSL WEVVKEGQYH KELGCFVRFI SSGDIVSITE NRGNQFDQLS YEGLAYPQYA DWRKKASKFV SLWQITKKNK KKEIVTVEAK EKFDAICKYQ PRLYKFNKEY AYLLRDIVRG KSLVELQQIR QEIFRFIEQD CGVTRLGSL	Variant Cas12i2 of SEQ ID NO: 5 of PCT/US2021/ 025257

	LSTLETVKAV KGIISYFST ALNASKNNPI SDEQRKEFDP ELFALLEKLE LIRTRKKQK VERIANSLIQ TCLENNIKFI RGEGLSTTN NATKKKANSR SMDWLARGVF NKIRQLAPMH NITLFGCGSL YTSHQDPLVH RNPDKAMKCR WAAIPVKDIG DWVLRKLSQN LRAKNRGTGE YYHQGVKEFL SHYELQDLEE ELLKWRSDRK SNIPCWLQN RLAEKLGKE AVVYIPVRGG RIYFATHKVA TGAVSIVFDQ KQVWVCNADH VAAANIALTG KGIGEQSSDE ENPDGGRIKL QLTS	
926	MSSAISKYSK VLRPNERKNQ LLKSTIQCLE DGSAFFFKML QGLFGGITPE IVRFSTEQEK QQQDIALWCA VNWFRPVSQD SLTHTIASDN LVEKFEEYYG GTASDAIKQY FSASIGESYY WNDCRQQYYD LCRELGVEVS DLTHDLEILC REKCLAVATE SNQNNISIISV LFGTGEKEDR SVKLRITKKI LEAISNLKEI PKNVAPIQEI ILNVAKATKE TFRQVYAGNL GAPSTLEKFI AKDGQKEFDL KKLQTDLKKV IRGKSKERDW CCQEELRSYV EQNTIQYDLW AWGEMFNKAH TALKIKSTRN YNFAKQRLAQ FKEIQSLNNL LVVKKLNDDF DSEFFSGEET YTICVHHLGG KDLSKLYKAW EDDPADPENA IVVLCDDLKN NFKKEPIRNI LRYIFTIRQE CSAQDILAAA KYNQQLDRYK SQKANPSVLG NQGFTWNAV ILPEKAQRND RPNSLDLRIW LYLKLRHPDG RWKKHHIPFY DTRFFQEIYA AGNSPVDTCQ FRTPRFGYHL PKLTDQTAIR VNKKHVKAAC TEARIRLAIQ QGTLPVSNLK ITEISATINS KGQVRIPVKF RVGRQKGTLO IGDRFCGYDQ NQTASHAYSL WEVVKEGQYH KELRCRVRFI SSGDIVSITE NRGNQFDQLS YEGLAYPQYA DWRKKASKFV SLWQITKKNK KKEIVTVEAK EKFDAICKYQ PRLYKFKEY AYLLRDIVRG KSLVELQQIR QEIFRFIEQD CGVTRLGSL LSTLETVKAV KGIISYFST ALNASKNNPI SDEQRKEFDP ELFALLEKLE LIRTRKKQK VERIANSLIQ TCLENNIKFI RGEGLSTTN NATKKKANSR SMDWLARGVF NKIRQLAPMH NITLFGCGSL YTSHQDPLVH RNPDKAMKCR WAAIPVKDIG DWVLRKLSQN LRAKNRGTGE YYHQGVKEFL SHYELQDLEE ELLKWRSDRK SNIPCWLQN RLAEKLGKE AVVYIPVRGG RIYFATHKVA TGAVSIVFDQ KQVWVCNADH VAAANIALTG KGIGRQSSDE ENPDGGRIKL QLTS	Variant Cas12i2 of SEQ ID NO: 495 of PCT/US2021/ 025257
927	MSSAISKYSK VLRPNERKNQ LLKSTIQCLE DGSAFFFKML QGLFGGITPE IVRFSTEQEK QQQDIALWCA VNWFRPVSQD SLTHTIASDN LVEKFEEYYG GTASDAIKQY FSASIGESYY WNDCRQQYYD LCRELGVEVS DLTHDLEILC REKCLAVATE SNQNNISIISV LFGTGEKEDR SVKLRITKKI LEAISNLKEI PKNVAPIQEI ILNVAKATKE TFRQVYAGNL GAPSTLEKFI AKDGQKEFDL KKLQTDLKKV IRGKSKERDW CCQEELRSYV EQNTIQYDLW AWGEMFNKAH TALKIKSTRN YNFAKQRLAQ FKEIQSLNNL LVVKKLNDDF DSEFFSGEET YTICVHHLGG KDLSKLYKAW EDDPADPENA IVVLCDDLKN NFKKEPIRNI LRYIFTIRQE CSAQDILAAA KYNQQLDRYK SQKANPSVLG NQGFTWNAV ILPEKAQRND RPNSLDLRIW LYLKLRHPDG RWKKHHIPFY DTRFFQEIYA AGNSPVDTCQ FRTPRFGYHL PKLTDQTAIR VNKKHVKAAC TEARIRLAIQ QGTLPVSNLK ITEISATINS KGQVRIPVKF RVGRQKGTLO IGDRFCGYDQ NQTASHAYSL WEVVKEGQYH KELRCRVRFI SSGDIVSITE NRGNQFDQLS YEGLAYPQYA DWRKKASKFV SLWQITKKNK KKEIVTVEAK EKFDAICKYQ PRLYKFKEY AYLLRDIVRG KSLVELQQIR QEIFRFIEQD CGVTRLGSL LSTLETVKAV KGIISYFST ALNASKNNPI SDEQRKEFDP ELFALLEKLE LIRTRKKQK VERIANSLIQ TCLENNIKFI RGEGLSTTN NATKKKANSR SMDWLARGVF NKIRQLATMH NITLFGCGSL YTSHQDPLVH RNPDKAMKCR WAAIPVKDIG DWVLRKLSQN LRAKNRGTGE YYHQGVKEFL SHYELQDLEE ELLKWRSDRK SNIPCWLQN RLAEKLGKE AVVYIPVRGG RIYFATHKVA TGAVSIVFDQ KQVWVCNADH VAAANIALTG KGIGRQSSDE ENPDGGRIKL QLTS	Variant Cas12i2 of SEQ ID NO: 496 of PCT/US2021/ 025257
955	ATGGCTTCCATCTCTAGGCCATACGGCACCAAGCTGCGACCGGACGCACGGAAGAA GGAGATGCTCGATAAGTTCTTTAATACACTGACTAAGGGTCAGCGCGTGTTCGCAG ACCTGGCCCTGTGCATCTATGGCTCCCTGACCTGGAGATGGCCAAGTCTCTGGAG CCAGAAAGTGATTCAGAACTGGTGTGCGCTATTGGGTGGTTTCGGCTGGTGACAA GACCATCTGGTCCAAGGATGGCATCAAGCAGGAGAATCTGGTGAAACAGTACGAAG CCTATTCGGAAAGGAGGCTTCTGAAGTGGTCAAACATACCTGAACAGCCCCAGC TCCGACAAGTACGTGTGGATCGATTGCAGGCAGAAATTCCTGAGGTTTCAGCGCGA	Nucleotide sequence encoding

	GCTCGGCACTCGCAACCTGTCCGAGGACTTCGAATGTATGCTCTTTGAACAGTACA TTAGACTGACCAAGGGCGAGATCGAAGGGTATGCCGCTATTTCAAATATGTTGGA AACGGCGAGAAGGAAGACCGGAGCAAGAAAAGAAATGTACGCTACACGGATGAAAGA TTGGCTGGAGGCAAACGAAAATATCACTTGGGAGCAGTATAGAGAGGCCCTGAAGA ACCAGCTGAATGCTAAAAACCTGGAGCAGGTTGTGGCCAATTACAAGGGGAACGCT GGCGGGGCGAGACCCCTTCTTTAAGTATAGCTTCTCCAAAGAGGGAATGGTGAGCAA GAAAGAACATGCACAGCAGCTCGACAAGTTCAAACCGCTGAGCAAACCTCTATAAGATCT GGGACCTGAATTTTCAAACAAGGAGAAGCTGAAGCAGTACCTGGAGGCCGAAATC GGCATTCCGGTTCGACGCTAACGTGTACTCCCAGATGTTCTTAACGGGGTGAGTGA GGTCCAGCCTAAGACCACACGGAATATGTCTTTTAGTAACGAGAACTGGATCTGC TCACTGAACTGAAGGACCTGAACAAGGGCGATGGGTTTCAGTACGCCAGAGAAGTG CTGAACGGGTTCTTTGACTCCGAGCTCCACACTACCGAGGATAAGTTTAAATATCAC CTCTAGGTACCTGGGAGGCGACAAATCAAACCGCTGAGCAAACCTCTATAAGATCT GGAAGAAAGAGGGTGTGGACTGCGAGGAAGGCATTACAGCAGTTCTGTGAAGCCGTC AAAGATAAGATGGGCCAGATCCCCATTGAAATGTGCTGAAGTACCTGTGGCAGTT CCGGGAGACAGTCAGTGCCGAGGATTTTGAAGCAGCCGCTAAGGCTAACCATCTGG AGGAAAAGATCAGCCGGGTGAAAGCCCACCCAATCGTGATTAGCAATAGGTACTGG GCTTTTGGGACTTCCGCACTGGTGGGAAACATTATGCCCCGAGACAAGAGGCATCA GGGAGAGTATGCCGGTCAGAATTTCAAATGTGGCTGGAGGCTGAAGTGCATACG ATGGCAAGAAAGCAAAGCACCATCTGCCTTTTTATAACGCCCGCTTCTTTGAGGAA GTGTACTGCTATCACCCCTCTGTGCGCGAGATCACTCCTTTCAAACCAAGCAGTT TGGCTGTGAAATCGGGAAGGACATTCCAGATTACGTGAGCGTCGCTCTGAAGGACA ATCCGTATAAGAAAGCAACCAACGAATCCTGCGTGCAATCTACAATCCCGTCGCC AACACAACCTGGCGTTGATAAGACCACAACTGCAGCTTCATGATCAAACGCGAGAA TGACGAATATAAGCTGGTCATCAACCGAAAAATTTCCGTGGATCGGCCTAAGAGAA TCGAAGTGGGCAGGACAATTATGGGGTACGACCGCAATCAGACAGCTAGCGATACT TATTGGATTGGCCGGCTGGTGCCACCTGGAACCGGGGCGCATACCGCATCGGAGA GTGGAGCGTCCAGTATATTAAGTCCGGGCCTGTCTGTCTAGTACTCAGGAGTTA ACAATTCCTACTACCGACAGCTGGTGTACAACGGCATGCCATCAAGCTCCGAGCGG TTCAAGGCCTGGAAGAAAGCCAGAATGGCTTTTATCCGAAAACCTATTCTGTCAGCT GAATGACGAGGGACTGGAATCTAAGGGTCAGGATTATATCCCCGAGAACCCTTCTA GTTTCGATGTGCGGGGCGAAACCCCTGTACGTCTTTAACAGTAATTATCTGAAGGCC CTGGTGAGCAAACACAGAAAGGCCAAGAAACCTGTTGAGGGGATCCTGGACGAGAT TGAAGCCTGGACATCTAAAGACAAGGATTTCATGCAGCCTGATGCGGCTGAGCAGCC TGAGCGATGCTTCCATGCAGGGAATCGCCAGCCTGAAGAGTCTGATTAACAGTACT TTCAACAAGAATGGCTGTAAACCATCGAGGACAAAGAAAAGTTTAATCCGTGCT GTATGCCAAGCTGGTTGAGGTGGAACAGCGGAGAAACAAACAGCGGTCTGAGAAAG TGGGAAGAATCGCAGGTAGTCTGGAGCAGCTGGCCCTGCTGAACGGGGTTGAGGTG GTCATCGGCGAAGCTGACCTGGGGGAGGTGCAAAAAGGAAAGAGTAAGAAACAGAA TTCACGGAACATGGATTGGTGCGCAAAGCAGGTGGCACAGCGGCTGGAGTACAAAC TGGCCTTCCATGGAATCGGTTACTTTGGAGTGAACCCCATGTATACCAGCCACCAG GACCCCTTCGAACATAGGCGCGTGGCTGATCAGTATCGTCATGCGAGCACGTTTGA GGAAGTCAACGTGGAGAACATTGCCGAATGGCACGTGCGAAATTTCTCAAACCTACC TGCGTGACAGACAGCGGCACTGGGCTGTACTATAAGCAGGCCACCATGGACTTCTCTG AAACATTACGGTCTGGAGGAACACGCTGAGGGCCTGGAAAAATAAGAAAATCAAGTT CTATGACTTTAGAAAGATCCTGGAGGATAAAAACCTGACAAGCGTGATCATTTCAA AGAGGGGCGGGCGCATCTACATGGCCACCAACCCAGTGACATCCGACTCTACCCCG ATTACATACGCCGGCAAGACTTATAATAGGTGTAACGCTGATGAGGTGGCAGCCGC TAATATCGTTATTTCTGTGCTGGCTCCCCGCAGTAAGAAAAACGAGGAACAGGACG ATATCCCTCTGATTACCAAGAAAGCCGAGAGTAAGTACCACCAGAAAGACCGGAAG AGATCAAAAACAAGCCAGCTGCCTCAGAAA	parent Cas12i4
956	MASISRPYGTKLRPDARKKEMLDKFNTLTGQRFVADLALCIYGLTLEMAKSLE PESDSELVCAIGWFRVLVDKTIWSDGDIKQENLVKQYEAYSKEASEVVKTYLNSPS SDKYVWIDCRQKFLRFQRELGTRNLSEDFECMLFEQYIRLTKEIEGYAAISNMF NGEKEDRSKKRMYATRMKDWLEANENITWEQYREALKNQLNAKNLEQVVANYKGN GGADPFFKYFSKEGMVSKKEHAQQLDKFKTVLKNKARDLNFNKEKLLQYLEAEI GIPVDANVYSQMFNSGVSEVQPKTRNMSFSNEKLDLLTELKDLNKGDFEYAREV LNGFFDSELHTTEDKFNITSRYLGDKSNRLSKLYKIWKKEGVDCEEIGIQQFCEAV KDKMGQIPIRNVLKYLWQFRETVSAEDFEAAAKANHLEEKISRVKAHPIVI SNRYW	Parent Cas12i4 amino acid sequence

	AFGTSALVGNIMPADKRHQGEYAGQNFKMWLEAELHYDGKKAKHHLFPYNARFFEE VYCYHPSVAEITPFTKQFGCEIGKDIPDYVSVALKDNPNYKKATKRILRAIYNPVA NTTGVDKTTNCSFMIKRENDEYKLVINRKISVDRPKRIEVGRTIMGYDRNQ TASDT YWIGRLVPPGTRGAYRIGESVQYIKSGPVLSSSTQGVNNSTTDQLVYNGMPSSSER FKAWKKARMAFIRKLIRQLNDEGLESGQDYIPENPSSFDVRGETLYVFNNSNYLKA LVSKHRKAKKPV EGI LDEIEAWTSKDKDSCSLMRLSSLSASMQGIASLKSLSINSY FNKNGCKTIEDKEKFNPNVLYAKLVEVEQRRTNKRSEKVGRIAGSLEQLALLNGVEV VIGEADLGEVEKGSKKQNSRNMDCAKQVAQRLEYKLAFHGIGYFGVNPMTYSHQ DPFEHRRVADHIVMRARFEEVNVENIAEWHVRNFSNYLRADSGTGLYYKQATMDFL KHYGLEEHAEGLENKKIKFYDFRKILEDKNLTSVIIPKRGGRIYMATNPVTS DSTP ITYAGKTYNRCNADEVAAANIVISVLAPRSKKNEEQDDIPLITKKAESKSPPKDRK RSKTSQLPQK	
957	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGS LTLE MAKSL EPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYS GKEA SEVVKT YLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMF SNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYLWQ FRET VSAEDF EAAAKANHLE EKISR VKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE AELHYDGKKA KHHLFPYNAR FFEEVYCYHP SVAEITPFTK KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESGK QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSAS MQGIASLKS L INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLNKKIK FYDFRKILED KNLT SVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	Variant Cas12i4 A
958	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGS LTLE MAKSL EPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYS GKEA SEVVKT YLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMF SNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYLWQ FRET VSAEDF EAAAKANHLE EKISR VKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLFPYNAR FFEEVYCYHP SVAEITPFTK KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESGK QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSAS MQGIASLKS L INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLNKKIK FYDFRKILED KNLT SVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	Variant Cas12i4 B
928	AACGA ACTCCATCTGGGATAGCAATAACCTGTGAAAATGCTCCCCGGCTAATTTG TATCAATGATTATGAACAACATGCTAAATCAGTACTTCCAAAGTCTATATATGACT	HAOI

ATTACAGGTCTGGGGCAAATGATGAAGAACTTTGGCTGATAATATTGCAGCATTT TCCAGGTAAGAAAATTTATTTTTTAAAAATCATGTTTTAAAAATTACACAAAGACCGT ACCAAAATAAGATCTCCTAGTTTTACGTTGGTGGTGTGTAATTATTTGTTTCAGATT TGTGCTTAGTAGAGAGGGAAAAGTCTTGGGGCTGTAAGAAATCTTGGGGCTTTAA ATTGTTAAAAAATATTCCAAGCCTGTGAATCTTGAGGAACCTGACTGCAAAAGCCAA ACCTATGTTACTTCACCTTGGAAATATGACAACAATTAATTTAACTACATGTAAAAA TAGCGATAAAATTCGGATGACTTTTTCTTTTCTTAGTATGACAGTAAATGCTTATGT TCATGGTGTAGGAAACAGCATTAAATGCCAGATAACCATCTTATCCGGATGAACCA GACTGGATTGTTGGCTCAAATGTTTTCTTCTGCTGGCTTTTCGTGTTATCATTCA TTTTGATTACTGTTGTCTAAACTTTCACCTTTAGATTTCAATTTGTCTATGCAGCAT TAATCTTTCAACTTTGCTGTTTCATCTCTCCTTCAAAGCACTTCATCTCTCTTCCC AAATTAGTTTTCTTTGACTTTTCATATTTCAAAGCACAAGATGGTGGGTGACATGG TTTTATGTTTTCTGTTTGTAATAAAAAACAAGAAATAAAATCATTTCAAAGGGTTTTT TTTTATAGCAGTTACAAAAATGGTTATTTGCTGGAGCAAGAGAGGAGTGCCTTCA CTACACTACACTCAGTCTCATCCATCTAACATTATGGCTGTAGTAAAGGCAATCG GTATTGTGGGTACTCATTGATGGTGATAAAGACAAAAAGGCAGAAAATATGCAGGG GGAGAGAATTAGCCTTCCCTCCCTGATTTCTTCTTTAGTCTACAACAAAATCACTCA AAATCAGTTTTCCATATTTAAATTAGGAGAAAATAAAATTATCCTGGCCAAGGTGGT CTCTGGTAGGCAGCACTGATTCACCCACAAATCCATGTAGAAGACTGAAAATGGCA ATGGGGTGAAGGATACGGCCTCTCCCCAACCCCTTCAAGCCTTGACTTTGTCTCAG GTTTTGCCTGGAACCCAAATGAGCTCAACAAATGCCAGGGAAGTCATGGGAAGGGA AGTTGACTGAGAGTAGAGGGGCTTAAAATTCTGCATCATTATTTACTATTTTGGAC TCATTTAAAAGTTTTCTGCTCTTGGGAAGATGCCCTTCTTGGGCCGATATTAACCTT GTCCACCAAAAATTTGCCTATGAGTGGTCTCTTGAAAACACTTTAACCCAAATAGGT TATTACAACCAAGGAAATTTTCAGACCCTTGACAGATTTATAGAGTTAGTGTCTCAG CATTGCTAGACCTCCAATGCTCAAGTGATTATTTATTTTCAATTTGTATACAGCTTTC CTTACTTCTTAATTCCTTTTGTGCGATGCTAGCTAATTAACCTAGAGCTAATTAGGA GTCTCCATGAGCTACACTGTGTACTACATGCTGAGGACAAAAGCAGTGAGCCAGACA AAGTTCCTGTCCCTAGGAACTTACATTCCTTGGATGCATATCAGCCTCCATAATG CTGTTGGGTTGAATTGATGCAAAATGGGCCCAAAATAGTTGGCCAAGTGGAGGCTCT CAGAGAGGATGCAAAGGGGCGCCCCAAAGCAGATGGATCACCTATGCAACCCCTTA AAATGTAGAAACTTTGGGAGACATAGAAGGCTTGGTGACTTCTAAGTTATGAAGTG GAAAAGTGCCCTCATGCCTTATGTGAATTACATGGTATTCAAGTGAGTATTCCTATC CTATGTGTGTACCGAGTAACCTTAGGGATAGGACACAGATAAATGAAAATGAATTTGC AGTGTCACCTTTTCCATGAACCTTGATCATTCTCTTTTGTTCAGCTTTAAATTTAA AAAAAAATCAATCAACTTTCTTTGGAGGACAGCTGATGCTATTTTATTATCAACT AGTTGAGTTTTTATTGCAATACATTTTGCAATGTGTCTCTTTTGTCTGTATGACTC GCTAGGTGAACCTTGATTCTCCTCACACTGCATCATGTAGCTGGTCACGTGAACTAA GAATAGAAATTTCTGCCAGGGTTGTGGAGACTTTGGGTTGATGGCATGAAGGAAATC AACCTGAAATTTACATTTCTGATTCTAATGAAAAGTGCAAAACAATCAAACCTCAG ATAACCCATTGTGATACAAAGCCAGAGTATTTCAAACACATTTATGAAATTTATAC ACCTCCCCATCTCGCAAGTACAACAAAAGGTCATTCACCGTGACAGCTTTTATTTC TCTGTACTCAGCTCTGATAATCACATTTTGGAGTTCTGGGGACATGGACCACTCAT GTGACCCAGCAGTTGCTTGGAGATATTTTGGGTAAGACTTCAGACTAATATTACT GTGGCAGTAGAAAAAATGTTTAAAGGACAAGTAAATGGAACCAACCCAGAACAAA ATTTCTTACGGTGGTTATAACAAAACAGGGTAAATGTCAACTTGCTACATTTTGCA TGGCTGGAATTGATTGGGATTAATCAACGAAGAACAGTAATTTGTTTCTCTTACA CATTTATTCAAAGTAGCCTTCTCAACTATGGTCTTCACGTTGTTGTAGCTTTTTTT TCTGAAATTATCAATGATGGAAGATGATTAACAATTTTCGACACTTAGAAGCCCTC ATGATTTCAAAAAGGAACTCTTTCTGCTGCGTTACCTATTGAGACTGAAGATG GCATCATTTTCTTTTAAATAACAGATGGGTAAAAGTGATGTCATTCTTTCACTTTA ATATTTGAGAAGTGATATGAAGTTACAGTGACATTGTGTTCTCATAGGCATAAAT GTCACAAAATAATTTATCTAGTATCCACAATAGGTGAATAAGGTGTTTTGCTTTA TATATTTTAACTGTTTAGAGTAAAAAATTAATGTGGAGAAAATTGGAATGCAGTAT TATAGGATTACACAACCTTACAAAACATGAATCCACTATGTCCAGTTAGTGTGATTC AGAAACAGCATGCAGTTATAAAGCTGGGTGAGGCATGGGTGTCTTCTTCAACAGG GCAGCTACTTTGTGAGGAGTGTATATATCATTGATTTTTTTATAAGTTAAATTTG AGGCCCTGTAGATGTGAGGGTGGGCCAAAATTCCTGTGAACAGATTCTCCCCGT TACCCCGCTTCCTTTACTCTGGCATCTCATTTTCTATCCTTTGAAAACGGTTATT	
---	--

ATTCAATTGGTTCAACTGTTTGGCAGTTGAACCAATTCTTTTCCAAAGTGAGGC CCAGGAAAGCACAGTCCGAGAATATAGTGAGGTGCTATTTTATGTATGATTGTGGG AAATTTACTTTAAATTTGGAGTGGGGTTGGGCAAGGCTTGGAAAGCTAGTGAGCTAT CTGACATAGTTGTTACTACTATTTGAAAAATATCAAAACATGGAGGACTCTTTAGA TAACATGCCTGTTCCCATTCATTGATTTTATCTAATTTTACGTAGCAATTACGTT TTGTGCATTGGTTGACAAGCCTCTGTATTATCCTCAGAACAGAAAATACTGTTTAA GGGAAATTAAGAGCCCGCAGTTACTAAAGTGACTGCGCCACCAAGTGGACAAGTGT AAAGCCACTGTCTGGAGATGGAAGGATTGAGCTTTGCTTTTATAAATGGGAATTTGA CCTTTAAAAATGTCCCTTTTGGCACGCACGCGCGCGCGCGCGGAACACACACA CACACACACACACACACACACACACACACACACACAGGCTGCTGCCCTGCAGATTG CTTGTCTTGTCTATAAAGCTTTTCATTGTTTCTCTAGCTCTAAGTAAATATTAATGC CTTCCAAGGCTGGCATGCCAATGGCTGCTATTAAGATCGTTTTCTCTCATTCTAAT AACACACTTAGAGATGATTGGTAATAAAAACTCTCTTCAAGGCTTCTGCTTCTCCC CCTTCAAAATGGAGATCAAAGAATCATGCTGTGAGGGTCCGTCAAGAAGAAAAGAC TTTTCAGCAACAGAGCATGTGGTGTGGCATAAAATAATGACAATTATAATGTTCAAA GGAATAGCATAGAAATCACACAGTAAAACTTCTTTATTATGCTTTTCAGGGACTGG ATGTTTTTACTTTTATTATGTGAGGAAGGGTTAGATTACAGACCCTTAGCTATTCCA CAAAGCAATAGAAGGCAGAATTTCTTCTTCCGCTACAGGAAGCAGCTTCGATTAA GGGCTTTTTCTTTTCTTCTTTTCTTTTCTTTAAGTTACTGCATTACTATATCATA CTTCACTATATTTACTAAAAAGTCATGCTGTTTCTGGAAGTAGAGTTACATCTAGG AAATACTAGGTGAATGCTGGTTAGATATGCATGTGTGCCTAAACAACACGTTTATT ATACTCATGCATACATAGAAATAGGGCTGTATTTCTTCAATTTTAATCAGTACTAA TGAGAATAATAAATCAAAACAAATAGGAGAGATATATTTTGGCAGGAGGAAAAGAGA ACTAGTCTTCTGTAAATTTTACTGGTGAATTTTGGTTGCTGGTTTATTGGTAAT TTTCAATTCACACAGAAAGATCACAGAAACATTCATTTAAATAAATTTTCCGGAG TCAAAACTTTTTAACACCCAAATTTTCAGTTTTTGTCAAATAACATTTTTGAGAAA AGTGTAAATTAACCTAATAAAAAACCTTCCCTCATCATTAGACTTTAATGAATAT GGCATATAACTAAATAATTTTGAAGAAACCAAATTATAATTTTAAAGTAATTGCC TGAAGCTGCTGTTTATCACATAAAAAAGAAGACAACTAGACATAGCATATCTTCTT AAACCTAATCTAACTCTATGCAATTTGTATACCATCTTGATTTTCAAGATTGGGG AAGTGAACGAAACTATGTTTACACAAAGAACCTGTACGTGAATGTTTGTAGTGGC TTTATTTAGAATTTCCCCCAAACCTGTAAGTATTCAAAATGTCTTTTAGCTTGGGA ATGACTGGACAAATGATAGTACCCCTGTATGATGGAATATTATTCATCAACCAAAA GGAACAACTATTGACACGTACAACAACATGAGAAAATCTCTAATGCGTTATGTTA AGTGAAGAAGCCAACTCAAAAGGCTACATACTGAATGATTTTGTTTACATGATA TTCTTGGAAAAGCAAAATTATCAGGACAAAGAAAAATGCATCAGTGGTTGTCAGGG GATTGAACCTGGGGAGAGTTTCTCTGCAAAAGAAAAATGGGGACTTTTTTGGGAATGA TTGAACTTTTCTAGATCTTGATTGTCATGGCAGTTACACCACTGTATGCATTTGT CAAAATTCACAAAACCTGCAGACTAAAATGAGTGAATACTATTATGTATTAGTTATA CTTTAATAAATAATTGCTTGGGAAATTCATTATCCTCTAATTGTTAACTTTCTAAC CAAACAAACAGTAAATTGCCCTCTTTTCCATTAGCTTTATGAAGTCATTTGCTTGT TTGGAAAAAATCCAATTATATTTTTTCTTTTAACTAAAATGTAATGTCAAAGTTTT GGTATGATTCTGAACTCTAAAGCCTTTTATTTTATTTTATTTTAAATTCTAGA TGGAAGCTGTATCCAAGGATGCTCCGGAATGTTGCTGAAACAGATCTGTCGACTTC TGTTTTAGGACAGAGGGTCAGCATGCCAATATGTGTGGGGGTACGGCCATGCAGC GCATGGCTCATGTGGACGGCGAGCTTGCCACTGTGAGAGGTAGGAGGAAGATTGTC ACCACAGGGACAGAAGGAGGCTAACGTTTATCGACCTCCTTCTCTGAATGCACCAA GCAATATGTTTCTTGATGTTTTTACACTCAGAAACATTAAGCTCATGGACTCTAT CATCAAAATACTTGTCTTGCATGTCCTGCTCCTCTCTTTCCAGCTGTGTGACTG GGCAAGATATCCTCTCTCTGCATTGGTTTCTTGGCTGTAAAAATAGGGACAAAAAT TGTACCTGCCTCATTTGGGTTATGGTGAGAATTGAATGAGTTCAGGTATACAAAGTT CATGGCAGAGAGTAGGGGCTCAGTAACTGTTGGTTATATTTAGGGTATTAATAGTA CTGCTCAGGAAATGATCTCTGACAGGTAGACTTGCCCAAAGTCACAGCTAGGTA GTTACAGAATTGGAATTCAGCCCTGTGGCTACCTTATCTCAAAACCCTCCTGCTTC CCCCCAACCAAGTGGTTCTCACAGCCAAATTGCAAATGGAGCAACGTGGTTGGTT GTGTTTTCTTCCGTGGTTTTTGGGTGATGATTCTTTTTTATGGATGAGTTATATTCC CAATAGAGCAGTTCCAGCTGTCTTAGGAGGGAGTGATGAGAAAATCAAAATATGATG TAAAGAAATCTCTTATTAGGGCTAATTTATTAACCTTCCAGTTCTCTAGCAACTGT GAACATTTGAAAGGCTGTGCAGAGTAAAAAATCTCCCCAAATGTGCTCCAGAAAC	
---	--

TAATATAAAAAGTTGGAATGAATTATTTTGATGCTAAGCAGAGCAGAAAAAGACA CGACTATATAATATTTTAAACATTTTAGTTTTAAGAATTAAGGATCTTGTGAATT CACTTCCCTTCTTGAAATGTCTGACATAAAATCTGTGAGGATATCAGAAATGGCA CAATGAGGTTTTGCTGGACAGACTTAGCAGCTTCCTTAATTCTAGGACCACATACA AATAAGTGGCTTTGGGGCCTCAGCCTTTTGTCTATGGTAATCCTGAAACATAAGTA GAGAGAAGAAAAAAGGGAATACTAAATGGGTAAATATCTATACAAAATCAAG ATAATAAAGGCCCTTTCAGGCTTGAAACTATAGGCAACAACCTTAGAACAAAAGAA AACAAATGAACATCAAAAACTAAACTTTAGTGCTCTTAAATCTCAATGAAAAATA AAAAGTAAATGGTAAACTGAAAGAAATGGAAAAAATATGAGACTGTGAAGGGTT AATGTCTTCCACGTAAAAAGCCTTATATTTGAAGAAGAAAATAATATATTGCT CAAAGGAAAAAGAGAAATAAGTGAACAAAAGATATAAATAGGAAATTTACAAATG GAGACATAAAAGTGACCAATAAACATATGAAAAATATTCAATTTCAATTAATAGCA AAGACATGGAAATTATGACCATCTATTTTATTTTCCGTATATCGAATTTTATTTT AAGATCAGGCAGTATGATTAGGTTAGGGAGAAAAATGTGCATTTCAAACAGTGTGA GAAAAGTATAAAGTGAATAATCTTCCTAGAAAAATAATCTGGCACTGTATATCAAA GCTCTAAAAATGTAAATTCATGTGATGTTAAAAATCTCTCTAGGAATTCGAAG GAAATAATTATGATTTTGGAGAAAAAATCATTCTGCAAGGATTTTCATGCTTC TTATTTTAGCAGGAAAAATAATTTGAAAAAATACCCAAACATCTTATAATTGGAG ATAGTTTGCAAAAAATATGATGCATAAAAAATGACATCAAATTTAAAAATTATACTA TAGGAAGAGTGCAATAATGTAGAATGATATTTTAATTTAAAAATGTGAGAAATCAG TTGCAAAACAATAGTCAGGTCCTAAAATACATTTAGTTTCAAAGATCACAATTTACA AATGTTTATTTATAAGTGATGAGATTACTCCTGACTTTTATACTCTTCTGATTTTGG GCTCAACCTTATAAACTCTTCTTTGAATTATTTTGTAAAGGAGGAAATGATAACAAT TAGATTTAAAAGAGTAGAGATAAAGGGACAAGGGACCATGAAGAGAATGGAAATAA AGAAAGGAAGCAGAGAAAAGCAAAAAGCAGAGCTCACTTGGTAAGGCACCTGGAGC CAGCAAATATTTTTTACCACATGTATTAGTTCCTTCTCACACTGATTTAAAGATAC TCTTCGAGACTGGGTAAATTTATTAAGGAAAGAGGTTTAACTGACTCACAGTCTTAC ATGGCTGGGGAGGCCCTCAGGAACTTACAATCATGGTGGAAGGCAAAGGGGAAGGA ACGACCTTCTTCCCATGGTGGCAGGAGAGAGAAGTGCAAGCAGGGAAATGCCAGAT ACTTATAAAACCATCTGATCTCATGAGAACTACCCACTATCATGAGAACAGCATG GGGGAACACACCCCATGATCCAATCACCTCCCACTAGGTCTTTCCCTCAACACCT GGGATTATAATTCAAGATGAGATTTGGATGGAGAAACAAAGCCTAACCATACCAAC ACATATTGCTTTATTTGATATTTGACAGGTGTTTCTGTCCCTGTTTTGTGGGCAAG TAGCTAAAGTTCAGAGAAAACAGTTTTTTCATAGCTCGTCAATGACAGACTTATTC TCCAAGTCACATTTGATGGTTCCAAGACCAGTCTTTATTCTTGGTGGAGTTGGGCT GAGAAAGAAAGAGGAGAAGAAAGAAAGAAAGCTTCCTTAGAAACTATGATTTG ACAGTGTAAGTAGGACTATTTCCCTCCAGAAGTAACCATAAGAAGATATTTAAATGCC TATTACAGTCTTATCCCTTAGATTTATTTAACACTTATAAAGCAATTATCATGTT CCAGACACTATTTTAAAGTATATTACGAGTATTATAGCATTGAAGGCTCAGAGCGGC CCAATAAAATCGATCATATTATTTAAACCTATTTTACACAGGAGAAACTGAGGTACA CGCCAGGTGAATAACCTTGCCTAGGGATGCACAATTCATAAGTGATAGAGATGGGA TTCAGACAGAGGTATTCTGTCTCCAGAATCTGGGCTCCTCACCACTTTGCAAGAGC TTTAATTTAGAAACTCCTATGAAGTGTCTGAGGAGAAGCCCATTATGATCCCTT AGAAGTAATTATAGTTTTAGGAGCATGCAAAGCAGACCCCTCAGGAAGATAAGTTA CACAATAGACATTTGGATAAGGTGGATCCAGCAGAACAAGAGAGGGTGGTGACAT CGAGATTGCAGAGGAATTGGAGAAGGCAATGGAAGTGTACACATGTTGCCCTCAAA AACATAGGGTCTTCCATTGGGTTCCTATCAGGGCAGCAACATCAGAGTTTCTATTCT TGTATTTATACTAGAAACCTCTCTCCAGGGTTCTAAGTTTTACCTATGTTTTAA AGACTATCTATAGGTTATTAGTCTATTTAATATTTAGGTGTATCCAGAAAGCTGAT GGTCAACAGCTCATAGCAGGTGTTCTTTGGCTGGTGTGTTTATGTTGTGGGACAGT GGGTTACTTGCAAGGAAAGGATGAATGGCTGGAGTAGATGGTGCTTGTGCTCTGCA TGTATTCCTTCTTACTTCCCATTTCCATCAGACCTACCACTTTTTGCTGACATT ATCTGTTGCAACATGAGCCCATGGATAGGTGTGTTTGAAGTAGGGGAATGGGAGAG AGGGTTCCCTAGCTAATGATGTACAGCAGTAGGTGGATAAATACCTCAGCTCTCTT TGCTCAGGTAAGTGAAGCATTCTTAATATGGTCACCCAGTGTTCCTTGAAGGAT TGAGTCCAGTTGGCCCTGAGGTGCTGCTGATGAACACACCCCTCTTTTATTGG CTTCTTCCCATTTCTTCTCATTCCCATTCCTTCAATTCATTGAGATTGTTTC CAAATAAGATGACTTGCTCTCACATCTCTGTGTCATTTTGGCTTCTTGAAGTATG CAAACCAGGATAATAGCTAACTGAAGGCTATAGATAGCCACAGGCAAATTTAAGTA	
--	--

ACAGTGTAAAGAATATTCATACTTGGCAGAGATTATTTATAAAAACTCAGAAAAATT CACATGGAATTATGAAGTTATTATTGTATTTATCCATCATTCCCAGAAAAGAATAT GGAAATCCTCTCAAGCAAGCCAGTCCTTGGGAATATTGGGAAATCTATGCAATTTG TTGTGGAGTATTTTTTTTTTTTGTACCCTCCTAAATATCTGGCCGCTAAGCATTCC TGTCTCCAGGGACTTAGACCCTAGCAAGGAAGAGAAGTTGGGGCCAGGTTTCAGAAA ACGGGTTAGTTATCAATCTCCCTGGAGAAGTGCCCCCTCAGCAGGGTCAGTGAGA GTAAGTGAAACCCATTGGTGCCACAGGCAATGGTCTGGCCTGAGTAATTAGAATG GGCCTCCAGAAAGTTCTGGGAATTGCTATGGTGCCATAGTCTCATTTCCTCCCGTTG ACTCTCCAGATTTATTCAGAGTCCAACCTTCAAGGGCCTTTCTGCCCTTCCCTCTCAC AACTGTGGAATAATAATAATCCACCTTATTAAGTGGGACCGAGAACTGAGCTCGAC TCTTATTTTTTTTGAGACAGAGTCTTGCTCTGTCCACAGACTGGAGTGCAGTGGCAC TATCTCAGCTCACTGCAACCTCTGCCTCCCAGGTTCAAGCGATTCCCCTGCCTCAG CCTCCTGGGTAGCTAGGACTATAGGCACGCACCGCGACGGCTGGCTAATTTTTTGT ATTTTAGTATAGACAGGGTTTCACCATGTTGGCCAGGATGGTCTTGATCTCCTGAC CTCATGATCTGCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACATGCGTGAGCCAC CGCGCCTGTCTGAACTCTACTTTTTTACACTGCTGCATGTTGTAGAGTGACCAA TGAAGCTATACTTTTTTCATTTTCAAAATGATGATGAATACAAGGTTATCAAATAA AACACAGAGGGCCCATATGTTTGAATTTAGATAAACAACAAATCATAGGTGTCC TGTATGTTTGCTCAATCTGGCAACCTGGATGAATAAGAGCTCTCACCTGAGGATT TCTTGTGAGGATTCATGAAATAAATGCTAGAAATGCTTACACACTATCTTTATTTG CCCCCTCAGAGCCCAAAGTCTCTGAAATCTTTATCTTTCACACACAAAACTCACTT TCAGAAAAGTATATTCCATTTACATCTAGTGGAAATAAAAAATGTTCTTTTTCTTT GTGAAAAATATTTTTATTTTAAAGCTTTATGCAGAAACCTCAGGGAAAAAAGGTAC TTTTAGGAGCCAGGCTTGTAATGTAAATGTCCAAAAAAGATGAAATTGAAACAAAC AAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAGTGCAAGCTCCTGTGTGGAG ACTGCAGTGAGTCTGAGATTGCATGTTCCATCAGAAGGGGGCAGCCACATCTTAGC TCTTGATGACCCAAGGGAGCAGGGATGTGGGGTGGCCAAATCTTCCAAAAATTTTAA GAAGCCAGAAATCTTGATTTCTATGTACAATCTCCTGGTTTTTAAATGTGGGCCAAA TAAATCAAAATTCCTTAAACACGTGTTGGGGCAACAATGTGTGGGCCAAAGTAAA TACTTTTGTGGGTACAAGTGTCCTCCTAGGCTGTACATCTGGGACATCTGATTTAT GTGGAATTTTACCGAGAAGTAGTTTATTTCTGTGGCAGGTCATTTTCACTTTCTA GGATTATGTTTCTTCATTGATAAAGTGAGCTACTTGAGCAAGACCAGTGGATTGAA TGCCACGTCCCAAGGAGGCTGGGGTTGTTTCCAGGGATCTTACAGAACTTAGGTGT GATACTGAGCATGAGCTACTTGTGTTGCATTTGGTGTTCAAAAGAAAAGTCTTT AAATAGTTCTGTGCTGGAAAGACAAAAAAGAAAAAAGAAAAAATTTTACAAACAAA ATCTCCAAAAACAAAACCCAGAAAACCTGGCATAGAAGTGGATGATCTTTGCAATT TTTTTTCAGTATATAAATAAATGATTTTGATCCCATTAAAAATTTTATCAAATGCAA AAAGAAACAATTCAAAGTATAGAGCTACCTTTTCTTACTCTACTGAAATCTACACT TTATGTGAGCCCTGGAGGGTTTAGACGCACCTTATGTGAGCCCACTTCTTTCGACT GCACTATGTGAGCCTTGGAGGGTTTAGATGAGGCAGTGAGCATTTGAATGCTTTTA ATTTCCATTTTTCAAAGTACATTTCTTGGTCTATAGGAGAGGAACAAGATATGTAAC TATCTCTGACTATTGCTAAAAACACAAACGCTCTTAATAAATGTTGCATAAACTCA GAAAGTGATACTTCAAAGTCTTGTGAAAAATGATGATCACCAGCATTTATACAGCA ATTAGTATGTGCCACGCAATTTGACTTTATTATTTATTCATCTATCTTTACCACCA TCTTAAATATGTGAGTGCAAAACCTGAGAACTTTCTCCAACCTCCTGTGGGTGT GGAAATCGAGGCTTAGAGAGGTTAATGCTTTGCTCAGATTATTAATCACTTAGGCA GTGCTACCTATAATATCCTGCTCTGTTACTGGTATTTCCAAACGTCATTAAGTGA GCAAGAATCCTAAGGCAAGCACTATGCTATCATCTTAAAAATATTTATTGCAAAACAT CCTATGTTTTATTGTTTTATCTTTTTAACTTTGAGAAGATAAAATAAGCCACAGAA GTGAAATTAATTGGGAAATCATTGCTTTTTTGCAAAATTTGGGAGCATAAACAATG GGTCATGAATTACAATCAAACAAAAGATAAAATCTAAGAAGTCTTTTAAAGTGGA AAAAATAACTGAAAAATACTGAATGGAGGGCAGTTTTTTCATGCACTGTGTTACGA ATAAAAAATTTGATTCAATGGATTACTTAATCAACATTTTAAATAGTTGTAAATCTT ATAATATTTAAGCTGTTTTATAAGTGCTCTACTTATAATGGCACATCCGTTTGAA ACTCTAGCAGATCATTTTTATTTATTTTTTTGAATTTTTTCTTTATATTTCTTTAA AGAAGGATACAAAATTTATTTCTATGAATATTTAACATATGGAAGGAAATAGCAATA ATAAACATAAATGCTAACACATATAAAATAGGTGGTATCATTAGGCTAAATTTTAG TCTTCCAGGATAAGTAGAACATCTCTGACTTCTCAAATATCCAATTAATAAAATGC TTACTATACCATTTGGTGCTTTAAGAACATTGCCATGGAACCTCTCAGGTTTTAT	
--	--

GCACAGTAGCTATAATAAAATTTTCCTTCATCTTTCATGGAGCTACTTGAGATTTT TTTTCTCCCTTTTAAACATGAGAAATCAAAAAGAAAGAGAAAAGAAGGATTAAATAT TCATTATATCCTTTTGTCTGACTTGTATGTGGGCAAGTGCCACATGAGGGAGTG CTGGGACCTCATATCAAGAAAAATTAAACCTACCTAATGCGTTCCAGGAATGTTT AGCATATTAGCAAATTCCTATTAAACTGTCAAAAAAAAAAAGTTTTAAAGAAA TTCCAGCCCCCTGGATGCAATTAGAGGCTACCACACTGGATTTGATGGGCCATAAAA CCATTAAATCTAAACACTTTCTTTTGGAGCCTAAAAGGCCAGAACATTCCAAAGTG AAGTTTTGGGACTCAGCTATGACTTGACCACCTATTAAGATGCAGGTGGAACAGAT TGCAGAGTAACACAAAGAGCCACACAGACCCAGATGACTGCATTAGGGTGTAGGT GAGAGTTTTAGCTGTTGAATTTTCTGGATTTTCCAAGATTAAAGTGATCAACCTTAA CAATGAGTGAAAGACCATTCAACAGGAAGAATTGTCATTTCCCTTGTCTTAAACCC AAACGATGTATTTTTTGAAAGCTTTATTGATTATATATTTATGTGTTGTGCTAGG CGACCGACTAGATATATGTTTCAGCATACCTACTAGGAAAAATATCCCCATTATTCT CAATTTTACCTAATCCAGGCAAAGCACTGGACTTGCTTTAAGGAACATTTTTACTC TTTCTGAAGTGGAGTGCCTGTCTATGATCAGGTGCAATGCTTGGACTTTACGTTCT TGTGATTAAATCCTTACAATAGGCCTGTGAAGTAATTCTCATTTCTGTTTGACAGTAG AGAAGAAGGAAGCCCTTGACCAAGGTCTAGTGCCAGTAATGGTGGTGATGGGGTTT GAACCTAAGTCTGTTTCACTCTAAAGTGTAACCAATTTTATGTTTGTAGACTTGCT TTTCTAACAATAAAAAAGTCAGTGACATGCTCTTCTGTGTGTAAGCACTCACACAC ACACACACAAACACACATCCGTATTACATATGCTTATATATGTATTAAGATTAT GGACATTTGATATATATACATATACTAAAATGTATAATTCATTGCTAAAGTATTTT CATATAAATAGTGGCTTCAGTGTTAAAATCAGTTTGCAATGAAACAAGATTGTTGA TTAAACACCTATTAATAAATTAGAATCTAGCCATATTAAGACAGTCATCGAATG GAGTGATTTCTACGATTTTGCACCAAAATTTAAGCTATTGGGTGGCTTTCTTGAGA GCATGAGATTGCTTCTCTCAGAATTATTAATGTGCTGATGACATTAAATGTGA CAGTGAAAAAAGTCAGAGGCTCACATGTGTATCCCAACACTGAAGTTGTTAAACAC TGGGAGGTTGGTTGAAGTTGTTGTGTGCAAACTCAATACTCCTTAAACCATTTATT TAAAGGCCTATCACTGTGTTATGGTCTCCATATGATCTGCCATTTATGCCAGGACT TGACAATTCAGTAAAAATGACAGAATAAACAACAGGAATCACTGCAGTAGAGCTAA TGTTTTAGTCTGTTGCAGAGTTCTGCCCTAGAAATACAGTGAAAACAAGGAAGGGA GAGCTAAGATGTCCCTGAGACTAATTGTTCCCTGAAAAATATTTTCATAAGTAAAAA AGAGGTCTAGAGGTGTAGTGGCAGTGTGATCACTCAAGATTATATAGCTCCGGATT CGTTCAATGGGCCATGATGAAAGCACGGCAACGATTAAATCTGTTTTCTTGGTCTT TCTTGGCAGTGTTTAAATTGGTTCAGTTCCATAAATTGTAAATTAAGATCTGTTTG ACAACTTTTAAGTATTTCAAGCATAATTGTAGTTGAAGGTTTGTCTTTTAGATCA CTGAGTCTCAGAACTTTATTTTTCTGGTTAATCTCAATTGTAATTTTAGACATTCAT AAAACATGTTGACTGCGTCTATGTGATGGTAGATCCTCTGTGAAGACCTTTATGA TGGTAGTTCCCTGTGAAGATAGGATGACACACTCAATGGACATTATGGTGCACAG TTATACAAACACTTCACTATGACAGGCCCTGAGTTTAGAACACACAACCTGCTTGG TACTTGGTCATCGCATATTTTCCCCATTACGTAATGACTTCCTGTGCAGATGACAA AATGCGTTTTCTCAACAAAATTTTTCAGTGCAGCTGTTTTGATGACTAAGTTTTT GTAGGAGCTTTTTAATCAAATGCACCTAAGAAAACCCCAACACTTTAGGCCCTTGG AACATATTACACTTTTTTGTCTCTCTTCTCTTTTCTTAAACCATTAATTTG GAAATTTGATTCTGCCTTCCCATAAAAGAGAATTATTTCAAAGAAATTTATTTGGG TCTAAATTAACATGTTACTTAATGTTCTGCTTGAATCTAGGTATATGATTAGTCC CATATGAATTGATGTTCCAAATAATTTACTCTCATTGATAACTAATATTTCTATT TCCCTCTATTGTTTTGTGGTGGTGGTGGTGGTGGTGGATGAACATCATTCTCAAAT ATATTATAATTCCCTTCCCTCATCAAGCCAGCATGATAAACTTCAGTTTTGCTTGA TGGTTTCATCATCTTATTTCTGTGTGTAAGATTGTTGGATTTGACATTAAACATTTG GAAACTATTTTATAATTGATAACTTGTGCTTTCTCAGCTTTGAGTAAGCGCTCTCT TCTTCATCTTATACCATTTTTATTTTATTTATTTCACTTCTGCTTCTGATCTGA GATCTAGGAAGCTGGACAAATCCAGATAAGCAAGCTAAACAAACAAACAACAACA ACAACAACAACAACAACAACAACAACCCAACTAAACCAACCAAAATCATGGG ATAATGGTTAAGTGTACTGAGGGGCCATTATGCGAACACAGTTTAATTCCTTGGCT TTAAACTAATAAGAGAAGAATACATAAACAATGTGGCAATGTACCTGTGACCC TCTCCAGAGGGTGCCAGGCTAGAAGAAAGGCAGATTTATCAGCAAGGCATGGCGGG CCATTGGCAAACCATGGGACAACCACTACCAACTTCACTGCCATTGCTCCATATTT CCCTCCCGTTTTCAATGAGCCCCAACTTTGCTCAGGACATCACACATATTTACTA ATTTGGATGAGTCCTTTTGAAAGAAAATATCTACCTCATGGTTCTCAAAGTATGGT	
---	--

CCTTGGAACATCAGCGTCAGCAGGACTCTGGAGCTTGTTAGAAATGCAGATCTTAG GTCGCACTACAGACCTACTGAGTCAGAATCTGAATTTTGTAAACATACCCATGTGA TTCCTCAAAGATTGAGAAGCCCTGATCAGAGCCTGGGATGAAAGTTCTGTGGTT CCAAGCCAAAGGCATAGTTCAGGCTTTCACACATGACACTATTAGATGTAGATGGA TATTGTTCCCTTCTGAAGACCCCTCAAGGTCTTCTGAGAGCCTATTAAGTTCAGAAT GACTGCCTGAAATGAGTGAGAAGTCACAAGGAGACTCTAGATAATTAAGAGATGTG TTCACAGTAGTCTTTGATAAAAAACCTGGGACAGGCAGGCTTAGTATGCAGGCCCT AAAAATTATGTACACAATGGATTTCCTATTTTGTCTTCTCACATCCAGATTACCT GGATCAGAAATAAATGTTTTTCATTAAGACTTGATGTGACAAACAAACAAACAAAA CTCTGCCAAGCTCTAGAAGAACAATTGCATTTCCAGCCAGAGGGAGAACACTGCC AGTTTTTGCTGTTTTCCAAAGCTGTTTACCTGTCTAGCTCATTTAAATCACTGTA CTTTGGAGTTCGGATTAGCGTCCCCAGAGGTAGCTGCATTCATACTTGATGAGTT CTTTTAAATCTCAGCCATTGATTGTAGGTTCCATAGTATAGGAAATTTAGCCAACC CTCTATTGAATGGCAGTTTAGAAAGGTCGAGCTACACTTACCTTATGTCAGGTTAT TGCAGACCCCTTGTGGCATTTTTCCACCCTAGGACATGTGATTTAACTCTAATAGAA ATCTTTATTATGGGTGGGTCTGAGATTAACTTTTATTCTATAAACAGAAATCATG CCACTGGCCGTAGCCCATTTTTTGGAGATGGAGTGGGGGGAATGGATGATAGTAAAC AAGGATATTAATCTCATTATTTTTATATCATTATATTTATAGTTACATTGCAAT GGAAGAGTAGAGAAACCAAAAACCTTACACTGGGAACCTTACAATTTTCTTCCAAG TATTACTGATTGATGTTTGGACTATGCAAGTGCTGCCAGCCCCCTAGACTCACTCT GCAGCTCCCCCATGAAATTTGTGAACAGGTTAGGGTGGGGATAGGGAAAAGCAT GTTCTGTGTTTCACTTCTTGGATTATTTGTTCCAGGCTCTCCAAAGTAATGTGTACC TTGGGAATGCAGAAATATCTCCTTAGATATTTCTCTCCCTATATATGTCTCACAG GGAATTTCTGGAATTGGAGAAGATTCCACTCTCCTTTAGGAGCTTTCTCCATAAAG GTATTGAGCATTGGACACTATATTTGCAAGGGAAAAGAGGAATGGGTCTCTTGAGC ATCAAAATCATTGTAGAAGAATCTCCAAACTGTTTTTCAAAATGTCTGTACTAACT TACATTCTGACATCAATGGGTTCCCTTTTCTCCACAAGGGTCCCTTTTCTTTGC ATCTTCACCAACACTTGTATCATTTGGTGTTTTTGATAATAACCAATTCTAACAGTT GGAGGTGATACTTCATTATGATTTTAATTTAAATTTCCCTGATAATTAGTGATACT GAGCTTCTTTTATATCTATTGGCCATTTATATCTCTTTTGGAGAAATGTCTG TTCAGATGCTTTTGCCAATTTTTTTCTTTTTTCAACTTTTATTTTAGAATCAGGGA GCCATGTGCAGGTTTGTACAAAGGTATATTGCATGATGCTGAGGTTTGGAGTGCA AATGAATCCATCACCTAGGTAGTGACCACAATTCCAAACAGGTAGTTTTTTTCAGC CCTTTTCCCCCTCCCAACCCCACTGTTGTATTCCCAGCATCTATTGTTACCATTT TTTTGACCATGTGTATCCAATATTCAGCTTCCATTTATAAGTGACAACATGTGGTA TTTGGTTTTTGGTTACTACATTAATTCACCTTAGGTTATTGATTTCAGCTGATCC ATGTTGGTGCAAAGGACATTATTTGTTATTTTTTATGGCTACATAGTATTCCATG GTGGATATGTACCACATTTTAAAAATTCAATCCACCATTGGTGGGCACCTGGATTG ATTCCATGTCTTTGCTATTGTGAATAGTGCTGTGATGAACATGCAGGTGCACGTGT CTCTTTGGTAGAATGACTTATTGTCTTTGGGAATATACCCAGTTAGTGGGATTGC TGGATCGAATGGTAGAAAACTCTCAGGCTTTTGAGAAATCTCCAAACTGCTCTCT ATAGTGGCTTATTTAATTTACATTCCCTACAGCAGTGATCAGCCTTCTCTTTCT CCACAGACTCACCAACATAGTATTTTTTGACTTTTTTAACAAAAGTAATTCGACTG GTATGAGATGGTATATCATTGTGGTTTTGATTGCAATTTCTTTGATGATTAGAGAT GATGAGCATCATTTTCATATATTATCAGCCTCTTTTATGCCTTTGTTTGAGAAAT ATCTGCAAAATGTCTTTGCCACTTTTTAATGGGGTTATCTGTTTTGTCTATGTTGA TTTTTTAAGTTTTCTAAAGATTCTGGATATTAGACCTTTGTTGGATGCATAGTAT GCAAAATATTTTCTCCAATTTTGTAGGTTGCCTGTTTACTCCTTTGATTGTTTCTCT TGCTGTCTTTGCTATTTTTTAATTTGGGTTATTTGTTTTCTGGCTATTGAGTTGT TTGAGTTCCTTATTTTTTTTTTTTGGATATTAGCACTCATTAGATATACACTTTACA AATATTTTCTCCCAATACCTGTGTGTCTCTTGATTCTGTTAATTGTTTTCTTTGC TGTGCAGAAACATTTTAGTTTCACACAATTCCTTAAAAAACTAAAAATAGAATTGC CATATGATCCAGAAATCTACTTCTGGATATTTATTGAGAGGAATTGAAATCAGCA TGTTGAAGAGATATCTGCACCTCTATGTTTCGTTATAGCATTATTCATAATAGTCAT GATATGCCATCAACCTAAGTATCCATTGACAGATGAATGGATAAAGAATGAGGTGT ATGTACACAAAGGAATACTATTACGCTTTTAAAAAGTGGGAAATCTGTAAACA TGGATAGACAGATACTATATGATCTTACTTATATGTGGAATCTAAAAAGGTAGGTC TCACAGAAACAGATCATAAAAGGTGGCTACCAGAGGCTGGGAGAGGAAGGAAAAG AATGAGGAAAGTGACATATTGATCAAAGTTGTACAAAGTTTCAGTGCGACTGGAGT	
--	--

AATAGGTTTTAGTGATCTATTGTACTGTCATGGTGTCCACAGTTAATAGTAATGTAT TGTATATCTTAAAATTACTAAACGATTAGGTATTTAATGTTCTCCCTACAAAAAAA TGGTAAGTTGGTGTATTAGTCCACTTTACACTGCTATAAGGAACTGCCCCGAGACT AAGTAATTTATAAAGAAAAGAGGTTTACTGGCTCACAGTTCTGTATGGCTGGGGAG GCCTCAGGAAGCTTACAATCATGGTGAAAGGGGAAAGCAGGTATGTCTTACATGGTG GCAGGTAAGAGATCCTGTGTGTGAAGTGAAGGGGGAAGAGTCCCTTATAAAACCAT CAGATCTCGTGTGAGCTCACTCACTAGCATGAAAACAGCATGGAGGAAATCACCCC TATGATCCAATCACCTCTTTCCCTCAACACATGGGGATTACAGTTCCCTGCCTTGA TGCGTGGGATTAAAAATTTGAGATAAGATTTGGGTGGGGACAAAAGCCAAACCTTA TCAGTTGGTGAGGTGATGAATATGGTCATTAGCTTGTCTAAATTTGTCTGCAATGT ATACATAGATCAAAACATCACTTTGTACCCCATAAACATGTGCAATTACTATTTTC CAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA GAGAAGATTAACAAATAATGACATTAAGAAATTCAGGTCTTGGCAAAATTAGAAA TACATACAAAGCTATCCAGAACTTATTTTTCCAAATGCATTAGGCGTCTCTCACC TTACCCCTTACAATTGCATGGCTTCAGAGATTACACAGAAAACGTTACAGAAACATT GCCCCAGTAGATGATCTTGCAATGCTATGAAGTAGGCAGAACAGCTGTGGCTATAG CAATTGTGCAGATAGAACGTACTTCATGGATGGCAAGACTGGGACTCTAGGACAGG CTTTCAATCCATTCTACCCGTGTTGTTGTTCTGAAATGAAAGTTTATCTCCAGTT TATATAGGTAGCCTTATCTTTGATGCTTCAATACCTGAGACCTGGCCAGTGTCCCT TTTTAGTGATTGTATGTGTGTGTGTGTGTGCATATGCAATTTCCCTTATAGCAATGG CACAGTGTATCACTGTTAATTAAGAAGAGAAAAGAAATGCCAAACATACGAATAA AGTCTGAATATATCTGTAACATTAAGAGTGTAGGTGTCTATCTTTGAAGATATGTC TTAAGGACAATGAAAGAGTCAGTGAGTAAGAGAAGAGAGTCTGGGATTTATACA AGATCAGTGTTACTTGATGGTGTAGGCTCCTAGGTATTTTATCTTTAGGATATACC GTCTATTACAAAAGCCAAGATTTTATAGATTTGGATCAACATTAGGGAACCTCATTC TAGGCAAGAGCCAGGTTTTGCCTTTATGTTAATATGACCTCAGCTGTGAGCTCCAT TTTGCCAGGCATCTTAAACTGCAACACATATCATTGGAATCTTCCGTTACAGTCT AATACATAGCCACACATTGGGAGCAAGAATGAAATCCAACCCCTGTCTTTGCAAAA ATGCAATGAGACAGTGTCTGCTTTGGGAGCAGGGAGTCAGAATTTTATTGTGGACA ATGGATAAGGTGAGTAAAAGGGCTTAAACATTTGTGCTTTCAAGCCATAGGCTAG GATAACGATAGTCAGAACTTTTTGATGAAGTCTGACCATGTCTACGCCATTTATAAA ATTTTGAAGCTTGTAAAGTATTACCCCAAAATGAGCAGTGTGAACCTCAAAGGGTTTA TCATTGTCTCTCAGGCAAAGGTAAATTTTGAATTATTTAGCAAAGGACTTTGAGCA ATTGGAGAGATACTCAGCTGCTGGTCTCTAGCGCTCTAACAGGGTGATGCCCCC CGCTCTGCCGGCACTGATGTTTTAGTTGCTGGATTATGAGGAAGTCTGGGGATTCC TTGGGGAGAAAAGGAGTGTATGACATATTGAAGCACAAACGACATATTGAAGAGACT CGGGGGCTGGGGTGATAAACTTCAGAGCCGTGGCTATTTACCAATTGGAGTGTAAG TATTTTAATATTTTAAACAAACATAATTGCCATTCTGGTATGTACCAACTTCATCTC AGATCTGTCTTAAAGAAATAGGCAAATTCTTTATTGCCTCTCTGAATGGTTCATAT AAATTTCCAGGCTCCCTTAGCTCATTCTAACATAAAACTGTATTAATAATAATGAA TGTAATTCATCAATAATTTTCCCTTTGTATAGCAAAATAGTCACAAGTGGATTGAGA TCAGAGTGATCACTCATATTTGTTCTGGGGAGAAGGGAGCCTGCTGTTTTGCTCCT GTTTTCTCCTAGGACTAGTATTTTAGCTTCAAATGATAATACCTTAGCACAGACTC TGATATTCCTCCTACATGCAGGAGCATTCTCTTGAATAATTTTGGGGATGCCAAT TCAAAATTTTCAAGCATGTATGATTACTTATTGGAATAATCACTGAGCAGCAAT AACTCCAGCAGTTACTTGTATCAAGGTAGAATCAAGAAATAGATGGTATGGACCAA ACTTGCTTCTCTCTAAATATGCATACCCAAGTGATTGGGTAAAAATGTTTGTGAAG GGCTTACATTTTCTGCAAGTCAGATGGTTTAAAGAGAAGTAGAAATTATGTGTGTTT TGCAGCATTTTGGTAATCTGTGTGGAGTGTCTGTAGATATTTCTCATGAGTTCAAG GGAATCCTTTTGTGGATTTTGTATGTTCTTATTGGCAGAGCTGCTGCTTGACTACAT GATGTCTTTGTATTAACACAAAACATGCCCTATCATCTGAGTGATTTTCTCTGC CAGACCCCTTTGTGCATCCACACTCTGCACCTCCAGTGACGGAGGACCTTCCCAC TGGATTCTAAGATTCCATGCCTTCCCAATGCATGGCAGTGTCTCTCATGCACATGG CAAACCTACTCTCTTGGATGTCAGTGCCTGAAATATTGAGGGAGTACATTATCT AGGCATGGTACCAGGGAGTCATTTAGACATGTAGGGAGTCTAGAAAGATCATTGCC CTGGGAGAGTGCTCAGCCATGCTGAGTTCTCCTACTTTGTTGCTCATTTCTGTGTG ACCTTAGGTAACATCCTCTTCAGGACTTTTTTTTTTTTTTTTTTTTGTACAGGGAG TCTCATCTGTCTATCCAGGCTGGAGTACAGTGGTGTGATCTCAGCTCACTGCAATC TCCGCTCTCTGGTTCAAGCAATCCTAGTGCTTCAGTCGCTGAGTAGCTGGGATT	
---	--

ACAGGCATGCGCCACTACGCCAGCTAATATTTGTATTTTCAGTAGAGATAGGGTT TTATCATGTTGATCAGGCTGGTCTTGAACCTCCTGACCTCAAGGGATCTGCCTACCT TGGCCTCCCAAAGTGCTGGGATTACAGATGAGAGCCACCAACCCTGGCCAGGACAT AATTTATTTAGGTGAATTGATTGTTGGAGGATTTTGATCCAAGCAATCAATGTCC CTTGGTGTTCCTTTCAAACAGCAGTAAGTGACCTGAATTTATTTTCCACATTTCCA AATCTTAATGAAAAATCAGACAATGGTCTATATGTTTCATTTGTGTTCTTACTTAATA AAATGTGGGTTTTAGACAATATTTTGCCAGTCATGAATTCCTATAGAAGGAACCTCT TTGGGAGAACAGACTAGTGATCTATAGACATGATGACCTCCAACCTCAGATCTTCTG TAGCTAACCACTGACCGGGAGAACATGTATGAAAAACATCTTCAAAGGCATTGAAA AATTAACATTTATCAAAAACAAAATACATTTTATTTTCATTTGAACCTAGACCTTTA CTATCTAATGGCTATGGTACTATTTAAATGTCAAAGTGTGATCTAGCATCAGCCTA ATCTGGTTAGAAATGCAAACCTCTTGGGCCACATCTCAGACTTACTGGACCAGAAGC TCTGTGGGTGGGACCAGAAAATCTGTGTTTCATTCACATGCCCTCCAGGGGATTGT CCTGTCTAAAGTTTGAGAATCATGGAAGCTTTTTAACCTCTCATTATAGCTTTATAA GCAGCAACTCACTGGATTCTATCAACATCCTGTGAGTGTCAATTTGGACAAGTATA TTTATACCCATTTGATGCATGGTAGGCACACAGATGAGTCAAATGACTTGAAAGGAA TAGAGTTTTACATAATATACTTTTATATATTTATACTTCTAATATATTTTATACTTT ATAACAGATTTGACTGTTTTATATATTTGCATATAAACATTATATCAGTTTTCTCCTC CACTAAGGCTGACTCCAATTTTACTCCAATTTTACTACCAATTTTGGAAAGAAAGC CTACCTATCACTCATGTTCTCTCAAGTACCCTCTAAAACCTATTAGTTAGATGACTC TATTTAATTTTCCATTTATTTGCCCGTTTCTGTCTACCTTTCCCCCAAAATGTAA CTGCTACCTTGCTCAAAGGATGTGTCTACTTGGGATATCTAGCACACACATTTTA TGAGATTTTAAAGACAACATAAATGGTAACTATATATTTAATACAATTTTGA GACAAAATTTAAATTTAAAGGAAGAAAAAATTAACCTAACCCATAAATCTC CCACCCATCATTAGCATGTAGTTTGTGTTAGATTCTATACCAATAAGTAGAATTGT ACAAATTTGATATCATGTAATACATGTCATTTTGTAACTTTTTCTTTCTTAAAT ATATCTATATATCATAAACATTTTCTATGTCTATATTATTTTAAATTTGTAATAC CCAGAGTTCTCCAGAGAAACAGAAATTAATAGGATCTCCCCCTTGGAGATTTCTCAT CTTTTCTCTCTCGATAGATACAGATAGATACATAAGTCTATCTCTCTATCT CTATCTTTATCTCTAAAACACCTATCCATAGATAGACATTTTATAGGAATTTGGCTCA TGTGTTTGTGGAAGCTTGCAAGTTCAAATGTGCAGAGTAGGTGGGCAAGCTACAG GGAAATGTTGATGTTGCAGTTCAGTCTGAAGGCAGGCTCCTTGCAGAATTTCTTCT TTTTCTTAGCAGTCTTACTTCCCCTTCTCTTCCCCTTCTCCTTCCCCTTCCCCTT CTTCTCTCCTTCCCCTTCTTCTTATTCTCCTTCACAGACTTATTTTTAAGGCCCT GAGCTGATTAGATAAGACCCACTCACATTATGAGGGATAATCTGCTTTACCTGTAG TCTACTAATTAATTTTAAATCTCATCTAAAAAACACCTTCATAGCAGCATTCAGA CATGTTTACCAAATATCTGGGCACCATGGTTTAGCATATTGATGCAGAAAAATTGA TTATCATAAATAATATTATTTTTTTTTGAGATGTAGTTTCACTCTTGTCAACCCAGGC TAGAGCGCAATGCTGCAATCTCAGCTCACTTCAACCTCTTCTCCTTAGGTTCAAGC GATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCATGACCACGC CCGGTTAATTTTGTGTTTTTTTAGTAGAGATGGGGTTTACCACGTTGGTCAGGCT GGTCTCGAAGTCTGACCTCAGGTGATCCACCCGCTCAGCCTCCCAAAGTGTG GATTACAGGCGTGAGCCACTGTGCCAGCAATAATATTAATTTAATGGGTGTGTT CATTTTCATTTTATATATGACCTACAATTTAACCAATCCCCTAAGGCTGGATGTTCA GGTTCTTAATATTTTTTGCCCGTATTTACAGACACCTTTGACTATTGGATTATTT TGTTCTTCAGGAACAATATACAAAGTGTGGAAGAAATGTATATTTCTAATCATTG GAAAAATAACACTGAGCAGAAATAACTCCAGTAGCCATTTGTATCAGAGGAGGTAG AATCAGGAAATAGATGGTATGGGCCAGACTTTCTCTCTCTTAAAGAGATTGACT TCATATTGCCAAATTGCCCTTCTAGATGTCTTTACTCATCCAACCTACAATTCAAAG GTTTGGGAGGGTAAGCAATGCCAGGCCCATCTTGATCATCCCCTTTCTTTCTCAGC CTGTCACTCCCTGGGAACGGGCATGATGTTGAGTTCTGGGCCACCTCCTCAATTG AAGAAGTGGCGGAAGCTGGTCTTGAGGCACCTCGTTGGCTGCAACTGTATATCTAC AAGGACCGAGAAGTACCAAGAAGCTAGTGCGGCAGGCAGAGAAGATGGGCTACAA GGCCATATTTGTGACAGTGGACACACCTTACCTGGGCAACCGTCTGGATGATGTGC GTAACAGATTCAAACCTGCCGCCACAACCTCAGGTAACCATGATCATGTGGGCCCCGA GCTGAGGCGAAAGGGATCTTGACTGGGAATGTTAGGGTCTGGGTTCTACTGATAGC AACGTTGCTAAACATCTAGTTAATCTTCAGCTAATCACATCCCTTTTGTAGACATC ACTTTTTTTGAGATACACAATAGAAACAGAAATGGCCTCTATAAAAGTCCAATAAA TTTTAGACCAGAGTGCATTAAGGGCTTTGGCTTTGGGAAGTATGAATTGCTATAC	
--	--

AGATGGAAGATACTGAATTTTGCCCAAGCAGCAGTTTATTATTATCATCCTGGTGC CCTATTTCCTTTGTTAAAGTCAAAGAGCCACCTTTACCTTTTATTTTAAATGGTACA TGGGACAGCTAAGGCTAAGAAGATTGAAGAAAAGAAAATAATGAAGGTTTAAAAAAG CCACATCTTTGATCCCTCACTGTCTACTTCTTCTTTCAGCAATATTCCCTTCACTG TGGTTTCATCCATGGGTCAAGATTCATTGATTCATTCACTCAAAATCATTTCATCTTAG CAAAAACAATATATCACATAATCTGATGTTGAACTATAAAGGTTTCATCAGGTCAT TCATTACCCCTGTCCACAAGCTGTGAATTATTATCTCTTTTCTGGTTGTATTTTG GATTACAATCATCTTGAGTCAAAGCTGGAAACTGAGTGGAACTCTCTGGGAAAGAC TCAAACCTCCTTAAGCTATACACCTCTTTTCCCCATCAGATTTTCTTCTCTCAGT TTCCACCAAAAATGTGCTCTTGGATTTTTCATATGAATGTATAATGTACCTCAGGCC TATAAGTATTTTAAAGGGATCAAAATCTTAGTTTAAATGGAGGACATTTTATGA TGGACTCCTACAGCATCCATCAGAATATGTAAGATGATGAGGAATGTCTTCTCTGTG TTCCCAGATCTCATGCCACAGAGGCCCTTGCTTACTCTATGTTGAATTGTATTTG GAAAAAAAAAAAAAAAAAAAACTAGGGCTAGCAAAATTGAAAAAGATAAAAAA CGAAAGAAGCCACATGTAAACATACTGTGTTTACTCTTCTAAAAATATTAAAAATG AAAAGATCCAAAATCAAATTAATATTCCCCTGGAATTTTCATATCTATTTTCACTGAC TGTGGAGTGAATCTCACCACGAAAGTTGCTGCAGTCTTGTATAAGTTTCACATAGT TTTACTGTGTTTGTGCTATGTGAGAATAAACTACTGTGCATAAAATCTTGCTGTT GAGCCATGTGTGAATTAGCTGTGTGATGTTACCTCCCTGTTACTACCAGGCTGGTT TAGGATATCATTTCTGTATGTGGCACCAGGATTAGACCAATGACAGAAAAAGAAAG TGCTCTCCCTGCCAAACTGGCCAATAAACTGTTCCACATATCCCAGACTCAGGGT TACCTAAACAACCTGTGTTTAAAGAGAACAAAACAAAAGCCTCTGACATAGTCTT ACTCCTTGCCAAATTCGTGAGAAAGCTGATGGATTCAAATTTCCCCCAATATGAATC CCGTATTTACATTATTTCTCTATTTTGAAGTACTTTTCTTAAAGACTTTT TAAATAGTTTCCCACTATCGAGGCTTCTTAGAGGAAACATTTCTCATTATTTCCCC TTGGCTATTTGAAAAGGAATTTGTTCTTCTTTTCTCCATCTCTTAACACTACTA CTACTAACAAATAGTAACAACAATAGTAAGTACAGTAGGGTTTTTGTGTTGTTTTT AACTTAAGACATATTTCTTGTCTGGATACCAAAATATGTTTCACAGAGGCATCT ACTTAGATGGGGTGCAGATGACACAGTTGTTAATCTGGCAGGTACCTCTTGCTTC TTCACTGCTGGGGCTACTCAGTGAGTGGCAGGAAGGTTGATTGCTTTCCCCCTT TTCTTTTGTCTCCTGGGCTCCTTCCAGATGATGTGACGGGCCATGAAACAAAGACT CTTTTCAGCTGTGCGGTGTGCATAGAACTGGCTGCGGCTTCTAGCTTGTACATCT CCGGTCTGAAGATGATCAAATAATGAGCAACACATCCAGGTTATAGGGAACCGGG AAACACCCCGCAGCTGGGTGTACCCAGCCCTCAGAGTGCACATTGGTGTGTTT GTCCTAGTGGACTTCGGAGTAGGCCAGTGCCCTTCTGGTCAGTTCTCAGTGGCCCA CATTCACTCTTTAAAGGCAGAGCATGCTAACGGGAGGTCCAGGCTTCCGCTGAGG CCAAATACACCCCAAAAGCTCATCTGTTATAGCCTGATATGAAATCGGTTTCTTTC TGCAACTGACCTGACTCATAGAAAGTGAAGCCTGGCTTTTTCATAAGTGAAGTTTGG CAGGCAAGGGAGGCAGGAAATCCAGAGGAGAATGAGCCTGTAAAGCATGGCTCCTT CCAGCCCTTGTTACTTCTCTGCCCAAGTGTGGGGGAGGGTCTGTCTCTTGGCAT CTGGGGCCAGCAAGAGTTTCAAGGTTTGGTAGTCTCTGCTTGGTCCATATGCAAAA CACATGTATGTGTATACATTATTAATGGCAAGGGGTTTCTGAAACTGAGAGGGAG TAAGGAGACTTCTCATCTGCTCTTGAAGAAGCAAGGAATGAAGCCAGTTCAGTAG ACTGATTCCTGAGGCTTTGGGGCAGGAACTTTTCTTTCTCCATATCCCCATGGAG ATGGTTCATTTACCCTGAATTAAGATTTGGCCCTTCGGTGCAGTGCCAAGGCAGTT TAAAGAGAAGAAAAGTAATTTCTGATCATTGACTAAGATCAAGGTAAATCATGACA CTTATCCTTTCTATGATTTGCCAGTGACATGTTTTCTTAAGCCCAGAAATGATTTA TTGATCGCAGCAGCCAGAATATATCACACTAAAACAGATCAGCCTGCCACTGTCTT CTCAGGCTCTCATGATTAAAGTGGCCTGCTTTAAAGTAGACTCAATGTGAATAGG TCTCCATGACCTCTGCCTCACTGCGTAGCACTCACATCCTCACCCTCTTGCCT CTGGCTTCCCTGCGGTTCTTTCAATATGCCAGGCATGCTGGAACCCCGGAGCCTTT GCACTGGCTGTTCCCTCTGTCTGTAAACAGTCATTGCGAGAATCAACGCATGACTAA TAGCCTCACTTCCATTGAGTCTTGACATTAGGAATGGATATACATGTCTATATTGG GAAACCACAATAAAAAATTGATGGTAGAGATGCAAAATATGAGCAAGATAACAGGGTG GGGGCAGGGGAGAGAGGTGAGTGGAGGACTTGGCACAGTAGCCTCTTAAATGGCAC AATAGCCTCTTAAATTTTGGTTAAGAAATCATTCACATTGATAAGTATGGCAGGA TAAAGGTGTCCATGAGTGAATCCCGGGAACCTGTACTTTATGTGGCAAAAGGGAC TTTGCAGATGTGATTAACCTAAGGGGCTTGAGATGGGAAGATTTTCTGTTTTTA TCAGTAGGCTTGATATAATTAAGGGGCTTATAAGAGGGAAGCAAGAGTGTACAG	
---	--

	AGTCAGAGAAAAGAGATGAAATGACAGATGTAGAGGTTGGAATGATGTGGTCAGGAA CCAGGGAAAAGCAGGGGGTATCTAGAAGCTGGAAAAGACAAGGGAATAGGGCTTCCC CTAGATTCTCCAGAAATACAGCCCTATTGATATCTTGAGTTTAGTCCAGTGAGACT TATTTTAGACTTCTGACATTTGGAACTGTAATATAATACCTTCATGTTATTTTAT TGCTGTTGTAACAAATAAACCACAAACACAGTTCTACTAATTTCTTTCTCAAGGTAG CTTCTCAATTTTGCCAACGCTGGTTACCATACTACTTAAAGTTTCATTTTGAGTC TCTGAAAACCTCACATCTCTTAACTGCTCTACTTTTTCTTGGCTTTGTATAGT GCTTATGTTCTGCTACACTTTGTAATTTATTGATTATGCTTACCATGGGCAGGGAT TCTTAACTGTTTTATTTATTTATATATCTTAAACATTGAAAACACTGGCATGTAGT AGATGCTTAATAAGTAATTGTTGACTCAATCGATAAAATATACTAGAACATACAAG ATTTTCCCAATGTAACATAACTAGTAAGAGGCTGAACCGGGATTTGAACTCAAAAT TCATTCCTGAACCTTCTTCTAGCAGCCACATTGAGGAAGAAAATTACCAGGGCTGT GTTCTCAACACAAGTGTTTTCCGAACACAGAAATTAAGGCTGGTGGCCATGTAT CAGTGCTGTGTTTTATGAGCCCCCTTTCAATCTCTTTCTTTTCATATTGTGTTGA TGCTGTAGCTTCTACTGGTCATGTTATTTTTTTGTTTCCCAAGACGGAATTATGTG GCTTTATCTTTAATGTTGCATTATCAATACTTATAATAAATAATATTATGTATTAC TCAATATTCATGATTAATAGTGTTACTATTGGTTATTTAATAATGTTTAACTTACA TTAGCAGTTGTTACTATTTTTATGATGCTAAATTACTAACAGCTAAAACAACCTTCT ATATTAATAAGTATATTTGAGTGCCACTCAAGAGATAATGAGTACCTTACAAAGAA GAAATCTGTTTCTCACCTTTGCGTCATTAAACAGATCAGGATTTGGAGAATTAAG CCCTAAGTAATAGTGTTATTATTTTGATCTCACCCCTTTTTTCTTATGAAATGGA ATACTTTGGTTATCAGAAGCCACTTTAAGCATATATATATATATATATATATAC ATATATATATATATATATGTCTAATCCGAATAAAAAATAGCATTTCATGGAGGTTTC TTTTGGAGCCTTTGGTAAAACACTCCATCGTGGGTCTCTGTCAAGATATCTGAAAA CTTTTTCTTGGCTTCTGGCTTTGAACAAAGTTTCAGAGTAACAACAAGGCTTCATT GTGCACTGAAATTTCTGTAAGGCAACATTCAATCAAGTGTTGATTTCGATTTCACC ATCCAAGAATAACAACAGTTATTTATATAATTTTATCCAGCTTCTGTTTTTTCCT ATCCATTTCACCCCTTTCACCCACCCCTGCTGAAACACTGGAGCTTGTTTGGGATG GGGTTGGGGTGCCATGCAGACTACATACACATACAGATGTTTTTCTTTTTCTTTTC CCGGTCTTGCTATGGGATAGACAGACTGGACTTTTTCTTATTAACAATATTATTTA AAAGCTTGGAATTTATTATCATTTAATCATTTGTATGTAATGAAATAGGTCTCCAT GGTAAAGATGTGTTTATTGACCAGCGGTTAGCTTTATTCAAATTAGGGTGACCATA GAAGACCAAGGACTATGATATAATGTACAATCCTAAGTGGTTTGATTTAAATAAAA AGAAAGACCAGGCATTTAGCTAAAATCCCCACCAAGCCCAATGACTAGATGGGC ATCCATATGACTCAATGAAATTTCTATGATCTTAAATGGCCATCTGAGTCCGTGA AATATAGGACTAATATTCAATCCTTATTGAGAAAGCCTTGTTAATAGCTTGAAT TGAGTTATATGGGATAGGAATGTTTCATATCTTTATGACAATATATGCCACCTAAGC TACATAACCAGCTGTGTTAGCTAAAATACTCTAAAGTGTAAAAAATCATAGTTTTTC TATTAAGGAAGTCATGATTGTTAAAAATAATTTTTAAATAGTGTGCCTAGATTCT TCTAGTATAATATATAATTTTTTTTTTTTTTTTTTATTTTGAGACAGAGTCTTGCTCT GTCACCCAGGCTGGAGTGCAGTGGCTGGAGTTGCTCACTGCAACCTCGCCTCCCGG GTTCAAGCGATTCTCGTGCCTCAGCCTCCCAAGTAGCTGAGATTACACGTGCCAC CACTATGTCCGGCTAATTTTTTTGAATTTTTTAGTAGAGACTGGGTTTCATCATGCT GGCCAGACTGGTCTTGAACCTCCTGACCTCAGGTGATCTGCCCACCTCGGCCTCCCA AAGTGCTGGGATTACAGGCATGAGCCATTGCGCCAGCCGATATATAAATTTTTAT ATGGCTCCATGATCTTCTCTACATTTAATGACAGAAGCTGGTGGAGGGGAAGAAAGA GATGGGACTAAGCCAGAGATCAATATACATACAACTATACTTTGACCAAAAAAAGG GAGATTGACTGGCAGGGGAATTAATAGTATGCAGAAGAGCAAGGTGAGTCCAGTCA CTGTCTATTATTCAAAAAACAGCCTTTCAGGAGAAGTTGCAACTGAATTTGGGACTG TGGGCAGATAAGTCACAGGAATGATTCTATTGTGTATCCTGAAGTCATCCATCCAG CTAGGAGTCAGAGGTGCAGGCTGAAAAGACATTGCCCTAGAGTGGGGAAGTGCCA AAATCTAGCCAGGATATTAGGCCAAGAGAAAAGACCTCAGGCACAGGGGAAGCCAG CTTCAGA	
929	AACGAACTCCATCTGGGATAGCAATAACCTGTGAAAATGCTCCCCGGCTAATTTG TATCAATGATTATGAACAACATGCTAAATCAGTACTTCCAAAGTCTATATATGACT ATTACAGGTCTGGGGCAAATGATGAAGAACTTTGGCTGATAATATTGCAGCATTT TCCAGGTAAGAAAATTTATTTTTTAAATCATGTTTTTAAATTTACACAAAGACCG	HAO1 exon 1

930	TGATTCTGAAACTCTAAAGCCTTTTATTTTATTTTATTTTAAATTCTAGATGGAA GCTGTATCCAAGGATGCTCCGGAATGTTGCTGAAACAGATCTGTCGACTTCTGTTT TAGGACAGAGGGTCAGCATGCCAATATGTGTGGGGGCTACGGCCATGCAGCGCATG GCTCATGTGGACGGCGAGCTTGCCACTGTGAGAGGTAGGAGGAAGATTGTCACCAC AGGGACAGAAGGAGGCTAACGTTTATCG	HAO1 exon 2
931	GGAGGGTAAGCAATGCCAGGCCCATCTTGATCATCCCCTTTCTTTCTCAGCCTGTC AGTCCCTGGGAACGGGCATGATGTTGAGTTCTGGGCCACCTCCTCAATTGAAGAA GTGGCGGAAGCTGGTCTGAGGCACTTCGTTGGCTGCAACTGTATATCTACAAGGA CCGAGAAGTCACCAAGAAGCTAGTGCAGGCAGGCAGAGAAGATGGGCTACAAGGCCA TATTTGTGACAGTGGACACACCTTACCTGGGCAACCGTCTGGATGATGTGCGTAAC AGATTCAAACCTGCCGCCACAACCTCAGGTAACCATGATCATGTGGGCCCCGAGCTGA GGCGAAAGGGATCTTGACTG	HAO1 exon 3
932	ACGTATTTCTAATTGGCAAATTTCTCATTATTTATGCATTTCTTATTTTAGGATGAA AAATTTTGAAACCACTACTTTATCATTTTCTCCTGAGGAAAATTTGGAGACGACA GTGGACTTGCTGCATATGTGGCTAAAGCAATAGACCCATCTATCAGCTGGGAAGAT ATCAAATGGCTGAGAAGACTGACATCATTTGCCAATTGTTGCAAAGGGCATTTTGAG AGGTTTCGTTTATTTCTCTACTTGAATTCATACCTGACTTTGTGATCCTTTGTG	HAO1 exon 4
933	CTGCCGTGTTAAGTTACAGTTTCCCTAAGGTGCTTGTGTTTACTCTCTCCAGGTGATG ATGCCAGGGAGGCTGTTAAACATGGCTTGAATGGGATCTTGGTGTGCAATCATGGG GCTCGACAACTCGATGGGGTGCCAGCCACTGTGAGTTTTGGCAGACGCTAAGATTT CCTTTTGGAGTTCCCATTTCCATC	HAO1 exon 5
934	TAACAATTCAGTGTTAATAGAGTCACATTATTGAACTTTTCTTTCCCCAGATTGAT GTTCTGCCAGAAAATTGTGGAGGCTGTGGAAGGGAAGGTGGAAGTCTTCTCGACGG GGGTGTGCGGAAAGGCACTGATGTTCTGAAAGCTCTGGCTCTTGGCGCCAAGGCTG TGTTTGTGGGAGACCAATCGTTTGGGGCTTAGCTTTCCAGGTAAGTGGACAAAGA AATGAATATATAAAATAGACAACCTTGACAGTAAACAAATGAATAAAACAAGTCAG ACTGATTTAGTTCTGAATCACTCTGTATCTTTTCACTTGGTTAGGGGGAGAAAGGT GTTCAAGATGTCTCGAGATACTAAAGGAAGAATCCGGTTGGCCATGGCTCTGAG TGATAAGACTCATTTCTGTTTACAACCTTTCTTTTATGATCCTTTAAGT	HAO1 exon 6
935	TGATTATTATTGCATTCAGTTTCATATTAATGTATGCATTATTTTTTTCAGGGTGCC AGAATGTGAAAGTCATCGACAAGACATTGGTGAGGAAAAATCCTTTGGCCGTTTCC AAGATCTGACAGTGACACAATATTTCCCATCTGTATTATTTTTTTTCAGCATGTAT TACTTGACAAAGAGACACTGTGCAGAGGGTGACCACAGTCTGTAATTTCCCACTTC AATACAAAGGGTGTCGTTCTTTTCCAACAAAATAGCAATCCCTTTTATTTTCATTGC TTTTGACTTTTCAATGGGTGTCTTAGGAACCTTTTAGAAAGAAATGGACTTTTCATC CTGGAAATATATTAAGTGTAAAGGAGAAACATTGAAAATGTGTTTAGACAACGTC ATCCCTGGCAGGCTAAAGTGCTGTATCTTTTAGTAAAATGGAGGTAGCAAACAC TAAGGTGAAAAGATAATGATCTCATTGTTTATTAACCTGTATTCTGTTTACATGTC TTTAAACAGTGTTCTTAAATTGTAAGCTCAGGTTCAAAGTGTGTTGTAATGCCTG ATTACAACTTTGAGAAGGTAGCACTGGAGAGAATTGGAATGGGTGGCGGTAATTG GTGATACCTTTTGAATGTAGATTCCAATCACATCTTTAGTGTCTGAATATATACC AAATGTTTTAGGATGTATGTTACTTCTTAGAGAGAAATAAAGCATTTTTGGGAAGA A	HAO1 exon 7
965	MSNKEKNASETRKAYTTKMIPRSHDRMKLLGNFMDYLMGTPIFFELWNQFGGGID RDIISGTANKDKISDDLLAVNWFVKMPPINSKPQGVSPSNLANLFQQYSGSEPDIO AQEYFASNFDTEKHQWKDMRVEYERLLAELQLSRSDMHDLKLMYKEKICIGLSLST AHYITSVMFGTGAKNNRQTKHQFYSKVIQLLEESTQINSVEQLASIIKAGDCDSY RKLRIKRSRKGATPSILKIVQDYELGTNHDDEVNVP SLIANLKEKLGRFEYECEWK CMEKIKAFASKVGPYYLGSYSAMLENALSP IKGMTTKNCKFVLKQIDAKNDIKYE NEPFGKIVEGFFDSPYFESDTNVKWLPHHHIGESNIKTLWEDLNAIHISKYEEDIA SLSEDKKEKRIKVYQGDVCQTINTYCEEVGKEAKTPLVQLLRLYLSRKDDIAVDKI IDGITFLSKKHKEKQKINPVIQKYP SFNFGNNSKLLGKIISPDKLKHNLKCNRN QVDNYIWEIKVLNTKTMREKHHYALSSTRFLEEYYPATSENPPDALAARFRK TNGYEGKPALSAEQIEQIRSAPVGLRKVKRQMRLEAARQQNLLPRYTWGKDFNIN ICKRGNNEFVTLATKVKKKKEKNYKVVLGYDANIVRKNTYAAIEAHANGDGVIDYN DLPVKPIESGFVTVESQVRDKSYDQLSYNGVKLLYCKPHVESRRSFLEKYRNGTMK	(Cas12i1 of SEQ ID NO: 3 of U.S. Patent No. 10,808,245)

	DNRGNNI QIDFMKDFEAIADDETSLYFFNMKYCKLLQSSIRNHSSQAKEYREEIFE LLRDGKLSVLKLSLSNLSFVMFKVAKSLIGTYFGHLLKKPKNSKSDVKAPPITDE DKQKADPEMFALRLALEEKRLNKVKSKEVIANKIVAKALELRDKYGPVLIKGENI SDTTKKGKSSSTNSFLMDWLARGVANKVKEMVMHQGLEFVEVNPNTSHQDPFVH KNPENTFRARYSRCTPSELTEKNRKEILSFLSDKPSKRPTNAYYNEGAMAFATYG LKKNDVLGVSLKFKQIMANILHQSEDDQLLFP SRGGMFYLATYKLDADATS VNWN GKQFWVCNADLVAAYNVGLVDIQKDFKKK	
966	MSISNNNIPYNPKLLPDDRKHMLVDTFNQDLIRNNLHDMIIALYGALKYDNIK QFASKEKPHISADALCSINWFRVLTNERKPAIESNQIISKFIQYSGHTPDKYALS HITGNHEPSPHKWIDCREYAINYARIMHLSFSQFQDLATACLNCKILILNGTLTSSW AWGANSALFGGSDKENFSVKAKILNSFIENLKDEMNTTKFQVVEKVCQQIGSSDAA DLFDLYRSTVKDGNRGPATGRNPKVMNLFSDQGEISSEQREDFIESF QKVMQEKNSKQIIPHLDKLKYHLVKQSGLYDIYSWAAAIAKNANSTIVASNSNLNT ILNKTEKQQTFEELRKDEKIVACSKILLSVNDTLPEDLHYNPSTSNLGNLDVFFD LLNENSVHTIENKEEKNKIVKECVNQYMEECKGLNKPPMPVLLTFISDYAHKHQAQ DFLSAAKMNFIDLKIKSIKVVP TVHGSSPYTWISNLSKKNKDKGMIRTPNSSLIGW IIPPEEIHDKQFAGQNP IIAVLRVYCINNKMWHHFPFSDSRFFTEVYAYKPNLPY LPGGENRSKRFGYRHSNLSNESRQILLDKSKYAKANKSVLRMCNMTHNVVDFPK TSLNIRIKTDKNNSPVLDDKGRITFVMQINHRILEKYNNTKIEIGDRILAYDQNS ENHTYAILQRTEEGSHAHQFNGWYVRVLETGKVT SIVQGLSGPIDQLNYDGMVTS HKFNCWQADRSFVSQFASLKISETETFEAYQAINAQGAYTNLFLYLRLRKALR VCHMENINQFREEILAI SKNRLSPMSLGSLSQNSLKMIRAFKSI INCYMSRMSFVD ELQKKEGDLELHTIMRLTDNKLNDKRVEKINRASSFLTNAKHSMGCKMIVGESDLP VADSKTSKKQNVDRMDWCARALSHKVEYACKLMGLAYRGIPAYMSSHQDPLVHLVE SKRSVLRPRFVVADKSDVKQHHLNLRMLNSKTKVGTAVYREAVELMCEELGIH KTDMAKGVSLSDFVDKFIGEKAIFPQRGGRFYMSTKRLTTGAKLICYSGSDVWLS DADEIAAINIGMFVCDQTGAFAKKKKKEKLDDEECDILPFRPM	(Cas12i3 of SEQ ID NO: 14 of U.S. Patent No. 10,808,245)
1024	ATGCTCCCCCGGCTAATTTGTATCAATGATTATGAACAACATGCTAAATCAGTACT TCCAAAGTCTATATATGACTATTACAGGTCTGGGGCAAATGATGAAGAACTTTGG CTGATAATATTGCAGCATTTTCCAGATGGAAGCTGTATCCAAGGATGCTCCGGAAT GTTGCTGAAACAGATCTGTGCGACTTCTGTTTTAGGACAGAGGGTCAGCATGCCAAT ATGTGTGGGGGCTACGGCCATGCAGCGCATGGCTCATGTGGACGGCGAGCTTGCCA CTGTGAGAGCCTGTGAGTCCCTGGGAACGGGCATGATGTTGAGTTCCTGGGCCACC TCCTCAATTGAAGAAGTGGCGGAAGCTGGTCCCTGAGGCACCTCGTTGGCTGCAACT GTATATCTACAAGGACCGAGAAGTCACCAAGAAGCTAGTGCGGCAGGCAGAGAAGA TGGGCTACAAGGCCATATTTGTGACAGTGGACACACCTTACCTGGGCAACCGTCTG GATGATGTGCGTAACAGATTCAAACCTGCCGCCACAACCTCAGGATGAAAAATTTTGA AACCAGTACTTTATCATTTTCTCCTGAGGAAAATTTTGGAGACGACAGTGGACTTG CTGCATATGTGGCTAAAGCAATAGACCCATCTATCAGCTGGGAAGATATCAAATGG CTGAGAAGACTGACATCATTGCCAATTGTTGCAAAGGGCATTTTGAGAGGTGATGA TGCCAGGGAGGCTGTTAAACATGGCTTGAATGGGATCTTGGTGTGCAATCATGGGG CTCGACAACCTCGATGGGGTGCCAGCCACTATTGATGTTCTGCCAGAAATTGTGGAG GCTGTGGAAGGGAAGGTGGAAGTCTTCTGGACGGGGGTGTGCGGAAAGGCACTGA TGTTCTGAAAGCTCTGGCTCTTGGCGCCAAGGCTGTGTTTGTGGGGAGACCAATCG TTTGGGGCTTAGCTTTCCAGGGGGAGAAAGGTGTTCAAGATGTCTCGAGATACTA AAGGAAGAATTCCGGTTGGCCATGGCTCTGAGTGGGTGCCAGAATGTGAAAGTCAT CGACAAGACATTGGTGAGGAAAAATCCTTTGGCCGTTTCCAAGATCTGA	HAO1 cDNA
1082	rArGrArArUrCrCrGrUrUrUrCrArUrUrGrArCrGrGrCrGrArG rCrArUrCrCrUrUrGrGrArUmA*mC*mA*rG	3' end modified RNA guide targeting HAO1 sequence of

		SEQ ID NO: 1047
1083	mA*mG*mA*rArArUrCrCrGrUrCrUrUrUrCrArUrUrGrArCrGrGrCrGrGrArGrCrArUrCrCrUrUrGrGrArUr mA*mC*mA*rG	5' and 3' end modified RNA guide targeting HAO1 sequence of SEQ ID NO: 1047
1084	rArGrArArArUrCrCrGrUrCrUrUrUrCrArUrUrGrArCrGrGrGrGrArArGrUrArCrUrGrArUrUrArGmC*mA*mU*rG	3' end modified RNA guide targeting HAO1 sequence of SEQ ID NO: 1026
1085	mA*mG*mA*rArArUrCrCrGrUrCrUrUrUrCrArUrUrGrArCrGrGrGrGrArArGrUrArCrUrGrArUrUrUrArGmC*mA*mU*rG	5' and 3' end modified RNA guide targeting HAO1 sequence of SEQ ID NO: 1026
1086	rArGrArArArUrCrCrGrUrCrUrUrUrCrArUrUrGrArCrGrGrCrArArArGrUrCrUrArUrArUrArUrGrAmC*mU*mA*rU	3' end modified RNA guide targeting HAO1 sequence of

		SEQ ID NO: 1025
1087	mA*mG*mA*rArArUrCrCrGrUrCrUrUrUrCrArUrUrGrArCrGrGrCrArAr ArGrUrCrUrArUrArUrArUrGrAmC*mU*mA*rU	5' and 3' end modified RNA guide targeting HAO1 sequence of SEQ ID NO: 1025

In some embodiments, the gene editing system disclosed herein may comprise a Cas12i polypeptide as disclosed herein. In other embodiments, the gene editing system may comprise a nucleic acid encoding the Cas12i polypeptide. For example, the gene editing system may

comprise a vector (*e.g.*, a viral vector such as an AAV vector, such as AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAVrh10, AAV11 and AAV12) encoding the Cas12i polypeptide. Alternatively, the gene editing system may comprise a mRNA molecule encoding the Cas12i polypeptide. In some instances, the mRNA molecule may be codon-

optimized.

II. Preparation of Gene Editing System Components

The present disclosure provides methods for production of components of the gene editing systems disclosed herein, *e.g.*, the RNA guide, methods for production of the Cas12i polypeptide, and methods for complexing the RNA guide and Cas12i polypeptide.

A. *RNA Guide*

In some embodiments, the RNA guide is made by *in vitro* transcription of a DNA template. Thus, for example, in some embodiments, the RNA guide is generated by *in vitro* transcription of a DNA template encoding the RNA guide using an upstream promoter sequence (*e.g.*, a T7 polymerase promoter sequence). In some embodiments, the DNA template encodes multiple RNA guides or the *in vitro* transcription reaction includes multiple different DNA templates, each encoding a different RNA guide. In some embodiments, the RNA guide is made

using chemical synthetic methods. In some embodiments, the RNA guide is made by expressing the RNA guide sequence in cells transfected with a plasmid including sequences that encode the RNA guide. In some embodiments, the plasmid encodes multiple different RNA guides. In some embodiments, multiple different plasmids, each encoding a different RNA guide, are transfected into the cells. In some embodiments, the RNA guide is expressed from a plasmid that encodes the RNA guide and also encodes a Cas12i polypeptide. In some embodiments, the RNA guide is expressed from a plasmid that expresses the RNA guide but not a Cas12i polypeptide. In some embodiments, the RNA guide is purchased from a commercial vendor. In some embodiments, the RNA guide is synthesized using one or more modified nucleotide, *e.g.*, as described above.

B. Cas12i Polypeptide

In some embodiments, the Cas12i polypeptide of the present disclosure can be prepared by (a) culturing bacteria which produce the Cas12i polypeptide of the present disclosure, isolating the Cas12i polypeptide, optionally, purifying the Cas12i polypeptide, and complexing the Cas12i polypeptide with an RNA guide. The Cas12i polypeptide can be also prepared by (b) a known genetic engineering technique, specifically, by isolating a gene encoding the Cas12i polypeptide of the present disclosure from bacteria, constructing a recombinant expression vector, and then transferring the vector into an appropriate host cell that expresses the RNA guide for expression of a recombinant protein that complexes with the RNA guide in the host cell. Alternatively, the Cas12i polypeptide can be prepared by (c) an *in vitro* coupled transcription-translation system and then complexing with an RNA guide.

In some embodiments, a host cell is used to express the Cas12i polypeptide. The host cell is not particularly limited, and various known cells can be preferably used. Specific examples of the host cell include bacteria such as *E. coli*, yeasts (budding yeast, *Saccharomyces cerevisiae*, and fission yeast, *Schizosaccharomyces pombe*), nematodes (*Caenorhabditis elegans*), *Xenopus laevis* oocytes, and animal cells (for example, CHO cells, COS cells and HEK293 cells). The method for transferring the expression vector described above into host cells, *i.e.*, the transformation method, is not particularly limited, and known methods such as electroporation, the calcium phosphate method, the liposome method and the DEAE dextran method can be used.

After a host is transformed with the expression vector, the host cells may be cultured, cultivated or bred, for production of the Cas12i polypeptide. After expression of the Cas12i

polypeptide, the host cells can be collected and Cas12i polypeptide purified from the cultures etc. according to conventional methods (for example, filtration, centrifugation, cell disruption, gel filtration chromatography, ion exchange chromatography, etc.).

In some embodiments, the methods for Cas12i polypeptide expression comprises
5 translation of at least 5 amino acids, at least 10 amino acids, at least 15 amino acids, at least 20 amino acids, at least 50 amino acids, at least 100 amino acids, at least 150 amino acids, at least 200 amino acids, at least 250 amino acids, at least 300 amino acids, at least 400 amino acids, at least 500 amino acids, at least 600 amino acids, at least 700 amino acids, at least 800 amino acids, at least 900 amino acids, or at least 1000 amino acids of the Cas12i polypeptide. In some
10 embodiments, the methods for protein expression comprises translation of about 5 amino acids, about 10 amino acids, about 15 amino acids, about 20 amino acids, about 50 amino acids, about 100 amino acids, about 150 amino acids, about 200 amino acids, about 250 amino acids, about 300 amino acids, about 400 amino acids, about 500 amino acids, about 600 amino acids, about 700 amino acids, about 800 amino acids, about 900 amino acids, about 1000 amino acids or
15 more of the Cas12i polypeptide.

A variety of methods can be used to determine the level of production of a Cas12i polypeptide in a host cell. Such methods include, but are not limited to, for example, methods that utilize either polyclonal or monoclonal antibodies specific for the Cas12i polypeptide or a labeling tag as described elsewhere herein. Exemplary methods include, but are not limited to,
20 enzyme-linked immunosorbent assays (ELISA), radioimmunoassays (MA), fluorescent immunoassays (FIA), and fluorescent activated cell sorting (FACS). These and other assays are well known in the art (See, *e.g.*, Maddox et al., J. Exp. Med. 158:1211 [1983]).

The present disclosure provides methods of *in vivo* expression of the Cas12i polypeptide in a cell, comprising providing a polyribonucleotide encoding the Cas12i polypeptide to a host
25 cell wherein the polyribonucleotide encodes the Cas12i polypeptide, expressing the Cas12i polypeptide in the cell, and obtaining the Cas12i polypeptide from the cell.

The present disclosure further provides methods of *in vivo* expression of a Cas12i polypeptide in a cell, comprising providing a polyribonucleotide encoding the Cas12i polypeptide to a host cell wherein the polyribonucleotide encodes the Cas12i polypeptide and
30 expressing the Cas12i polypeptide in the cell. In some embodiments, the polyribonucleotide encoding the Cas12i polypeptide is delivered to the cell with an RNA guide and, once expressed in the cell, the Cas12i polypeptide and the RNA guide form a complex. In some embodiments,

the polyribonucleotide encoding the Cas12i polypeptide and the RNA guide are delivered to the cell within a single composition. In some embodiments, the polyribonucleotide encoding the Cas12i polypeptide and the RNA guide are comprised within separate compositions. In some embodiments, the host cell is present in a subject, *e.g.*, a human patient.

5

C. Complexes

In some embodiments, an RNA guide targeting HAO1 is complexed with a Cas12i polypeptide to form a ribonucleoprotein. In some embodiments, complexation of the RNA guide and Cas12i polypeptide occurs at a temperature lower than about any one of 20°C, 21°C, 22°C, 23°C, 24°C, 25°C, 26°C, 27°C, 28°C, 29°C, 30°C, 31°C, 32°C, 33°C, 34°C, 35°C, 36°C, 37°C, 38°C, 39°C, 40°C, 41°C, 42°C, 43°C, 44°C, 45°C, 50°C, or 55°C. In some embodiments, the RNA guide does not dissociate from the Cas12i polypeptide at about 37°C over an incubation period of at least about any one of 10mins, 15mins, 20mins, 25mins, 30mins, 35mins, 40mins, 45mins, 50mins, 55mins, 1hr, 2hr, 3hr, 4hr, or more hours.

15 In some embodiments, the RNA guide and Cas12i polypeptide are complexed in a complexation buffer. In some embodiments, the Cas12i polypeptide is stored in a buffer that is replaced with a complexation buffer to form a complex with the RNA guide. In some embodiments, the Cas12i polypeptide is stored in a complexation buffer.

In some embodiments, the complexation buffer has a pH in a range of about 7.3 to 8.6. In one embodiment, the pH of the complexation buffer is about 7.3. In one embodiment, the pH of the complexation buffer is about 7.4. In one embodiment, the pH of the complexation buffer is about 7.5. In one embodiment, the pH of the complexation buffer is about 7.6. In one embodiment, the pH of the complexation buffer is about 7.7. In one embodiment, the pH of the complexation buffer is about 7.8. In one embodiment, the pH of the complexation buffer is about 7.9. In one embodiment, the pH of the complexation buffer is about 8.0. In one embodiment, the pH of the complexation buffer is about 8.1. In one embodiment, the pH of the complexation buffer is about 8.2. In one embodiment, the pH of the complexation buffer is about 8.3. In one embodiment, the pH of the complexation buffer is about 8.4. In one embodiment, the pH of the complexation buffer is about 8.5. In one embodiment, the pH of the complexation buffer is about 8.6.

30 In some embodiments, the Cas12i polypeptide can be overexpressed and complexed with the RNA guide in a host cell prior to purification as described herein. In some embodiments,

mRNA or DNA encoding the Cas12i polypeptide is introduced into a cell so that the Cas12i polypeptide is expressed in the cell. In some embodiments, the RNA guide is also introduced into the cell, whether simultaneously, separately, or sequentially from a single mRNA or DNA construct, such that the ribonucleoprotein complex is formed in the cell.

5

III. Genetic Editing Methods

The disclosure also provides methods of modifying a target site within the HAO1 gene. In some embodiments, the methods comprise introducing an HAO1-targeting RNA guide and a Cas12i polypeptide into a cell. The HAO1-targeting RNA guide and Cas12i polypeptide can be introduced as a ribonucleoprotein complex into a cell. The HAO1-targeting RNA guide and Cas12i polypeptide can be introduced on a nucleic acid vector. The Cas12i polypeptide can be introduced as an mRNA. The RNA guide can be introduced directly into the cell. In some embodiments, the composition described herein is delivered to a cell/tissue/liver/person to reduce HAO1 in the cell/tissue/liver/person. In some embodiments, the composition described herein is delivered to a cell/tissue/liver/person to reduce oxalate production in the cell/tissue/liver/person. In some embodiments, the composition described herein is delivered to a cell/tissue/liver/person to correct calcium oxalate crystal deposition in the cell/tissue/liver/person. In some embodiments, the composition described herein is delivered to a person with primary hyperoxaluria.

Any of the gene editing systems disclosed herein may be used to genetically engineered an HAO1 gene. The gene editing system may comprise an RNA guide and a Cas12i2 polypeptide. The RNA guide comprises a spacer sequence specific to a target sequence in the HAO1 gene, *e.g.*, specific to a region in exon1 or exon 2 of the HAO1 gene.

A. *Target Sequence*

In some embodiments, an RNA guide as disclosed herein is designed to be complementary to a target sequence that is adjacent to a 5'-TTN-3' PAM sequence or 5'-NTTN-3' PAM sequence.

In some embodiments, the target sequence is within an HAO1 gene or a locus of an HAO1 gene (*e.g.*, in exon1 or exon 2), to which the RNA guide can bind *via* base pairing. In some embodiments, a cell has only one copy of the target sequence. In some embodiments, a cell

has more than one copy, such as at least about any one of 2, 3, 4, 5, 10, 100, or more copies of the target sequence.

In some embodiments, the HAO1 gene is a mammalian gene. In some embodiments, the HAO1 gene is a human gene. For example, in some embodiments, the target sequence is within the sequence of SEQ ID NO: 928 (or the reverse complement thereof). In some embodiments, the target sequence is within an exon of the HAO1 gene set forth in SEQ ID NO: 928, *e.g.*, within a sequence of SEQ ID NO: 929, 930, 931, 932, 933, 934, or 935 (or a reverse complement thereof). Target sequences within an exon region of the HAO1 gene of SEQ ID NO: 928 are set forth in **Table 5**. In some embodiments, the target sequence is within an intron of the HAO1 gene set forth in SEQ ID NO: 928 (or the reverse complement thereof). In some embodiments, the target sequence is within a variant (*e.g.*, a polymorphic variant) of the HAO1 gene sequence set forth in SEQ ID NO: 928 (or the reverse complement thereof). In some embodiments, the HAO1 gene sequence is a homolog of the sequence set forth in SEQ ID NO: 928 (or the reverse complement thereof). For examples, in some embodiments, the HAO1 gene sequence is a non-human HAO1 sequence. In some embodiments, the HAO1 gene sequence is a coding sequence set forth in SEQ ID NO: 1024 (or the reverse complement thereof). In some embodiments, the HAO1 gene sequence is a homolog of a coding sequence set forth in SEQ ID NO: 1024 (or the reverse complement thereof).

In some embodiments, the target sequence is adjacent to a 5'-TTN-3' PAM sequence or a 5'-NTTN-3' PAM sequence, wherein N is any nucleotide. The 5'-NTTN-3' sequence may be immediately adjacent to the target sequence or, for example, within a small number (*e.g.*, 1, 2, 3, 4, or 5) of nucleotides of the target sequence. In some embodiments the 5'-NTTN-3' sequence is 5'-NTTY-3', 5'-NTTC-3', 5'-NTTT-3', 5'-NTTA-3', 5'-NTTB-3', 5'-NTTG-3', 5'-CTTY-3', 5'-DTTR-3', 5'-CTTR-3', 5'-DTTT-3', 5'-ATTN-3', or 5'-GTTN-3', wherein Y is C or T, B is any nucleotide except for A, D is any nucleotide except for C, and R is A or G. In some embodiments, the 5'-NTTN-3' sequence is 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3'. The PAM sequence may be 5' to the target sequence.

The 5'-NTTN-3' sequence may be immediately adjacent to the target sequence or, for example, within a small number (*e.g.*, 1, 2, 3, 4, or 5) of nucleotides of the target sequence. In some embodiments the 5'-NTTN-3' sequence is 5'-NTTY-3', 5'-NTTC-3', 5'-NTTT-3', 5'-

NTTA-3', 5'-NTTB-3', 5'-NTTG-3', 5'-CTTY-3', 5'-DTTR-3', 5'-CTTR-3', 5'-DTTT-3', 5'-ATTN-3', or 5'-GTTN-3', wherein Y is C or T, B is any nucleotide except for A, D is any nucleotide except for C, and R is A or G. In some embodiments, the 5'-NTTN-3' sequence is 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3'. In some embodiments, the RNA guide is designed to bind to a first strand of a double-stranded target nucleic acid (*i.e.*, the non-PAM strand), and the 5'-NTTN-3' PAM sequence is present in the second, complementary strand (*i.e.*, the PAM strand). In some embodiments, the RNA guide binds to a region on the non-PAM strand that is complementary to a target sequence on the PAM strand, which is adjacent to a 5'-NAAN-3' sequence.

In some embodiments, the target sequence is present in a cell. In some embodiments, the target sequence is present in the nucleus of the cell. In some embodiments, the target sequence is endogenous to the cell. In some embodiments, the target sequence is a genomic DNA. In some embodiments, the target sequence is a chromosomal DNA. In some embodiments, the target sequence is a protein-coding gene or a functional region thereof, such as a coding region, or a regulatory element, such as a promoter, enhancer, a 5' or 3' untranslated region, etc.

In some embodiments, the target sequence is present in a readily accessible region of the target sequence. In some embodiments, the target sequence is in an exon of a target gene. In some embodiments, the target sequence is across an exon-intron junction of a target gene. In some embodiments, the target sequence is present in a non-coding region, such as a regulatory region of a gene.

B. Gene Editing

In some embodiments, the Cas12i polypeptide has enzymatic activity (*e.g.*, nuclease activity). In some embodiments, the Cas12i polypeptide induces one or more DNA double-stranded breaks in the cell. In some embodiments, the Cas12i polypeptide induces one or more DNA single-stranded breaks in the cell. In some embodiments, the Cas12i polypeptide induces one or more DNA nicks in the cell. In some embodiments, DNA breaks and/or nicks result in formation of one or more indels (*e.g.*, one or more deletions).

In some embodiments, an RNA guide disclosed herein forms a complex with the Cas12i polypeptide and directs the Cas12i polypeptide to a target sequence adjacent to a 5'-NTTN-3' sequence. In some embodiments, the complex induces a deletion (*e.g.*, a nucleotide deletion or

DNA deletion) adjacent to the 5'-NTTN-3' sequence. In some embodiments, the complex induces a deletion adjacent to a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the complex induces a deletion adjacent to a T/C-rich sequence.

In some embodiments, the deletion is downstream of a 5'-NTTN-3' sequence. In some embodiments, the deletion is downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion is downstream of a T/C-rich sequence.

In some embodiments, the deletion alters expression of the HAO1 gene. In some embodiments, the deletion alters function of the HAO1 gene. In some embodiments, the deletion inactivates the HAO1 gene. In some embodiments, the deletion is a frameshifting deletion. In some embodiments, the deletion is a non-frameshifting deletion. In some embodiments, the deletion leads to cell toxicity or cell death (*e.g.*, apoptosis).

In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) downstream of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) downstream of a T/C-rich sequence.

In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a T/C-rich sequence.

In some embodiments, the deletion ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the 5'-NTTN-3' sequence. In some embodiments, the deletion ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of a T/C-rich sequence.

In some embodiments, the deletion ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) downstream of the 5'-NTTN-3' sequence. In some embodiments, the deletion ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-

CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) downstream of a T/C-rich sequence.

In some embodiments, the deletion ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the 5'-NTTN-3' sequence. In some embodiments, the deletion ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of a T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 20 to

about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of the 5'-NTTN-3' sequence and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a T/C-rich sequence and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of the 5'-NTTN-3' sequence and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) downstream of the 5'-NTTN-3' sequence. In some
 5 embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23,
 10 24, 25, 26, 27, or 28 nucleotides) downstream of the 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a T/C-rich sequence and
 15 ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) downstream of the T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of
 20 the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-
 25 CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*,
 30 about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of the 5'-NTTN-3' sequence and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the 5'-NTTN-3' sequence. In some

embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 15 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a T/C-rich sequence and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) downstream of the 5'-NTTN-3' sequence and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence and

ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) downstream of a T/C-rich sequence and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) downstream of the 5'-NTTN-3' sequence and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) downstream of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) downstream of the 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts

within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) downstream of a T/C-rich sequence and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) downstream of the T/C-rich sequence.

5 In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) and ends within about 25 to
10 about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) downstream of the 5'-NTTN-3' sequence and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31,
15 32, or 33 nucleotides) downstream of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence and ends within about 25 to
20 about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 5 to about 10 nucleotides (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 nucleotides)
25 downstream of a T/C-rich sequence and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the T/C-rich sequence.

In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 20 to about 30
30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and

ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some
 5 embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*,
 10 about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of the 5'-NTTN-3' sequence and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-
 15 ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3',
 20 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a T/C-rich sequence and ends within about 20 to about 30 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the T/C-rich sequence.

In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*,
 25 about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about
 30 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-

3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) of a T/C-rich sequence.

5 In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of the 5'-NTTN-3' sequence and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) downstream of the 5'-NTTN-3' sequence. In some
 10 embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides) downstream of the 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3',
 15 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a T/C-rich sequence and ends within about 20 to about 25 nucleotides (*e.g.*, about 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28
 20 nucleotides) downstream of the T/C-rich sequence.

In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of the 5'-NTTN-3' sequence. In some embodiments, the deletion starts within about 10 to about 15
 25 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some embodiments, the deletion starts
 30 within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) of a T/C-rich sequence.

In some embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of the 5'-NTTN-3' sequence and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the 5'-NTTN-3' sequence. In some

5 embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence and ends within about 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31,

10 32, or 33 nucleotides) downstream of the 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3' sequence. In some

embodiments, the deletion starts within about 10 to about 15 nucleotides (*e.g.*, about 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides) downstream of a T/C-rich sequence and ends within about

15 25 to about 30 nucleotides (*e.g.*, about 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, or 33 nucleotides) downstream of the T/C-rich sequence.

In some embodiments, the deletion is up to about 40 nucleotides in length (*e.g.*, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, or 45 nucleotides). In some

20 embodiments, the deletion is between about 4 nucleotides and about 40 nucleotides in length (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, or 45 nucleotides). In some

embodiments, the deletion is between about 4 nucleotides and about 25 nucleotides in length (*e.g.*, about 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27,

25 or 28 nucleotides). In some embodiments, the deletion is between about 10 nucleotides and about 25 nucleotides in length (*e.g.*, about 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 nucleotides). In some embodiments, the deletion is between about 10

nucleotides and about 15 nucleotides in length (*e.g.*, about 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 nucleotides).

30 In some embodiments, the methods described herein are used to engineer a cell comprising a deletion as described herein in an HAO1 gene. In some embodiments, the methods are carried out using a complex comprising a Cas12i enzyme as described herein and an RNA

guide comprising a direct repeat and a spacer as described herein. In some embodiments, the sequence of the RNA guide has at least 90% identity (*e.g.*, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to a sequence of any one of SEQ ID NOs: 967-1023. In some embodiments, an RNA guide has a sequence of any one of SEQ ID NOs: 967-1023.

5 In some embodiments, the RNA guide targeting HAO1 is encoded in a plasmid. In some embodiments, the RNA guide targeting HAO1 is synthetic or purified RNA. In some embodiments, the Cas12i polypeptide is encoded in a plasmid. In some embodiments, the Cas12i polypeptide is encoded by an RNA that is synthetic or purified.

10 *C. Delivery*

Components of any of the gene editing systems disclosed herein may be formulated, for example, including a carrier, such as a carrier and/or a polymeric carrier, *e.g.*, a liposome, and delivered by known methods to a cell (*e.g.*, a prokaryotic, eukaryotic, plant, mammalian, etc.). Such methods include, but not limited to, transfection (*e.g.*, lipid-mediated, cationic polymers, calcium phosphate, dendrimers); electroporation or other methods of membrane disruption (*e.g.*, nucleofection), viral delivery (*e.g.*, lentivirus, retrovirus, adenovirus, adeno-associated virus (AAV)), microinjection, microprojectile bombardment ("gene gun"), fugene, direct sonic loading, cell squeezing, optical transfection, protoplast fusion, impalefection, magnetofection, exosome-mediated transfer, lipid nanoparticle-mediated transfer, and any combination thereof.

20 In some embodiments, the method comprises delivering one or more nucleic acids (*e.g.*, nucleic acids encoding the Cas12i polypeptide, RNA guide, donor DNA, etc.), one or more transcripts thereof, and/or a pre-formed RNA guide/Cas12i polypeptide complex to a cell, where a ternary complex is formed. In some embodiments, an RNA guide and an RNA encoding a Cas12i polypeptide are delivered together in a single composition. In some embodiments, an RNA guide and an RNA encoding a Cas12i polypeptide are delivered in separate compositions. In some embodiments, an RNA guide and an RNA encoding a Cas12i polypeptide delivered in separate compositions are delivered using the same delivery technology. In some embodiments, an RNA guide and an RNA encoding a Cas12i polypeptide delivered in separate compositions are delivered using different delivery technologies. Exemplary intracellular delivery methods, 30 include, but are not limited to: viruses, such as AAV, or virus-like agents; chemical-based transfection methods, such as those using calcium phosphate, dendrimers, liposomes, lipid nanoparticles, or cationic polymers (*e.g.*, DEAE-dextran or polyethylenimine); non-chemical

methods, such as microinjection, electroporation, cell squeezing, sonoporation, optical transfection, impalefection, protoplast fusion, bacterial conjugation, delivery of plasmids or transposons; particle-based methods, such as using a gene gun, magnetofection or magnet assisted transfection, particle bombardment; and hybrid methods, such as nucleofection. In some
5 embodiments, a lipid nanoparticle comprises an mRNA encoding a Cas12i polypeptide, an RNA guide, or an mRNA encoding a Cas12i polypeptide and an RNA guide. In some embodiments, the mRNA encoding the Cas12i polypeptide is a transcript of the nucleotide sequence set forth in SEQ ID NO: 921 or SEQ ID NO: 955 or a variant thereof. In some embodiments, the present application further provides cells produced by such methods, and organisms (such as animals,
10 plants, or fungi) comprising or produced from such cells.

D. Genetically Modified Cells

Any of the gene editing systems disclosed herein can be delivered to a variety of cells. In some embodiments, the cell is an isolated cell. In some embodiments, the cell is in cell culture or
15 a co-culture of two or more cell types. In some embodiments, the cell is *ex vivo*. In some embodiments, the cell is obtained from a living organism and maintained in a cell culture. In some embodiments, the cell is a single-cellular organism.

In some embodiments, the cell is a prokaryotic cell. In some embodiments, the cell is a bacterial cell or derived from a bacterial cell. In some embodiments, the cell is an archaeal cell or
20 derived from an archaeal cell.

In some embodiments, the cell is a eukaryotic cell. In some embodiments, the cell is a plant cell or derived from a plant cell. In some embodiments, the cell is a fungal cell or derived from a fungal cell. In some embodiments, the cell is an animal cell or derived from an animal cell. In some embodiments, the cell is an invertebrate cell or derived from an invertebrate cell. In
25 some embodiments, the cell is a vertebrate cell or derived from a vertebrate cell. In some embodiments, the cell is a mammalian cell or derived from a mammalian cell. In some embodiments, the cell is a human cell. In some embodiments, the cell is a zebra fish cell. In some embodiments, the cell is a rodent cell. In some embodiments, the cell is synthetically made, sometimes termed an artificial cell.

30 In some embodiments, the cell is derived from a cell line. A wide variety of cell lines for tissue culture are known in the art. Examples of cell lines include, but are not limited to, 293T, MF7, K562, HeLa, CHO, and transgenic varieties thereof. Cell lines are available from a variety

of sources known to those with skill in the art (see, *e.g.*, the American Type Culture Collection (ATCC) (Manassas, Va.)). In some embodiments, the cell is an immortal or immortalized cell.

In some embodiments, the cell is a primary cell. In some embodiments, the cell is a stem cell such as a totipotent stem cell (*e.g.*, omnipotent), a pluripotent stem cell, a multipotent stem cell, an oligopotent stem cell, or an unipotent stem cell. In some embodiments, the cell is an induced pluripotent stem cell (iPSC) or derived from an iPSC. In some embodiments, the cell is a differentiated cell. For example, in some embodiments, the differentiated cell is a liver cell (*e.g.*, a hepatocyte), a biliary cell (*e.g.*, a cholangiocyte), a stellate cell, a Kupffer cell, a liver sinusoidal endothelial cell, a muscle cell (*e.g.*, a myocyte), a fat cell (*e.g.*, an adipocyte), a bone cell (*e.g.*, an osteoblast, osteocyte, osteoclast), a blood cell (*e.g.*, a monocyte, a lymphocyte, a neutrophil, an eosinophil, a basophil, a macrophage, a erythrocyte, or a platelet), a nerve cell (*e.g.*, a neuron), an epithelial cell, an immune cell (*e.g.*, a lymphocyte, a neutrophil, a monocyte, or a macrophage), a fibroblast, or a sex cell. In some embodiments, the cell is a terminally differentiated cell. For example, in some embodiments, the terminally differentiated cell is a neuronal cell, an adipocyte, a cardiomyocyte, a skeletal muscle cell, an epidermal cell, or a gut cell. In some embodiments, the cell is an immune cell. In some embodiments, the immune cell is a T cell. In some embodiments, the immune cell is a B cell. In some embodiments, the immune cell is a Natural Killer (NK) cell. In some embodiments, the immune cell is a Tumor Infiltrating Lymphocyte (TIL). In some embodiments, the cell is a mammalian cell, *e.g.*, a human cell or a murine cell. In some embodiments, the murine cell is derived from a wild-type mouse, an immunosuppressed mouse, or a disease-specific mouse model. In some embodiments, the cell is a cell within a living tissue, organ, or organism.

Any of the genetically modified cells produced using any of the gene editing system disclosed herein is also within the scope of the present disclosure. Such modified cells may comprise a disrupted HAO1 gene.

Compositions, vectors, nucleic acids, RNA guides and cells disclosed herein may be used in therapy. Compositions, vectors, nucleic acids, RNA guides and cells disclosed herein may be used in methods of treating a disease or condition in a subject. In some embodiments, the disease or condition is Any suitable delivery or administration method known in the art may be used to deliver compositions, vectors, nucleic acids, RNA guides and cells disclosed herein. Such methods may involve contacting a target sequence with a composition, vector, nucleic acid, or RNA guide disclosed herein. Such methods may involve a method of editing an HAO1 sequence

as disclosed herein. In some embodiments, a cell engineered using an RNA guide disclosed herein is used for *ex vivo* gene therapy.

IV. Therapeutic Applications

5 Any of the gene editing systems or modified cells generated using such a gene editing system as disclosed herein may be used for treating a disease that is associated with the HAO1 gene, for example, primary hyperoxaluria (PH). In some embodiments, the PH is PH1, PH2, or PH3. In specific examples, the target disease is PH1.

10 The gene editing system, pharmaceutical composition or kit comprising such, and any of the RNA guides disclosed herein may be used for treating primary hyperoxaluria (PH) in a subject. PH is a rare genetic disorder effecting subjects of all ages from infants to elderly. PH includes three subtypes involving genetic defects that alter the expression of three distinct proteins. PH1 involves alanine-glyoxylate aminotransferase, or AGT/AGT1. PH2 involves glyoxylate/hydroxypyruvate reductase, or GR/HPR, and PH3 involves 4-hydroxy-2-oxoglutarate
15 aldolase, or HOGA.

In PH1, excess oxalate can also combine with calcium to form calcium oxalate in the kidney and other organs. Deposits of calcium oxalate can produce widespread deposition of calcium oxalate (nephrocalcinosis) or formation of kidney and bladder stones (urolithiasis) and lead to kidney damage. Common kidney complications in PH1 include blood in the urine
20 (hematuria), urinary tract infections, kidney damage, and end-stage renal disease (ESRD). Over time, kidneys in patients with PH1 may begin to fail, and levels of oxalate may rise in the blood. Deposition of oxalate in tissues throughout the body, *e.g.*, systemic oxalosis, may occur due to high blood levels of oxalate and can lead to complications in bone, skin, and eye. Patients with PH1 normally have kidney failure at an early age, with renal dialysis or dual kidney/liver organ
25 transplant as the only treatment options.

In some embodiments, provided herein is a method for treating a target disease as disclosed herein (*e.g.*, PH such as PH1) comprising administering to a subject (*e.g.*, a human patient) in need of the treatment any of the gene editing systems disclosed herein. The gene editing system may be delivered to a specific tissue or specific type of cells where the gene edit
30 is needed. The gene editing system may comprise LNPs encompassing one or more of the components, one or more vectors (*e.g.*, viral vectors) encoding one or more of the components, or a combination thereof. Components of the gene editing system may be formulated to form a

pharmaceutical composition, which may further comprise one or more pharmaceutically acceptable carriers.

In some embodiments, modified cells produced using any of the gene editing systems disclosed herein may be administered to a subject (*e.g.*, a human patient) in need of the treatment. The modified cells may comprise a substitution, insertion, and/or deletion described herein. In some examples, the modified cells may include a cell line modified by a CRISPR nuclease, reverse transcriptase polypeptide, and editing template RNA (*e.g.*, RNA guide and RT donor RNA). In some instances, the modified cells may be a heterogenous population comprising cells with different types of gene edits. Alternatively, the modified cells may comprise a substantially homogenous cell population (*e.g.*, at least 80% of the cells in the whole population) comprising one particular gene edit in the HAO1 gene. In some examples, the cells can be suspended in a suitable media.

In some embodiments, provided herein is a composition comprising the gene editing system or components thereof. Such a composition can be a pharmaceutical composition. A pharmaceutical composition that is useful may be prepared, packaged, or sold in a formulation suitable for oral, rectal, vaginal, parenteral, topical, pulmonary, intranasal, intra-lesional, buccal, ophthalmic, intravenous, intra-organ or another route of administration. A pharmaceutical composition of the disclosure may be prepared, packaged, or sold in bulk, as a single unit dose, or as a plurality of single unit doses. As used herein, a “unit dose” is discrete amount of the pharmaceutical composition (*e.g.*, the gene editing system or components thereof), which would be administered to a subject or a convenient fraction of such a dosage such as, for example, one-half or one-third of such a dosage.

In some embodiments, a pharmaceutical composition comprising the gene editing system or components thereof as described herein may be administered to a subject in need thereof, *e.g.*, one who suffers from a liver disease associated with the HAO1 gene. In some instances, the gene editing system or components thereof may be delivered to specific cells or tissue (*e.g.*, to liver cells), where the gene editing system could function to genetically modify the HAO1 gene in such cells.

A formulation of a pharmaceutical composition suitable for parenteral administration may comprise the active agent (*e.g.*, the gene editing system or components thereof or the modified cells) combined with a pharmaceutically acceptable carrier, such as sterile water or sterile isotonic saline. Such a formulation may be prepared, packaged, or sold in a form suitable

for bolus administration or for continuous administration. Some injectable formulations may be prepared, packaged, or sold in unit dosage form, such as in ampules or in multi-dose containers containing a preservative. Some formulations for parenteral administration include, but are not limited to, suspensions, solutions, emulsions in oily or aqueous vehicles, pastes, and implantable sustained-release or biodegradable formulations. Some formulations may further comprise one or more additional ingredients including, but not limited to, suspending, stabilizing, or dispersing agents.

The pharmaceutical composition may be in the form of a sterile injectable aqueous or oily suspension or solution. This suspension or solution may be formulated according to the known art, and may comprise, in addition to the cells, additional ingredients such as the dispersing agents, wetting agents, or suspending agents described herein. Such sterile injectable formulation may be prepared using a non-toxic parenterally-acceptable diluent or solvent, such as water or saline. Other acceptable diluents and solvents include, but are not limited to, Ringer's solution, isotonic sodium chloride solution, and fixed oils such as synthetic mono- or di-glycerides. Other parentally-administrable formulations which that are useful include those which may comprise the cells in a packaged form, in a liposomal preparation, or as a component of a biodegradable polymer system. Some compositions for sustained release or implantation may comprise pharmaceutically acceptable polymeric or hydrophobic materials such as an emulsion, an ion exchange resin, a sparingly soluble polymer, or a sparingly soluble salt.

V. Kits and Uses Thereof

The present disclosure also provides kits that can be used, for example, to carry out a method described herein for genetical modification of the HAO1 gene. In some embodiments, the kits include an RNA guide and a Cas12i polypeptide. In some embodiments, the kits include a polynucleotide that encodes such a Cas12i polypeptide, and optionally the polynucleotide is comprised within a vector, *e.g.*, as described herein. The Cas12i polypeptide and the RNA guide (*e.g.*, as a ribonucleoprotein) can be packaged within the same or other vessel within a kit or system or can be packaged in separate vials or other vessels, the contents of which can be mixed prior to use. The kits can additionally include, optionally, a buffer and/or instructions for use of the RNA guide and Cas12i polypeptide.

In some embodiments, the kit may be useful for research purposes. For example, in some embodiments, the kit may be useful to study gene function.

All references and publications cited herein are hereby incorporated by reference.

Additional Embodiments

Provided below are additional embodiments, which are also within the scope of the present disclosure.

5 Embodiment 1: A composition comprising an RNA guide, wherein the RNA guide comprises (i) a spacer sequence that is substantially complementary or completely complementary to a region on a non-PAM strand (the complementary sequence of a target sequence) within an HAO1 gene and (ii) a direct repeat sequence; wherein the target sequence is adjacent to a protospacer adjacent motif (PAM) comprising the sequence 5'-NTTN-3'.

10 In Embodiment 1, the target sequence may be within exon 1, exon 2, exon 3, exon 4, exon 5, exon 6, or exon 7 of the HAO1 gene. In some examples, the HAO1 gene comprises the sequence of SEQ ID NO: 928, the reverse complement of SEQ ID NO: 928, a variant of SEQ ID NO: 928, or the reverse complement of a variant of SEQ ID NO: 928.

 In Embodiment 1, the spacer sequence may comprise: (a) nucleotide 1 through nucleotide
15 16 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (b) nucleotide 1 through nucleotide 17 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (c) nucleotide 1 through nucleotide 18 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (d) nucleotide 1 through nucleotide 19 of a sequence that is at least 90% identical to a sequence of
20 any one of SEQ ID NOs: 466-920; (e) nucleotide 1 through nucleotide 20 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (f) nucleotide 1 through nucleotide 21 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (g) nucleotide 1 through nucleotide 22 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (h) nucleotide 1 through nucleotide 23 of a
25 sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (i) nucleotide 1 through nucleotide 24 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (j) nucleotide 1 through nucleotide 25 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (k) nucleotide 1 through nucleotide 26 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID
30 NOs: 466-920; (l) nucleotide 1 through nucleotide 27 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (m) nucleotide 1 through nucleotide 28 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (n)

nucleotide 1 through nucleotide 29 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; or (o) nucleotide 1 through nucleotide 30 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920.

In any of the composition of Embodiment 1, the spacer sequence may comprise: (a) nucleotide 1 through nucleotide 16 of any one of SEQ ID NOs: 466-920; (b) nucleotide 1 through nucleotide 17 of any one of SEQ ID NOs: 466-920; (c) nucleotide 1 through nucleotide 18 of any one of SEQ ID NOs: 466-920; (d) nucleotide 1 through nucleotide 19 of any one of SEQ ID NOs: 466-920; (e) nucleotide 1 through nucleotide 20 of any one of SEQ ID NOs: 466-920; (f) nucleotide 1 through nucleotide 21 of any one of SEQ ID NOs: 466-920; (g) nucleotide 1 through nucleotide 22 of any one of SEQ ID NOs: 466-920; (h) nucleotide 1 through nucleotide 23 of any one of SEQ ID NOs: 466-920; (i) nucleotide 1 through nucleotide 24 of any one of SEQ ID NOs: 466-920; (j) nucleotide 1 through nucleotide 25 of any one of SEQ ID NOs: 466-920; (k) nucleotide 1 through nucleotide 26 of any one of SEQ ID NOs: 466-920; (l) nucleotide 1 through nucleotide 27 of any one of SEQ ID NOs: 466-920; (m) nucleotide 1 through nucleotide 28 of any one of SEQ ID NOs: 466-920; (n) nucleotide 1 through nucleotide 29 of any one of SEQ ID NOs: 466-920; or (o) nucleotide 1 through nucleotide 30 of any one of SEQ ID NOs: 466-920.

In any of the composition of Embodiment 1, the direct repeat sequence may comprises: (a) nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (d) nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (i) nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (k) nucleotide 11 through

nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (n) nucleotide 14
 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (o) nucleotide 1 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (p) nucleotide 2 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (q) nucleotide 3 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (r) nucleotide 4
 10 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (s) nucleotide 5 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (t) nucleotide 6 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (u) nucleotide 7 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (v) nucleotide 8 through
 15 nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (w) nucleotide 9 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (x) nucleotide 10 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (y) nucleotide 11 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (z) nucleotide 12 through
 20 nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; or (aa) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 10 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (b) nucleotide 2 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (c) nucleotide 3 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (d)
 25 nucleotide 4 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (e) nucleotide 5 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (f) nucleotide 6 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (g) nucleotide 7 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (h) nucleotide 8 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (i) nucleotide 9 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (j) nucleotide 10 through nucleotide 36 of any one
 30 of SEQ ID NOs: 1-8; (k) nucleotide 11 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (l) nucleotide 12 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (m) nucleotide 13 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (n) nucleotide 14 through nucleotide 36

of any one of SEQ ID NOs: 1-8; (o) nucleotide 1 through nucleotide 34 of SEQ ID NO: 9; (p) nucleotide 2 through nucleotide 34 of SEQ ID NO: 9; (q) nucleotide 3 through nucleotide 34 of SEQ ID NO: 9; (r) nucleotide 4 through nucleotide 34 of SEQ ID NO: 9; (s) nucleotide 5 through nucleotide 34 of SEQ ID NO: 9; (t) nucleotide 6 through nucleotide 34 of SEQ ID NO: 9; (u) nucleotide 7 through nucleotide 34 of SEQ ID NO: 9; (v) nucleotide 8 through nucleotide 34 of SEQ ID NO: 9; (w) nucleotide 9 through nucleotide 34 of SEQ ID NO: 9; (x) nucleotide 10 through nucleotide 34 of SEQ ID NO: 9; (y) nucleotide 11 through nucleotide 34 of SEQ ID NO: 9; (z) nucleotide 12 through nucleotide 34 of SEQ ID NO: 9; or (aa) SEQ ID NO: 10 or a portion thereof).

10 In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (d) 15 nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (i) 20 nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (k) nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (n) nucleotide 14 through nucleotide 36 of a sequence that is at least 90% 30 identical to a sequence of any one of SEQ ID NOs: 936-953; or (o) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 954 or a portion thereof).

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (b) nucleotide 2 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (c) nucleotide 3 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (d) nucleotide 4 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (e) nucleotide 5 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (f) nucleotide 6 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (g) nucleotide 7 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (h) nucleotide 8 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (i) nucleotide 9 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (j) nucleotide 10 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (k) nucleotide 11 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (l) nucleotide 12 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (m) nucleotide 13 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (n) nucleotide 14 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (or o) SEQ ID NO: 954 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (d) nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (i) nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (k) nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (n) nucleotide 14 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; or (o) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 960 or SEQ ID NO: 961 or a portion thereof).

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of SEQ ID NO: 959; (b) nucleotide 2 through nucleotide 36 of SEQ ID NO: 959; (c) nucleotide 3 through nucleotide 36 of SEQ ID NO: 959; (d) nucleotide 4 through nucleotide 36 of SEQ ID NO: 959; (e) nucleotide 5 through nucleotide 36 of SEQ ID NO: 959; (f) nucleotide 6 through nucleotide 36 of SEQ ID NO: 959; (g) nucleotide 7 through nucleotide 36 of SEQ ID NO: 959; (h) nucleotide 8 through nucleotide 36 of SEQ ID NO: 959; (i) nucleotide 9 through nucleotide 36 of SEQ ID NO: 959; (j) nucleotide 10 through nucleotide 36 of SEQ ID NO: 959; (k) nucleotide 11 through nucleotide 36 of SEQ ID NO: 959; (l) nucleotide 12 through nucleotide 36 of SEQ ID NO: 959; (m) nucleotide 13 through nucleotide 36 of SEQ ID NO: 959; (n) nucleotide 14 through nucleotide 36 of SEQ ID NO: 959; or (o) SEQ ID NO: 960 or SEQ ID NO: 961 or a portion thereof).

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (d) nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (i) nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (k) nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (n) nucleotide 14 through nucleotide 36 of a sequence that is at least 90%

identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (o) nucleotide 15 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; or (p) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 964 or a portion thereof).

5 In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (b) nucleotide 2 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (c) nucleotide 3 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (d) nucleotide 4 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (e) nucleotide 5 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (f) nucleotide 6 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (g) nucleotide 7 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (h) nucleotide 8 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (i) nucleotide 9 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (j) nucleotide 10 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (k) nucleotide 11 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (l) nucleotide 12 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (m) nucleotide 13 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (n) nucleotide 14 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (o) nucleotide 15 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; or (p) SEQ ID NO: 964 or a portion thereof).

20 In some examples, the spacer sequence is substantially complementary or completely complementary to the complement of a sequence of any one of SEQ ID NOs: 11-465.

In any of the composition of Embodiment 1, the PAM may comprise the sequence 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3'.

In some examples, the target sequence is immediately adjacent to the PAM sequence.

In some examples, the RNA guide has a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 967-1023.

30 In some examples, the RNA guide has the sequence of any one of SEQ ID NOs: 967-1023.

Embodiment 2: The composition of Embodiment 1 may further comprise a Cas12i polypeptide or a polyribonucleotide encoding a Cas12i polypeptide, which can be one of the

following: (a) a Cas12i2 polypeptide comprising a sequence that is at least 90% identical to the sequence of SEQ ID NO: 922, SEQ ID NO: 923, SEQ ID NO: 924, SEQ ID NO: 925, SEQ ID NO: 926, or SEQ ID NO: 927; (b) a Cas12i4 polypeptide comprising a sequence that is at least 90% identical to the sequence of SEQ ID NO: 956, SEQ ID NO: 957, or SEQ ID NO: 958; (c) a
5 Cas12i1 polypeptide comprising a sequence that is at least 90% identical to the sequence of SEQ ID NO: 965; or (d) a Cas12i3 polypeptide comprising a sequence that is at least 90% identical to the sequence of SEQ ID NO: 966.

In specific examples, the Cas12i polypeptide is: (a) a Cas12i2 polypeptide comprising a sequence of SEQ ID NO: 922, SEQ ID NO: 923, SEQ ID NO: 924, SEQ ID NO: 925, SEQ ID
10 NO: 926, or SEQ ID NO: 927; (b) a Cas12i4 polypeptide comprising a sequence of SEQ ID NO: 956, SEQ ID NO: 957, or SEQ ID NO: 958; (c) a Cas12i1 polypeptide comprising a sequence of SEQ ID NO: 965; or (d) a Cas12i3 polypeptide comprising a sequence of SEQ ID NO: 966.

In any of the compositions of Embodiment 2, the RNA guide and the Cas12i polypeptide may form a ribonucleoprotein complex. In some examples, the ribonucleoprotein complex binds
15 a target nucleic acid. In some examples, the composition is present within a cell.

In any of the compositions of Embodiment 2, the RNA guide and the Cas12i polypeptide may be encoded in a vector, *e.g.*, expression vector. In some examples, the RNA guide and the Cas12i polypeptide are encoded in a single vector. In other examples, the RNA guide is encoded in a first vector and the Cas12i polypeptide is encoded in a second vector.

20 Embodiment 3: A vector system comprising one or more vectors encoding an RNA guide disclosed herein and a Cas12i polypeptide. In some examples, the vector system comprises a first vector encoding an RNA guide disclosed herein and a second vector encoding a Cas12i polypeptide. The vectors may be expression vectors.

Embodiment 4: A composition comprising an RNA guide and a Cas12i polypeptide,
25 wherein the RNA guide comprises (i) a spacer sequence that is substantially complementary or completely complementary to a region on a non-PAM strand (the complementary sequence of a target sequence) within an HAO1 gene; and (ii) a direct repeat sequence.

In some examples, the target sequence is within exon 1, exon 2, exon 3, exon 4, exon 5, exon 6, or exon 7 of the HAO1 gene, which may comprise the sequence of SEQ ID NO: 928, the
30 reverse complement of SEQ ID NO: 928, a variant of the sequence of SEQ ID NO: 928, or the reverse complement of a variant of SEQ ID NO: 928.

In some examples, the spacer sequence comprises: (a) nucleotide 1 through nucleotide 16 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (b) nucleotide 1 through nucleotide 17 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (c) nucleotide 1 through nucleotide 18 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (d) nucleotide 1 through nucleotide 19 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (e) nucleotide 1 through nucleotide 20 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (f) nucleotide 1 through nucleotide 21 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (g) nucleotide 1 through nucleotide 22 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (h) nucleotide 1 through nucleotide 23 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (i) nucleotide 1 through nucleotide 24 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (j) nucleotide 1 through nucleotide 25 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (k) nucleotide 1 through nucleotide 26 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (l) nucleotide 1 through nucleotide 27 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (m) nucleotide 1 through nucleotide 28 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (n) nucleotide 1 through nucleotide 29 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; or (o) nucleotide 1 through nucleotide 30 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920.

In some examples, the spacer sequence comprises: (a) nucleotide 1 through nucleotide 16 of any one of SEQ ID NOs: 466-920; (b) nucleotide 1 through nucleotide 17 of any one of SEQ ID NOs: 466-920; (c) nucleotide 1 through nucleotide 18 of any one of SEQ ID NOs: 466-920; (d) nucleotide 1 through nucleotide 19 of any one of SEQ ID NOs: 466-920; (e) nucleotide 1 through nucleotide 20 of any one of SEQ ID NOs: 466-920; (f) nucleotide 1 through nucleotide 21 of any one of SEQ ID NOs: 466-920; (g) nucleotide 1 through nucleotide 22 of any one of SEQ ID NOs: 466-920; (h) nucleotide 1 through nucleotide 23 of any one of SEQ ID NOs: 466-920; (i) nucleotide 1 through nucleotide 24 of any one of SEQ ID NOs: 466-920; (j) nucleotide 1 through nucleotide 25 of any one of SEQ ID NOs: 466-920; (k) nucleotide 1 through nucleotide 26 of any one of SEQ ID NOs: 466-920; (l) nucleotide 1 through nucleotide 27 of any one of

SEQ ID NOs: 466-920; (m) nucleotide 1 through nucleotide 28 of any one of SEQ ID NOs: 466-920; (n) nucleotide 1 through nucleotide 29 of any one of SEQ ID NOs: 466-920; or (o) nucleotide 1 through nucleotide 30 of any one of SEQ ID NOs: 466-920.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through
5 nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID
NOs: 1-8; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a
sequence of any one of SEQ ID NOs: 1-8; (c) nucleotide 3 through nucleotide 36 of a sequence
that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (d) nucleotide 4
10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of
SEQ ID NOs: 1-8; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90%
identical to a sequence of any one of SEQ ID NOs: 1-8; (f) nucleotide 6 through nucleotide 36 of
a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (g)
nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of
any one of SEQ ID NOs: 1-8; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least
15 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (i) nucleotide 9 through nucleotide
36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (j)
nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of
any one of SEQ ID NOs: 1-8; (k) nucleotide 11 through nucleotide 36 of a sequence that is at
least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (l) nucleotide 12 through
20 nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID
NOs: 1-8; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to
a sequence of any one of SEQ ID NOs: 1-8; (n) nucleotide 14 through nucleotide 36 of a
sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (o)
nucleotide 1 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of
25 SEQ ID NO: 9; (p) nucleotide 2 through nucleotide 34 of a sequence that is at least 90% identical
to a sequence of SEQ ID NO: 9; (q) nucleotide 3 through nucleotide 34 of a sequence that is at
least 90% identical to a sequence of SEQ ID NO: 9; (r) nucleotide 4 through nucleotide 34 of a
sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (s) nucleotide 5 through
nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (t)
30 nucleotide 6 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of
SEQ ID NO: 9; (u) nucleotide 7 through nucleotide 34 of a sequence that is at least 90% identical
to a sequence of SEQ ID NO: 9; (v) nucleotide 8 through nucleotide 34 of a sequence that is at

least 90% identical to a sequence of SEQ ID NO: 9; (w) nucleotide 9 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (x) nucleotide 10 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (y) nucleotide 11 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of
 5 SEQ ID NO: 9; (z) nucleotide 12 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; or (aa) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 10 or a portion thereof).

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (b) nucleotide 2 through nucleotide 36 of any one
 10 of SEQ ID NOs: 1-8; (c) nucleotide 3 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (d) nucleotide 4 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (e) nucleotide 5 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (f) nucleotide 6 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (g) nucleotide 7 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (h) nucleotide 8 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (i) nucleotide 9 through
 15 nucleotide 36 of any one of SEQ ID NOs: 1-8; (j) nucleotide 10 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (k) nucleotide 11 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (l) nucleotide 12 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (m) nucleotide 13 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (n) nucleotide 14 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (o) nucleotide 1 through nucleotide 34 of SEQ ID NO: 9; (p)
 20 nucleotide 2 through nucleotide 34 of SEQ ID NO: 9; (q) nucleotide 3 through nucleotide 34 of SEQ ID NO: 9; (r) nucleotide 4 through nucleotide 34 of SEQ ID NO: 9; (s) nucleotide 5 through nucleotide 34 of SEQ ID NO: 9; (t) nucleotide 6 through nucleotide 34 of SEQ ID NO: 9; (u) nucleotide 7 through nucleotide 34 of SEQ ID NO: 9; (v) nucleotide 8 through nucleotide 34 of SEQ ID NO: 9; (w) nucleotide 9 through nucleotide 34 of SEQ ID NO: 9; (x) nucleotide 10
 25 through nucleotide 34 of SEQ ID NO: 9; (y) nucleotide 11 through nucleotide 34 of SEQ ID NO: 9; (z) nucleotide 12 through nucleotide 34 of SEQ ID NO: 9; or (aa) SEQ ID NO: 10 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID
 30 NOs: 936-953; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (d)

nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (i) nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (k) nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (n) nucleotide 14 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; or (o) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 954 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (b) nucleotide 2 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (c) nucleotide 3 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (d) nucleotide 4 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (e) nucleotide 5 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (f) nucleotide 6 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (g) nucleotide 7 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (h) nucleotide 8 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (i) nucleotide 9 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (j) nucleotide 10 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (k) nucleotide 11 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (l) nucleotide 12 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (m) nucleotide 13 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (n) nucleotide 14 through nucleotide 36 of any one of SEQ ID NOs: 936-953; or (o) SEQ ID NO: 954 or a portion thereof.

In some embodiments, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (b) nucleotide 2

through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (d) nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (i) nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (k) nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (n) nucleotide 14 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; or (o) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 960 or SEQ ID NO: 961 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of SEQ ID NO: 959; (b) nucleotide 2 through nucleotide 36 of SEQ ID NO: 959; (c) nucleotide 3 through nucleotide 36 of SEQ ID NO: 959; (d) nucleotide 4 through nucleotide 36 of SEQ ID NO: 959; (e) nucleotide 5 through nucleotide 36 of SEQ ID NO: 959; (f) nucleotide 6 through nucleotide 36 of SEQ ID NO: 959; (g) nucleotide 7 through nucleotide 36 of SEQ ID NO: 959; (h) nucleotide 8 through nucleotide 36 of SEQ ID NO: 959; (i) nucleotide 9 through nucleotide 36 of SEQ ID NO: 959; (j) nucleotide 10 through nucleotide 36 of SEQ ID NO: 959; (k) nucleotide 11 through nucleotide 36 of SEQ ID NO: 959; (l) nucleotide 12 through nucleotide 36 of SEQ ID NO: 959; (m) nucleotide 13 through nucleotide 36 of SEQ ID NO: 959; (n) nucleotide 14 through nucleotide 36 of SEQ ID NO: 959; or (o) SEQ ID NO: 960 or SEQ ID NO: 961 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or

SEQ ID NO: 963; (d) nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (i) nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (k) nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (n) nucleotide 14 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (o) nucleotide 15 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; or (p) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 964 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (b) nucleotide 2 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (c) nucleotide 3 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (d) nucleotide 4 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (e) nucleotide 5 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (f) nucleotide 6 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (g) nucleotide 7 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (h) nucleotide 8 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (i) nucleotide 9 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (j) nucleotide 10 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (k) nucleotide 11 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (l) nucleotide 12 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (m) nucleotide 13 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (n) nucleotide 14

through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (o) nucleotide 15 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; or (p) SEQ ID NO: 964 or a portion thereof.

In any of the compositions of Embodiment 4, the spacer sequence may be substantially
5 complementary or completely complementary to the complement of a sequence of any one of SEQ ID NOs: 11-465.

In some examples, the target sequence is adjacent to a protospacer adjacent motif (PAM) comprising the sequence 5'-NTTN-3'. In some examples, the PAM comprises the sequence 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-
10 TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3'.

In some examples, the target sequence is immediately adjacent to the PAM sequence. In some examples, the target sequence is within 1, 2, 3, 4, or 5 nucleotides of the PAM sequence.

In any of the composition of Embodiment 4, the Cas12i polypeptide is: (a) a Cas12i2
15 polypeptide comprising a sequence that is at least 90% identical to the sequence of SEQ ID NO: 922, SEQ ID NO: 923, SEQ ID NO: 924, SEQ ID NO: 925, SEQ ID NO: 926, or SEQ ID NO: 927; (b) a Cas12i4 polypeptide comprising a sequence that is at least 90% identical to the sequence of SEQ ID NO: 956, SEQ ID NO: 957, or SEQ ID NO: 958; (c) a Cas12i1 polypeptide comprising a sequence that is at least 90% identical to the sequence of SEQ ID NO: 965; or (d) a
20 Cas12i3 polypeptide comprising a sequence that is at least 90% identical to the sequence of SEQ ID NO: 966.

In some examples, the Cas12i polypeptide is: (a) a Cas12i2 polypeptide comprising a sequence of SEQ ID NO: 922, SEQ ID NO: 923, SEQ ID NO: 924, SEQ ID NO: 925, SEQ ID NO: 926, or SEQ ID NO: 927; (b) a Cas12i4 polypeptide comprising a sequence of SEQ ID NO:
25 956, SEQ ID NO: 957, or SEQ ID NO: 958; (c) a Cas12i1 polypeptide comprising a sequence of SEQ ID NO: 965; or (d) a Cas12i3 polypeptide comprising a sequence of SEQ ID NO: 966.

In any of the composition of Embodiment 4, the RNA guide and the Cas12i polypeptide may form a ribonucleoprotein complex. In some examples, the ribonucleoprotein complex binds a target nucleic acid.

30 In any of the composition of Embodiment 4, the composition may be present within a cell.

In any of the composition of Embodiment 4, the RNA guide and the Cas12i polypeptide may be encoded in a vector, *e.g.*, expression vector. In some examples, the RNA guide and the Cas12i polypeptide are encoded in a single vector. In other examples, the RNA guide is encoded in a first vector and the Cas12i polypeptide is encoded in a second vector.

5 Embodiment 5: A vector system comprising one or more vectors encoding an RNA guide disclosed herein and a Cas12i polypeptide. In some examples, the vector system comprises a first vector encoding an RNA guide disclosed herein and a second vector encoding a Cas12i polypeptide. In some examples, the vectors are expression vectors.

10 Embodiment 6: An RNA guide comprising (i) a spacer sequence that is substantially complementary or completely complementary to a region on a non-PAM strand (the complementary sequence of a target sequence) within an HAO1 gene, and (ii) a direct repeat sequence.

In some examples, the target sequence is within exon 1, exon 2, exon 3, exon 4, exon 5, exon 6, or exon 7 of the HAO1 gene, which may comprise the sequence of SEQ ID NO: 928, the reverse complement of SEQ ID NO: 928, a variant of the sequence of SEQ ID NO: 928, or the reverse complement of a variant of SEQ ID NO: 928.

In some examples, the spacer sequence comprises: (a) nucleotide 1 through nucleotide 16 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (b) nucleotide 1 through nucleotide 17 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (c) nucleotide 1 through nucleotide 18 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (d) nucleotide 1 through nucleotide 19 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (e) nucleotide 1 through nucleotide 20 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (f) nucleotide 1 through nucleotide 21 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (g) nucleotide 1 through nucleotide 22 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (h) nucleotide 1 through nucleotide 23 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (i) nucleotide 1 through nucleotide 24 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (j) nucleotide 1 through nucleotide 25 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (k) nucleotide 1 through nucleotide 26 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID

NOs: 466-920; (l) nucleotide 1 through nucleotide 27 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (m) nucleotide 1 through nucleotide 28 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; (n) nucleotide 1 through nucleotide 29 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920; or (o) nucleotide 1 through nucleotide 30 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 466-920.

In some examples, the spacer sequence comprises: (a) nucleotide 1 through nucleotide 16 of any one of SEQ ID NOs: 466-920; (b) nucleotide 1 through nucleotide 17 of any one of SEQ ID NOs: 466-920; (c) nucleotide 1 through nucleotide 18 of any one of SEQ ID NOs: 466-920; (d) nucleotide 1 through nucleotide 19 of any one of SEQ ID NOs: 466-920; (e) nucleotide 1 through nucleotide 20 of any one of SEQ ID NOs: 466-920; (f) nucleotide 1 through nucleotide 21 of any one of SEQ ID NOs: 466-920; (g) nucleotide 1 through nucleotide 22 of any one of SEQ ID NOs: 466-920; (h) nucleotide 1 through nucleotide 23 of any one of SEQ ID NOs: 466-920; (i) nucleotide 1 through nucleotide 24 of any one of SEQ ID NOs: 466-920; (j) nucleotide 1 through nucleotide 25 of any one of SEQ ID NOs: 466-920; (k) nucleotide 1 through nucleotide 26 of any one of SEQ ID NOs: 466-920; (l) nucleotide 1 through nucleotide 27 of any one of SEQ ID NOs: 466-920; (m) nucleotide 1 through nucleotide 28 of any one of SEQ ID NOs: 466-920; (n) nucleotide 1 through nucleotide 29 of any one of SEQ ID NOs: 466-920; or (o) nucleotide 1 through nucleotide 30 of any one of SEQ ID NOs: 466-920.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (d) nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (i) nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (j)

nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (k) nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (n) nucleotide 14 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (o) nucleotide 1 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (p) nucleotide 2 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (q) nucleotide 3 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (r) nucleotide 4 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (s) nucleotide 5 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (t) nucleotide 6 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (u) nucleotide 7 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (v) nucleotide 8 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (w) nucleotide 9 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (x) nucleotide 10 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (y) nucleotide 11 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (z) nucleotide 12 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; or (aa) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 10 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (b) nucleotide 2 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (c) nucleotide 3 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (d) nucleotide 4 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (e) nucleotide 5 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (f) nucleotide 6 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (g) nucleotide 7 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (h) nucleotide 8 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (i) nucleotide 9 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (j) nucleotide 10 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (k) nucleotide 11 through nucleotide 36 of any one of SEQ ID NOs: 1-8;

(l) nucleotide 12 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (m) nucleotide 13 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (n) nucleotide 14 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (o) nucleotide 1 through nucleotide 34 of SEQ ID NO: 9; (p) nucleotide 2 through nucleotide 34 of SEQ ID NO: 9; (q) nucleotide 3 through nucleotide 34 of SEQ ID NO: 9; (r) nucleotide 4 through nucleotide 34 of SEQ ID NO: 9; (s) nucleotide 5 through nucleotide 34 of SEQ ID NO: 9; (t) nucleotide 6 through nucleotide 34 of SEQ ID NO: 9; (u) nucleotide 7 through nucleotide 34 of SEQ ID NO: 9; (v) nucleotide 8 through nucleotide 34 of SEQ ID NO: 9; (w) nucleotide 9 through nucleotide 34 of SEQ ID NO: 9; (x) nucleotide 10 through nucleotide 34 of SEQ ID NO: 9; (y) nucleotide 11 through nucleotide 34 of SEQ ID NO: 9; (z) nucleotide 12 through nucleotide 34 of SEQ ID NO: 9; or (aa) SEQ ID NO: 10 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (d) nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (i) nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (k) nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (n) nucleotide 14 through nucleotide 36 of a sequence that is at least 90%

identical to a sequence of any one of SEQ ID NOs: 936-953; or (o) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 954 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (b) nucleotide 2 through nucleotide 36 of
5 any one of SEQ ID NOs: 936-953; (c) nucleotide 3 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (d) nucleotide 4 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (e) nucleotide 5 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (f) nucleotide 6 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (g) nucleotide 7 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (h) nucleotide 8 through nucleotide 36 of any one of SEQ ID
10 NOs: 936-953; (i) nucleotide 9 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (j) nucleotide 10 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (k) nucleotide 11 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (l) nucleotide 12 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (m) nucleotide 13 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (n) nucleotide 14 through nucleotide 36 of any one of SEQ ID NOs: 936-
15 953; or (o) SEQ ID NO: 954 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO:
20 959; (d) nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (h) nucleotide 8 through nucleotide 36 of a sequence that is at
25 least 90% identical to SEQ ID NO: 959; (i) nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (k) nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (n)
30 nucleotide 14 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO:

959; or (o) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 960 or SEQ ID NO: 961 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of SEQ ID NO: 959; (b) nucleotide 2 through nucleotide 36 of SEQ ID NO: 959; 5 (c) nucleotide 3 through nucleotide 36 of SEQ ID NO: 959; (d) nucleotide 4 through nucleotide 36 of SEQ ID NO: 959; (e) nucleotide 5 through nucleotide 36 of SEQ ID NO: 959; (f) nucleotide 6 through nucleotide 36 of SEQ ID NO: 959; (g) nucleotide 7 through nucleotide 36 of SEQ ID NO: 959; (h) nucleotide 8 through nucleotide 36 of SEQ ID NO: 959; (i) nucleotide 9 through nucleotide 36 of SEQ ID NO: 959; (j) nucleotide 10 through nucleotide 36 of SEQ ID 10 NO: 959; (k) nucleotide 11 through nucleotide 36 of SEQ ID NO: 959; (l) nucleotide 12 through nucleotide 36 of SEQ ID NO: 959; (m) nucleotide 13 through nucleotide 36 of SEQ ID NO: 959; (n) nucleotide 14 through nucleotide 36 of SEQ ID NO: 959; or (o) SEQ ID NO: 960 or SEQ ID NO: 961 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through 15 nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (d) nucleotide 4 through nucleotide 36 of a sequence that is at least 90% 20 identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or 25 SEQ ID NO: 963; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (i) nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (k) nucleotide 11 through 30 nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (m) nucleotide 13 through

nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (n) nucleotide 14 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (o) nucleotide 15 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; or (p) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 964 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (b) nucleotide 2 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (c) nucleotide 3 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (d) nucleotide 4 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (e) nucleotide 5 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (f) nucleotide 6 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (g) nucleotide 7 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (h) nucleotide 8 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (i) nucleotide 9 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (j) nucleotide 10 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (k) nucleotide 11 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (l) nucleotide 12 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (m) nucleotide 13 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (n) nucleotide 14 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (o) nucleotide 15 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; or (p) SEQ ID NO: 964 or a portion thereof.

In any of the RNA guide of Embodiment 6, the spacer sequence may be substantially complementary or completely complementary to the complement of a sequence of any one of SEQ ID NOs: 11-465.

In any of the RNA guide of Embodiment 6, the target sequence may be adjacent to a protospacer adjacent motif (PAM) comprising the sequence 5'-NTTN-3', wherein N is any nucleotide. In some examples, the PAM comprises the sequence 5'-ATTA-3', 5'-ATTT-3', 5'-ATTG-3', 5'-ATTC-3', 5'-TTTA-3', 5'-TTTT-3', 5'-TTTG-3', 5'-TTTC-3', 5'-GTTA-3', 5'-GTTT-3', 5'-GTTG-3', 5'-GTTC-3', 5'-CTTA-3', 5'-CTTT-3', 5'-CTTG-3', or 5'-CTTC-3'.

In some examples, the target sequence is immediately adjacent to the PAM sequence. In other examples, the target sequence is within 1, 2, 3, 4, or 5 nucleotides of the PAM sequence.

In some examples, the RNA guide has a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 967-1023. In some specific examples, the RNA guide has the sequence of any one of SEQ ID NOs: 967-1023.

Embodiment 7: A nucleic acid encoding an RNA guide as described herein.

5 Embodiment 8: A vector comprising an RNA guide as described herein.

Embodiment 9: A cell comprising a composition, an RNA guide, a nucleic acid, or a vector as described herein. In some examples, the cell is a eukaryotic cell, an animal cell, a mammalian cell, a human cell, a primary cell, a cell line, a stem cell, or a hepatocyte.

10 Embodiment 10: A kit comprising a composition, an RNA guide, a nucleic acid, or a vector as described herein.

Embodiment 11: A method of editing an HAO1 sequence, the method comprising contacting an HAO1 sequence with a composition or an RNA guide as described herein. In some examples, the method is carried out *in vitro*. In other examples, the method is carried out *ex vivo*.

15 In some examples, the HAO1 sequence is in a cell.

In some examples, the composition or the RNA guide induces a deletion in the HAO1 sequence. In some examples, the deletion is adjacent to a 5'-NTTN-3' sequence, wherein N is any nucleotide. In some specific examples, the deletion is downstream of the 5'-NTTN-3' sequence. In some specific examples, the deletion is up to about 40 nucleotides in length. In
20 some instances, the deletion is from about 4 nucleotides to 40 nucleotides, about 4 nucleotides to 25 nucleotides, about 10 nucleotides to 25 nucleotides, or about 10 nucleotides to 15 nucleotides in length.

In some examples, the deletion starts within about 5 nucleotides to about 15 nucleotides, about 5 nucleotides to about 10 nucleotides, or about 10 nucleotides to about 15 nucleotides of
25 the 5'-NTTN-3' sequence.

In some examples, the deletion starts within about 5 nucleotides to about 15 nucleotides, about 5 nucleotides to about 10 nucleotides, or about 10 nucleotides to about 15 nucleotides downstream of the 5'-NTTN-3' sequence.

In some examples, the deletion ends within about 20 nucleotides to about 30 nucleotides, about 20 nucleotides to about 25 nucleotides, or about 25 nucleotides to about 30 nucleotides of
30 the 5'-NTTN-3' sequence.

In some examples, the deletion ends within about 20 nucleotides to about 30 nucleotides, about 20 nucleotides to about 25 nucleotides, about 25 nucleotides to about 30 nucleotides downstream of the 5'-NTTN-3' sequence.

5 In some examples, the deletion starts within about 5 nucleotides to about 15 nucleotides downstream of the 5'-NTTN-3' sequence and ends within about 20 nucleotides to about 30 nucleotides downstream of the 5'-NTTN-3' sequence.

In some examples, the deletion starts within about 5 nucleotides to about 15 nucleotides downstream of the 5'-NTTN-3' sequence and ends within about 20 nucleotides to about 25 nucleotides downstream of the 5'-NTTN-3' sequence.

10 In some examples, the deletion starts within about 5 nucleotides to about 15 nucleotides downstream of the 5'-NTTN-3' sequence and ends within about 25 nucleotides to about 30 nucleotides downstream of the 5'-NTTN-3' sequence.

In some examples, the deletion starts within about 5 nucleotides to about 10 nucleotides downstream of the 5'-NTTN-3' sequence and ends within about 20 nucleotides to about 30 nucleotides downstream of the 5'-NTTN-3' sequence.

15 In some examples, the deletion starts within about 5 nucleotides to about 10 nucleotides downstream of the 5'-NTTN-3' sequence and ends within about 20 nucleotides to about 25 nucleotides downstream of the 5'-NTTN-3' sequence.

In some examples, the deletion starts within about 5 nucleotides to about 10 nucleotides downstream of the 5'-NTTN-3' sequence and ends within about 25 nucleotides to about 30 nucleotides downstream of the 5'-NTTN-3' sequence.

In some examples, the deletion starts within about 10 nucleotides to about 15 nucleotides downstream of the 5'-NTTN-3' sequence and ends within about 20 nucleotides to about 30 nucleotides downstream of the 5'-NTTN-3' sequence.

25 In some examples, the deletion starts within about 10 nucleotides to about 15 nucleotides downstream of the 5'-NTTN-3' sequence and ends within about 20 nucleotides to about 25 nucleotides downstream of the 5'-NTTN-3' sequence.

In some examples, the deletion starts within about 10 nucleotides to about 15 nucleotides downstream of the 5'-NTTN-3' sequence and ends within about 25 nucleotides to about 30 nucleotides downstream of the 5'-NTTN-3' sequence.

30 In some examples, the 5'-NTTN-3' sequence is 5'-CTTT-3', 5'-CTTC-3', 5'-GTTT-3', 5'-GTTC-3', 5'-TTTC-3', 5'-GTTA-3', or 5'-GTTG-3'.

In some examples, the deletion overlaps with a mutation in the HAO1 sequence. In some instances, the deletion overlaps with an insertion in the HAO1 sequence. In some instances, the deletion removes a repeat expansion of the HAO1 sequence or a portion thereof. In some instances, the deletion disrupts one or both alleles of the HAO1 sequence.

5 In any of the compositions, RNA guides, nucleic acids, vectors, cells, kits, or methods of Embodiments 1-10 described herein, the RNA guide may comprise the sequence of any one of SEQ ID NOs: 967-1023.

Embodiment 12: A method of treating primary hyperoxaluria (PH), which optionally is PH1, PH2, or PH3, in a subject, the method comprising administering any of the compositions,
10 RNAs, or cells as described herein to the subject.

In any of the compositions, RNA guides, cells, kits, or methods described herein, the RNA guide and/or the polyribonucleotide encoding the Cas12i polypeptide may be comprised within a lipid nanoparticle. In some examples, the RNA guide and the polyribonucleotide encoding the Cas12i polypeptide are comprised within the same lipid nanoparticle. In other
15 examples, the RNA guide and the polyribonucleotide encoding the Cas12i polypeptide are comprised within separate lipid nanoparticles.

Embodiment 13: An RNA guide comprising (i) a spacer sequence that is complementary to a target site within an HAO1 gene (the target site being on the non-PAM strand and complementary to a target sequence), and (ii) a direct repeat sequence, wherein the target
20 sequence is any one of SEQ ID NOs: 1047, 1026, or 1025 or the reverse complement thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (c) nucleotide 3 through nucleotide 36 of a sequence
25 that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (d) nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (g)
30 nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (i) nucleotide 9 through nucleotide

36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (k) nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (n) nucleotide 14 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1-8; (o) nucleotide 1 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (p) nucleotide 2 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (q) nucleotide 3 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (r) nucleotide 4 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (s) nucleotide 5 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (t) nucleotide 6 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (u) nucleotide 7 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (v) nucleotide 8 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (w) nucleotide 9 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (x) nucleotide 10 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (y) nucleotide 11 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; (z) nucleotide 12 through nucleotide 34 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 9; or (aa) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 10 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (b) nucleotide 2 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (c) nucleotide 3 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (d) nucleotide 4 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (e) nucleotide 5 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (f) nucleotide 6 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (g) nucleotide 7 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (h) nucleotide 8 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (i) nucleotide 9 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (j) nucleotide 10 through nucleotide 36 of any one

of SEQ ID NOs: 1-8; (k) nucleotide 11 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (l) nucleotide 12 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (m) nucleotide 13 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (n) nucleotide 14 through nucleotide 36 of any one of SEQ ID NOs: 1-8; (o) nucleotide 1 through nucleotide 34 of SEQ ID NO: 9; (p) nucleotide 2 through nucleotide 34 of SEQ ID NO: 9; (q) nucleotide 3 through nucleotide 34 of SEQ ID NO: 9; (r) nucleotide 4 through nucleotide 34 of SEQ ID NO: 9; (s) nucleotide 5 through nucleotide 34 of SEQ ID NO: 9; (t) nucleotide 6 through nucleotide 34 of SEQ ID NO: 9; (u) nucleotide 7 through nucleotide 34 of SEQ ID NO: 9; (v) nucleotide 8 through nucleotide 34 of SEQ ID NO: 9; (w) nucleotide 9 through nucleotide 34 of SEQ ID NO: 9; (x) nucleotide 10 through nucleotide 34 of SEQ ID NO: 9; (y) nucleotide 11 through nucleotide 34 of SEQ ID NO: 9; (z) nucleotide 12 through nucleotide 34 of SEQ ID NO: 9; or (aa) SEQ ID NO: 10 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (d) nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (i) nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (k) nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 936-953; (n) nucleotide 14 through nucleotide 36 of a sequence that is at least 90%

identical to a sequence of any one of SEQ ID NOs: 936-953; or (o) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 954 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (b) nucleotide 2 through nucleotide 36 of
5 any one of SEQ ID NOs: 936-953; (c) nucleotide 3 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (d) nucleotide 4 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (e) nucleotide 5 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (f) nucleotide 6 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (g) nucleotide 7 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (h) nucleotide 8 through nucleotide 36 of any one of SEQ ID
10 NOs: 936-953; (i) nucleotide 9 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (j) nucleotide 10 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (k) nucleotide 11 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (l) nucleotide 12 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (m) nucleotide 13 through nucleotide 36 of any one of SEQ ID NOs: 936-953; (n) nucleotide 14 through nucleotide 36 of any one of SEQ ID NOs: 936-
15 953; or (o) SEQ ID NO: 954 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO:
20 959; (d) nucleotide 4 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (h) nucleotide 8 through nucleotide 36 of a sequence that is at
25 least 90% identical to SEQ ID NO: 959; (i) nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (k) nucleotide 11 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (m) nucleotide 13 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO: 959; (n)
30 nucleotide 14 through nucleotide 36 of a sequence that is at least 90% identical to SEQ ID NO:

959; or (o) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 960 or SEQ ID NO: 961 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of SEQ ID NO: 959; (b) nucleotide 2 through nucleotide 36 of SEQ ID NO: 959; 5 (c) nucleotide 3 through nucleotide 36 of SEQ ID NO: 959; (d) nucleotide 4 through nucleotide 36 of SEQ ID NO: 959; (e) nucleotide 5 through nucleotide 36 of SEQ ID NO: 959; (f) nucleotide 6 through nucleotide 36 of SEQ ID NO: 959; (g) nucleotide 7 through nucleotide 36 of SEQ ID NO: 959; (h) nucleotide 8 through nucleotide 36 of SEQ ID NO: 959; (i) nucleotide 9 through nucleotide 36 of SEQ ID NO: 959; (j) nucleotide 10 through nucleotide 36 of SEQ ID 10 NO: 959; (k) nucleotide 11 through nucleotide 36 of SEQ ID NO: 959; (l) nucleotide 12 through nucleotide 36 of SEQ ID NO: 959; (m) nucleotide 13 through nucleotide 36 of SEQ ID NO: 959; (n) nucleotide 14 through nucleotide 36 of SEQ ID NO: 959; or (o) SEQ ID NO: 960 or SEQ ID NO: 961 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through 15 nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (b) nucleotide 2 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (c) nucleotide 3 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (d) nucleotide 4 through nucleotide 36 of a sequence that is at least 90% 20 identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (e) nucleotide 5 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (f) nucleotide 6 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (g) nucleotide 7 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or 25 SEQ ID NO: 963; (h) nucleotide 8 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (i) nucleotide 9 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (j) nucleotide 10 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (k) nucleotide 11 through 30 nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (l) nucleotide 12 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (m) nucleotide 13 through

nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (n) nucleotide 14 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or SEQ ID NO: 963; (o) nucleotide 15 through nucleotide 36 of a sequence that is at least 90% identical to a sequence of SEQ ID NO: 962 or
 5 SEQ ID NO: 963; or (p) a sequence that is at least 90% identical to a sequence of SEQ ID NO: 964 or a portion thereof.

In some examples, the direct repeat sequence comprises: (a) nucleotide 1 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (b) nucleotide 2 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (c) nucleotide 3 through nucleotide 36 of SEQ ID NO:
 10 962 or SEQ ID NO: 963; (d) nucleotide 4 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (e) nucleotide 5 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (f) nucleotide 6 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (g) nucleotide 7 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (h) nucleotide 8 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (i) nucleotide 9 through nucleotide 36 of
 15 SEQ ID NO: 962 or SEQ ID NO: 963; (j) nucleotide 10 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (k) nucleotide 11 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (l) nucleotide 12 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (m) nucleotide 13 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (n) nucleotide 14 through nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; (o) nucleotide 15 through
 20 nucleotide 36 of SEQ ID NO: 962 or SEQ ID NO: 963; or (p) SEQ ID NO: 964 or a portion thereof.

In some examples, the RNA guide has a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 989, 968, or 967. In some specific examples, the RNA guide has the sequence of any one of SEQ ID NOs: 989, 968, or 967.

25 In some examples, each of the first three nucleotides of the RNA guide comprises a 2'-*O*-methyl phosphorothioate modification.

In some examples, each of the last four nucleotides of the RNA guide comprises a 2'-*O*-methyl phosphorothioate modification.

30 In some examples, each of the first to last, second to last, and third to last nucleotides of the RNA guide comprises a 2'-*O*-methyl phosphorothioate modification, and wherein the last nucleotide of the RNA guide is unmodified.

In some examples, the RNA guide has a sequence that is at least 90% identical to a sequence of any one of SEQ ID NOs: 1082-1087. In some specific examples, the RNA guide has a sequence of any one of SEQ ID NOs: 1082-1087.

In some embodiments, an HAO1-targeting RNA guide comprises at least 90% identity to any one of SEQ ID NOs: 1082-1087. In some embodiments, an HAO1-targeting RNA guide comprises any one of SEQ ID NOs: 1082-1087. In some embodiments, an HAO1-targeting RNA guide comprising at least 90% identity to SEQ ID NO: 1083 or SEQ ID NO: 1084 binds the complementary region of HAO1 target sequence of SEQ ID NO: 1047 *via* base-pairing. In some embodiments, the HAO1-targeting RNA guide of SEQ ID NO: 1083 or SEQ ID NO: 1084 binds the complementary region of HAO1 target sequence of SEQ ID NO: 1047 *via* base-pairing. In some embodiments, an HAO1-targeting RNA guide comprising at least 90% identity to SEQ ID NO: 1085 or SEQ ID NO: 1086 binds the complementary region of HAO1 target sequence of SEQ ID NO: 1026 *via* base-pairing. In some embodiments, the HAO1-targeting RNA guide of SEQ ID NO: 1085 or SEQ ID NO: 1086 binds the complementary region of HAO1 target sequence of SEQ ID NO: 1026 *via* base-pairing. In some embodiments, an HAO1-targeting RNA guide comprising at least 90% identity to SEQ ID NO: 1087 or SEQ ID NO: 2293 binds the complementary region of HAO1 target sequence of SEQ ID NO: 1025 *via* base-pairing. In some embodiments, the HAO1-targeting RNA guide of SEQ ID NO: 1087 or SEQ ID NO: 2293 binds the complementary region of HAO1 target sequence of SEQ ID NO: 1025 *via* base-pairing.

Embodiment 14: A nucleic acid encoding the RNA guide of Embodiment 13 as described herein.

Embodiment 15: A vector comprising the nucleic acid of Embodiment 14 as described herein.

Embodiment 16: A vector system comprising one or more vectors encoding (i) the RNA guide of Embodiment 13 as described herein, and (ii) a Cas12i polypeptide. In some examples, the vector system comprises a first vector encoding the RNA guide and a second vector encoding the Cas12i polypeptide.

Embodiment 17: A cell comprising the RNA guide, the nucleic acid, the vector, or the vector system of Embodiments 13-16 as described herein. In some examples, the cell is a eukaryotic cell, an animal cell, a mammalian cell, a human cell, a primary cell, a cell line, a stem cell, or a T cell.

Embodiment 18: A kit comprising the RNA guide, the nucleic acid, the vector, or the vector system of Embodiments 13-16 as described herein.

Embodiment 19: A method of editing an HAO1 sequence, the method comprising contacting an HAO1 sequence with the RNA guide of Embodiment 13 as described herein. In
5 some examples, the HAO1 sequence is in a cell.

In some examples, the RNA guide induces an indel (*e.g.*, an insertion or deletion) in the HAO1 sequence.

Embodiment 20: A method of treating primary hyperoxaluria (PH), which optionally is PH1, PH2, or PH3, in a subject, the method comprising administering the RNA guide of
10 Embodiment 12 as described herein to the subject.

General techniques

The practice of the present disclosure will employ, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques),
15 microbiology, cell biology, biochemistry, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature, such as *Molecular Cloning: A Laboratory Manual*, second edition (Sambrook, et al., 1989) Cold Spring Harbor Press; *Oligonucleotide Synthesis* (M. J. Gait, ed. 1984); *Methods in Molecular Biology*, Humana Press; *Cell Biology: A Laboratory Notebook* (J. E. Cellis, ed., 1989) Academic Press; Animal
20 Cell Culture (R. I. Freshney, ed. 1987); Introduction to Cell and Tissue Culture (J. P. Mather and P. E. Roberts, 1998) Plenum Press; Cell and Tissue Culture: Laboratory Procedures (A. Doyle, J. B. Griffiths, and D. G. Newell, eds. 1993-8) J. Wiley and Sons; Methods in Enzymology (Academic Press, Inc.); Handbook of Experimental Immunology (D. M. Weir and C. C. Blackwell, eds.); Gene Transfer Vectors for Mammalian Cells (J. M. Miller and M.
25 P. Calos, eds., 1987); Current Protocols in Molecular Biology (F. M. Ausubel, et al. eds. 1987); PCR: The Polymerase Chain Reaction, (Mullis, et al., eds. 1994); Current Protocols in Immunology (J. E. Coligan et al., eds., 1991); Short Protocols in Molecular Biology (Wiley and Sons, 1999); Immunobiology (C. A. Janeway and P. Travers, 1997); Antibodies (P. Finch, 1997); Antibodies: a practice approach (D. Catty., ed., IRL Press, 1988-1989);
30 Monoclonal antibodies: a practical approach (P. Shepherd and C. Dean, eds., Oxford University Press, 2000); Using antibodies: a laboratory manual (E. Harlow and D. Lane (Cold Spring Harbor Laboratory Press, 1999); The Antibodies (M. Zanetti and J. D. Capra, eds.

Harwood Academic Publishers, 1995); *DNA Cloning: A practical Approach*, Volumes I and II (D.N. Glover ed. 1985); *Nucleic Acid Hybridization* (B.D. Hames & S.J. Higgins eds.(1985»; *Transcription and Translation* (B.D. Hames & S.J. Higgins, eds. (1984»; *Animal Cell Culture* (R.I. Freshney, ed. (1986»; *Immobilized Cells and Enzymes* (IRL Press, (1986»; and B. Perbal, *A practical Guide To Molecular Cloning* (1984); F.M. Ausubel *et al.* (eds.).

Without further elaboration, it is believed that one skilled in the art can, based on the above description, utilize the present invention to its fullest extent. The following specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever. All publications cited herein are incorporated by reference for the purposes or subject matter referenced herein.

EXAMPLES

The following examples are provided to further illustrate some embodiments of the present invention but are not intended to limit the scope of the invention; it will be understood by their exemplary nature that other procedures, methodologies, or techniques known to those skilled in the art may alternatively be used.

Example 1 - Cas12i2-Mediated Editing of HAO1 Target Sites in HEK293T Cells

This Example describes the genomic editing of the HAO1 gene using Cas12i2 introduced into HEK293T cells.

Cas12i2 RNA guides (crRNAs) were designed and ordered from Integrated DNA Technologies (IDT). For initial guide screening in HEK293T cells, target sequences were designed by tiling the coding exons of HAO1 for 5'-NTTN-3' PAM sequences, and then spacer sequences were designed for the 20-bp target sequences downstream of the PAM sequence. The HAO1-targeting RNA guide sequences are shown in **Table 7**. In the figures, "E#T#" can also be represented as "exon # target #."

Table 7. crRNA Sequences for HAO1

Guide Name	PAM*	Target strand (Non-PAM Strand)	crRNA Sequence	Target Sequences
HAO1_E1T2	CTTC	BS	AGAAAUCCGUCUUUCAU	CAAAGTCTATATATGA

			UGACGGCAAAGUCUAUA UAUGACUAU (SEQ ID NO: 967)	CTAT (SEQ ID NO: 1025)
HAO1_E1T3	CTTT	TS	AGAAAUCCGUCUUUCAU UGACGGGGAAGUACUGA UUUAGCAUG (SEQ ID NO: 968)	GGAAGTACTGATTTAG CATG (SEQ ID NO: 1026)
HAO1_E1T4	CTTC	TS	AGAAAUCCGUCUUUCAU UGACGGAUCAUUUGCCC CAGACCUGU (SEQ ID NO: 969)	ATCATTGCCCCAGAC CTGT (SEQ ID NO: 1027)
HAO1_E1T5	CTTT	BS	AGAAAUCCGUCUUUCAU UGACGGGGCUGAUAUA UUGCAGCAU (SEQ ID NO: 970)	GGCTGATAATATTGCA GCAT (SEQ ID NO: 1028)
HAO1_E1T6	ATTT	BS	AGAAAUCCGUCUUUCAU UGACGGGUAUCAUAUGAU UAUGAACAA (SEQ ID NO: 971)	GTATCAATGATTATGA ACAA (SEQ ID NO: 1029)
HAO1_E1T7	TTTG	BS	AGAAAUCCGUCUUUCAU UGACGGUAUCAUAUGAU AUGAACAA (SEQ ID NO: 972)	TATCAATGATTATGAA CAAC (SEQ ID NO: 1030)
HAO1_E1T9	ATTA	BS	AGAAAUCCGUCUUUCAU UGACGGUGAACACAUG CUAAUUCAG (SEQ ID NO: 973)	TGAACAACATGCTAAA TCAG (SEQ ID NO: 1031)
HAO1_E1T11	TTTA	TS	AGAAAUCCGUCUUUCAU UGACGGGCAUGUUGUUC AUAUCAUU (SEQ ID NO: 974)	GCATGTTGTTTCATAAT CATT (SEQ ID NO: 1032)
HAO1_E1T12	ATTT	TS	AGAAAUCCGUCUUUCAU UGACGGAGCAUGUUGUU CAUAAUCAU (SEQ ID NO: 975)	AGCATGTTGTTTCATAA TCAT (SEQ ID NO: 1033)
HAO1_E1T13	TTTG	TS	AGAAAUCCGUCUUUCAU UGACGGGGAAGUACUGAU UUAGCAUGU (SEQ ID NO: 976)	GAAGTACTGATTTAGC ATGT (SEQ ID NO: 1034)
HAO1_E1T14	ATTA	BS	AGAAAUCCGUCUUUCAU UGACGGCAGGUCUGGGG CAAAUGAUG (SEQ ID NO: 977)	CAGGTCTGGGGCAAAT GATG (SEQ ID NO: 1035)

HAO1_E1T15	TTTG	TS	AGAAAUCCGUCUUUCAU UGACGGCCCCAGACCUG UAAUAGUCA (SEQ ID NO: 978)	CCCCAGACCTGTAATA GTCA (SEQ ID NO: 1036)
HAO1_E1T16	ATTT	TS	AGAAAUCCGUCUUUCAU UGACGGGCCCCAGACCU GUAAUAGUC (SEQ ID NO: 979)	GCCCCAGACCTGTAAT AGTC (SEQ ID NO: 1037)
HAO1_E1T17	TTTC	TS	AGAAAUCCGUCUUUCAU UGACGGUUCAUCAUUUG CCCCAGACC (SEQ ID NO: 980)	TTCATCATTTGCCCCAG ACC (SEQ ID NO: 1038)
HAO1_E1T18	GTTT	TS	AGAAAUCCGUCUUUCAU UGACGGCUUCAUCAUUU GCCCCAGAC (SEQ ID NO: 981)	CTTCATCATTTGCCCCA GAC (SEQ ID NO: 1039)
HAO1_E1T19	TTTG	BS	AGAAAUCCGUCUUUCAU UGACGGAAUGCUGCAAU AUUAUCAGC (SEQ ID NO: 982)	AATGCTGCAATATTAT CAGC (SEQ ID NO: 1040)
HAO1_E1T23	TTTT	TS	AGAAAUCCGUCUUUCAU UGACGGCUUACCUGGAA AAUGCUGCA (SEQ ID NO: 983)	CTTACCTGGAAAATGC TGCA (SEQ ID NO: 1041)
HAO1_E1T24	ATTT	TS	AGAAAUCCGUCUUUCAU UGACGGUCUUACCUGGA AAAUGCUGC (SEQ ID NO: 984)	TCTTACCTGGAAAATG CTGC (SEQ ID NO: 1042)
HAO1_E2T1	CTTC	BS	AGAAAUCCGUCUUUCAU UGACGGUGUUUUAGGAC AGAGGGUCA (SEQ ID NO: 985)	TGTTTTAGGACAGAGG GTCA (SEQ ID NO: 1043)
HAO1_E2T2	CTTC	TS	AGAAAUCCGUCUUUCAU UGACGGCUCCUACCUCU CACAGUGGC (SEQ ID NO: 986)	CTCCTACCTCTCACAGT GGC (SEQ ID NO: 1044)
HAO1_E2T3	TTTA	BS	AGAAAUCCGUCUUUCAU UGACGGAUUCUAGAUGG AAGCUGUAU (SEQ ID NO: 987)	ATTCTAGATGGAAGCT GTAT (SEQ ID NO: 1045)
HAO1_E2T4	ATTC	BS	AGAAAUCCGUCUUUCAU UGACGGUAGAUGGAAGC UGUAUCCAA (SEQ ID NO: 988)	TAGATGGAAGCTGTAT CCAA (SEQ ID NO: 1046)

			988)	
HAO1_E2T5	ATTC	TS	AGAAAUCCGUCUUUCAU UGACGGCGGAGCAUCCU UGGAUACAG (SEQ ID NO: 989)	CGGAGCATCCTTGGAT ACAG (SEQ ID NO: 1047)
HAO1_E2T6	GTTG	BS	AGAAAUCCGUCUUUCAU UGACGGCUGAAACAGAU CUGUCGACU (SEQ ID NO: 990)	CTGAAACAGATCTGTC GACT (SEQ ID NO: 1048)
HAO1_E2T7	TTTC	TS	AGAAAUCCGUCUUUCAU UGACGGAGCAACAUUCC GGAGCAUCC (SEQ ID NO: 991)	AGCAACATTCCGGAGC ATCC (SEQ ID NO: 1049)
HAO1_E2T8	GTTT	TS	AGAAAUCCGUCUUUCAU UGACGGCAGCAACAUUC CGGAGCAUC (SEQ ID NO: 992)	CAGCAACATTCCGGAG CATC (SEQ ID NO: 1050)
HAO1_E2T9	GTTT	BS	AGAAAUCCGUCUUUCAU UGACGGUAGGACAGAGG GUCAGCAUG (SEQ ID NO: 993)	TAGGACAGAGGGTCAG CATG (SEQ ID NO: 1051)
HAO1_E2T10	TTTT	BS	AGAAAUCCGUCUUUCAU UGACGGAGGACAGAGGG UCAGCAUGC (SEQ ID NO: 994)	AGGACAGAGGGTCAGC ATGC (SEQ ID NO: 1052)
HAO1_E2T11	TTTA	BS	AGAAAUCCGUCUUUCAU UGACGGGGACAGAGGGU CAGCAUGCC (SEQ ID NO: 995)	GGACAGAGGGTCAGCA TGCC (SEQ ID NO: 1053)
HAO1_E3T1	CTTT	BS	AGAAAUCCGUCUUUCAU UGACGGCUUUCUCAGCC UGUCAGUCC (SEQ ID NO: 996)	CTTTCTCAGCCTGTCAG TCC (SEQ ID NO: 1054)
HAO1_E3T2	CTTT	BS	AGAAAUCCGUCUUUCAU UGACGGCUCAGCCUGUC AGUCCCUGG (SEQ ID NO: 997)	CTCAGCCTGTCAGTCC CTGG (SEQ ID NO: 1055)
HAO1_E3T3	GTTC	TS	AGAAAUCCGUCUUUCAU UGACGGCCAGGGACUGA CAGGCUGAG (SEQ ID NO: 998)	CCAGGGACTGACAGGC TGAG (SEQ ID NO: 1056)
HAO1_E3T4	GTTC	BS	AGAAAUCCGUCUUUCAU	CTGGGCCACCTCCTCA

			UGACGGCUGGGCCACCU CCUCAAUUG (SEQ ID NO: 999)	ATTG (SEQ ID NO: 1057)
HAO1_E3T5	CTTC	TS	AGAAAUCCGUCUUUCAU UGACGGAAUUGAGGAGG UGGCCAGG (SEQ ID NO: 1000)	AATTGAGGAGGTGGCC CAGG (SEQ ID NO: 1058)
HAO1_E3T6	CTTC	TS	AGAAAUCCGUCUUUCAU UGACGGUUCAAUUGAGG AGGUGGCC (SEQ ID NO: 1001)	TTCAATTGAGGAGGTG GCCC (SEQ ID NO: 1059)
HAO1_E3T7	CTTC	TS	AGAAAUCCGUCUUUCAU UGACGGCGCCACUUCUU CAAUUGAGG (SEQ ID NO: 1002)	CGCCACTTCTTCAATTG AGG (SEQ ID NO: 1060)
HAO1_E3T8	CTTC	BS	AGAAAUCCGUCUUUCAU UGACGGGUUGGCUGCAA CUGUAUAUC (SEQ ID NO: 1003)	GTTGGCTGCAACTGTA TATC (SEQ ID NO: 1061)
HAO1_E3T9	CTTC	TS	AGAAAUCCGUCUUUCAU UGACGGUCGGUCCUUGU AGAUUAACA (SEQ ID NO: 1004)	TCGGTCCTTGTAGATA TACA (SEQ ID NO: 1062)
HAO1_E3T11	CTTC	TS	AGAAAUCCGUCUUUCAU UGACGGUCUGCCUGCCG CACUAGCUU (SEQ ID NO: 1005)	TCTGCCTGCCGCACTA GCTT (SEQ ID NO: 1063)
HAO1_E3T12	TTTC	BS	AGAAAUCCGUCUUUCAU UGACGGUUCUCAGCCU GUCAGUCCC (SEQ ID NO: 1006)	TTTCTCAGCCTGTCAGT CCC (SEQ ID NO: 1064)
HAO1_E3T13	TTTC	BS	AGAAAUCCGUCUUUCAU UGACGGUCAGCCUGUCA GUCCUGGG (SEQ ID NO: 1007)	TCAGCCTGTCAGTCCC TGGG (SEQ ID NO: 1065)
HAO1_E3T14	GTTG	BS	AGAAAUCCGUCUUUCAU UGACGGAGUCCUGGGC CACCUCUC (SEQ ID NO: 1008)	AGTTCCTGGGCCACCT CCTC (SEQ ID NO: 1066)
HAO1_E3T15	ATTG	TS	AGAAAUCCGUCUUUCAU UGACGGAGGAGGUGGCC CAGGAACUC (SEQ ID NO: 1009)	AGGAGGTGGCCCAGGA ACTC (SEQ ID NO: 1067)

HAO1_E3T16	ATTG	BS	AGAAAUCCGUCUUUCAU UGACGGAAGAAGUGGCG GAAGCUGGU (SEQ ID NO: 1010)	AAGAAGTGGCGGAAG CTGGT (SEQ ID NO: 1068)
HAO1_E3T17	GTTG	BS	AGAAAUCCGUCUUUCAU UGACGGGUGCAACUGU AUAUCUACA (SEQ ID NO: 1011)	GCTGCAACTGTATATC TACA (SEQ ID NO: 1069)
HAO1_E3T18	GTTG	TS	AGAAAUCCGUCUUUCAU UGACGGCAGCCAACGAA GUGCCUCAG (SEQ ID NO: 1012)	CAGCCAACGAAGTGCC TCAG (SEQ ID NO: 1070)
HAO1_E3T19	CTTG	TS	AGAAAUCCGUCUUUCAU UGACGGUAGAUAUACAG UUGCAGCCA (SEQ ID NO: 1013)	TAGATATACAGTTGCA GCCA (SEQ ID NO: 1071)
HAO1_E3T20	CTTG	TS	AGAAAUCCGUCUUUCAU UGACGGGUGACUUCUCG GUCCUUGUA (SEQ ID NO: 1014)	GTGACTTCTCGGTCCTT GTA (SEQ ID NO: 1072)
HAO1_E3T22	ATTT	BS	AGAAAUCCGUCUUUCAU UGACGGGUGACAGUGGA CACACCUUA (SEQ ID NO: 1015)	GTGACAGTGGACACAC CTTA (SEQ ID NO: 1073)
HAO1_E3T23	TTTG	BS	AGAAAUCCGUCUUUCAU UGACGGUGACAGUGGAC ACACCUUAC (SEQ ID NO: 1016)	TGACAGTGGACACACC TTAC (SEQ ID NO: 1074)
HAO1_E3T24	CTTA	BS	AGAAAUCCGUCUUUCAU UGACGGCCUGGGCAACC GUCUGGAUG (SEQ ID NO: 1017)	CCTGGGCAACCGTCTG GATG (SEQ ID NO: 1075)
HAO1_E3T25	GTTG	TS	AGAAAUCCGUCUUUCAU UGACGGCCCAGGUAAGG UGUGUCCAC (SEQ ID NO: 1018)	CCCAGGTAAGGTGTGT CCAC (SEQ ID NO: 1076)
HAO1_E3T26	GTTA	TS	AGAAAUCCGUCUUUCAU UGACGGCGCACAUCAUC CAGACGGUU (SEQ ID NO: 1019)	CGCACATCATCCAGAC GGTT (SEQ ID NO: 1077)
HAO1_E3T27	TTTG	TS	AGAAAUCCGUCUUUCAU UGACGGAAUCUGUUACG CACAUCAUC (SEQ ID NO: 1020)	AATCTGTTACGCACAT CATC (SEQ ID NO: 1078)

			1020)	
HAO1_E3T28	GTTT	TS	AGAAAUCCGUCUUUCAU UGACGGGAAUCUGUUAC GCACAUCAU (SEQ ID NO: 1021)	GAATCTGTTACGCACA TCAT (SEQ ID NO: 1079)
HAO1_E3T29	GTTG	TS	AGAAAUCCGUCUUUCAU UGACGGUGGCGGCAGUU UGAAUCUGU (SEQ ID NO: 1022)	TGGCGGCAGTTTGAAT CTGT (SEQ ID NO: 1080)
HAO1_E3T30	GTTA	TS	AGAAAUCCGUCUUUCAU UGACGGCCUGAGUUGUG GCGGCAGUU (SEQ ID NO: 1023)	CCTGAGTTGTGGCGGC AGTT (SEQ ID NO: 1081)

* The 3' three nucleotides represent the 5'-TTN-3' motif.

Cas12i2 RNP complexation reactions were made by mixing purified Cas12i2 polypeptide of SEQ ID NO: 924 (400 μ M) with an HAO1-targeting crRNA (1 mM in 250 mM NaCl) at a 1:1 (Cas12i2:crRNA) volume ratio (2.5:1 crRNA:Cas12i2 molar ratio). Complexations were incubated on ice for 30-60 min.

HEK293T cells were harvested using TRYPLE™ (recombinant cell-dissociation enzymes; ThermoFisher) and counted. Cells were washed once with PBS and resuspended in SF buffer + supplement (SF CELL LINE 4D-NUCLEOFECTOR™ X KIT S; Lonza #V4XC-2032) at a concentration of 16,480 cells/ μ L. Resuspended cells were dispensed at 3×10^5 cells/reaction into Lonza 16-well NUCLEOCUVETTE® strips. Complexed Cas12i2 RNP was added to each reaction at a final concentration of 10 μ M (Cas12i2), and transfection enhancer oligos were then added at a final concentration of 4 μ M. The final volume of each electroporated reaction was 20 μ L. Non-targeting guides were used as negative controls.

The strips were electroporated using an electroporation device (program CM-130, Lonza 4D-NUCLEOFECTOR™). Immediately following electroporation, 80 μ L of pre-warmed DMEM + 10% FBS was added to each well and mixed gently by pipetting. For each technical replicate plate, plated 10 μ L (30,000 cells) of diluted nucleofected cells into pre-warmed 96-well plate with wells containing 100 μ L DMEM + 10% FBS. Editing plates were incubated for 3 days at 37°C with 5% CO₂.

After 3 days, wells were harvested using TRYPLE™ (recombinant cell-dissociation enzymes; ThermoFisher) and transferred to 96-well TWIN.TEC® PCR plates (Eppendorf).

Media was flicked off and cells were resuspended in 20 μ L QUICKEXTRACT™ (DNA extraction buffer; Lucigen). Samples were then cycled in a PCR machine at 65°C for 15 min, 68°C for 15 min, 98°C for 10 min. Samples were then frozen at -20°C.

Samples for Next Generation Sequencing (NGS) were prepared by rounds of PCR. The first round (PCR I) was used to amplify the genomic regions flanking the target site and add NGS adapters. The second round (PCR II) was used to add NGS indexes. Reactions were then pooled, purified by column purification, and quantified on a fluorometer (Qubit). Sequencing runs were done using a 150 cycle NGS instrument (NEXTSEQ™ v2.5) mid or high output kit (Illumina) and run on an NGS instrument (NEXTSEQ™ 550; Illumina).

For NGS analysis, the indel mapping function used a sample's fastq file, the amplicon reference sequence, and the forward primer sequence. For each read, a kmer-scanning algorithm was used to calculate the edit operations (match, mismatch, insertion, deletion) between the read and the reference sequence. In order to remove small amounts of primer dimer present in some samples, the first 30 nt of each read was required to match the reference and reads where over half of the mapping nucleotides are mismatches were filtered out as well. Up to 50,000 reads passing those filters were used for analysis, and reads were counted as an indel read if they contained an insertion or deletion. The % indels was calculated as the number of indel-containing reads divided by the number of reads analyzed (reads passing filters up to 50,000). The QC standard for the minimum number of reads passing filters was 10,000.

FIG. 1 shows HAO1 indels in HEK293T cells following RNP delivery. Error bars represent the average of three technical replicates across one biological replicate. Following delivery, indels were detected within and/or adjacent to each of the HAO1 target sites with each of the RNA guides. Delivery of E1T2, E1T3, E1T6, E1T7, E1T13, T1T17, E2T4, E2T5, E2T9, E2T10, E3T6, E3T19, E3T22, and E3T28 resulted in indels in over 70% of the NGS reads. Therefore, HAO1-targeting RNA guides induced indels in exon 1, exon 2, and exon 3 in HEK293T cells.

This Example thus shows that HAO1 can be individually targeted by Cas12i2 RNPs in mammalian cells such as HEK293T cells.

Example 2 - Cas12i2-Mediated Editing of HAO1 Target Sites in HepG2 cells

This Example describes the genomic editing of the HAO1 gene using Cas12i2 introduced into HepG2 cells by RNP.

RNP complexation reactions were performed as described in Example 1 with various RNA guides of **Table 7**. HepG2 cells were harvested using TRYPLE™ (recombinant cell-dissociation enzymes; ThermoFisher) and counted. Cells were washed once with PBS and resuspended in SF buffer + supplement (SF CELL LINE 4D-NUCLEOFECTOR™ X KIT S; Lonza #V4XC-2032) at a concentration of 13,889 cells/μL. Resuspended cells were dispensed at 2.5e5 cells/reaction into Lonza 16-well NUCLEOCUVETTE® strips. Complexed Cas12i2 RNP was added to each reaction at a final concentration of 20 μM (Cas12i2), with no transfection enhancer oligo. The final volume of each electroporated reaction was 20 μL. Non-targeting guides were used as negative controls.

The strips were electroporated using an electroporation device (program DJ-100, Lonza 4D-NUCLEOFECTOR™). Immediately following electroporation, 80 μL of pre-warmed EMEM + 10% FBS was added to each well and mixed gently by pipetting. For each technical replicate plate, plated 10 μL (25,000 cells) of diluted nucleofected cells into pre-warmed 96-well plate with wells containing 100 μL EMEM + 10% FBS. Editing plates were incubated for 3 days at 37°C with 5% CO₂.

After 3 days, wells were harvested using TRYPLE™ (recombinant cell-dissociation enzymes; ThermoFisher) and transferred to 96-well TWIN.TEC® PCR plates (Eppendorf). Media was flicked off and cells were resuspended in 20 μL QUICKEXTRACT™ (DNA extraction buffer; Lucigen). Samples were then cycled in a PCR machine at 65°C for 15 min, 68°C for 15 min, 98°C for 10 min. Samples were then frozen at -20°C. Samples were analyzed by NGS as described in Example 1.

FIG. 2 shows HAO1 indels in HepG2 cells following RNP delivery. Error bars represent the average of three technical replicates across one biological replicate. Following delivery, indels were detected within and/or adjacent to each of the HAO1 target sites with each of the RNA guides. Therefore, HAO1-targeting RNA guides induced indels in exon 1, exon 2, and exon 3 in HepG2 cells.

This Example thus shows that HAO1 can be targeted by Cas12i2 RNPs in mammalian cells such as HepG2 cells.

Example 3 - Cas12i2-Mediated Editing of HAO1 Target Sites in Primary Hepatocytes

This Example describes the genomic editing of the HAO1 using Cas12i2 introduced into primary hepatocytes cells by RNP.

RNP complexation reactions were performed as described in Example 1 with RNA guides of **Table 7**. Primary hepatocyte cells from human donors were thawed from liquid nitrogen very quickly in a 37°C water bath. The cells were added to pre-warmed hepatocyte recovery media (Thermofisher, CM7000) and centrifuged at 100g for 10 minutes. The cell pellet was resuspended in appropriate volume of hepatocyte plating Medium (Williams' Medium E, Thermofisher A1217601 supplemented with Hepatocyte Plating Supplement Pack (serum-containing), Thermofisher CM3000). The cells were subjected to trypan blue viability count with an INCUCYTE® disposable hemocytometer (Fisher scientific, 22-600-100). The cells were then washed in PBS and resuspended in P3 buffer + supplement (P3 PRIMARY CELL 4D-NUCLEOFECTOR™ X Kit; Lonza, VXP-3032) at a concentration of ~7,500 cells/μL. Resuspended cells were dispensed at 150,000 cells/reaction into the 16 well Lonza NUCLEOCUVETTE strips or 500,000 cells/reaction into the single Lonza NUCLEOCUVETTES® for the mRNA readout. Complexed Cas12i2 RNP was added to each reaction at a final concentration of 20 μM (Cas12i2), and transfection enhancer oligos were then added at a final concentration of 4 μM. The final volume of each electroporated reaction was either 20 μL in the 16 well nucleocuvette strip format or 100 μL in the single nucleocuvette format. Non-targeting guides were used as negative controls.

The strips were electroporated using DS-150 program, while the single nucleocuvettes were electroporated using CA137 program (Lonza 4D-NUCLEOFECTOR™). Immediately following electroporation, pre-warmed Hepatocyte plating medium was added to each well and mixed very gently by pipetting. For each technical replicate plate, plated all the cell suspension of diluted nucleofected cells into a pre-warmed collagen-coated 96-well plate or 24-well plate (Thermofisher) with wells containing Hepatocyte plating medium. The cells were then incubated in a 37°C incubator. The media was changed to hepatocyte maintenance media (Williams' Medium E, Thermofisher A1217601 supplemented with William's E medium Cell Maintenance Cocktail, Thermofisher CM 4000) after the cells attached after 4 hours. Fresh hepatocyte maintenance media was replaced after 2 days.

After 4-5 days post RNP electroporation, media was aspirated and the cells were harvested by shaking (500 rpm) in a 37°C incubator with 2mg/ml collagenase IV (Thermofisher, 17104019) dissolved in PBS containing Ca/Mg (Thermofisher). After cells were dissociated from the plate, they were transferred to 96-well TWIN.TEC® PCR plates (Eppendorf) and centrifuged. Media was flicked off and cell pellets for the NGS readout were resuspended in 20

μL QUICKEXTRACT™ (DNA extraction buffer; Lucigen). Samples were then cycled in a PCR machine at 65°C for 15 min, 68°C for 15 min, 98°C for 10 min and analyzed by NGS as described in Example 1.

For the mRNA readout, cell pellets were frozen at -80°C and subsequently resuspended in lysis buffer and DNA/RNA extracted with the RNeasy kit (Qiagen) following manufacturer's instructions. The DNA extracted from the samples were analyzed by NGS. The RNA isolated was checked for quantity and purity using nanodrop, and subsequently used for cDNA synthesis using 5x iScript reverse transcription reaction mix (Bio-Rad laboratories), following manufacturer's recommendations. cDNA templated was appropriately diluted to be in linear range of the subsequent analysis. Diluted cDNA was used to set up a 20 μL Digital Droplet PCR (ddPCR- BioRad laboratories) reaction using target-specific primer and probe for HAO1, ATTGTGCACTGTCAGATCTTGGAAACGGCCAAAGGATTTTCCTCACCAATGTCTTG TCGATGACTTTTACATTCTGGCACCCACTCAGAGCCATGGCCAACCGGAATTCTTCC TTTAGTAT (SEQ ID NO: 1088), and 2x ddPCR Supermix for Probes No dUTP (BioRad laboratories) following manufacturer's instructions. The reaction was used to generate droplets using Automated Droplet Generator (BioRad Laboratories), following manufacture's recommendations. The plate was sealed using PX1 PCR Plate Sealer (BioRad Laboratories) generated droplets were subjected to PCR amplification using C1000 Touch Thermal Cycler (BioRad Laboratories) using conditions recommended by the manufacturer. The PCR amplified droplets were read on QX200 Droplet Reader (BioRad Laboratories) and the acquired data was analyzed using QX Manager version 1.2 (BioRad Laboratories) to determine presence of absolute copy number of mRNA present in each reaction for the appropriate targets.

As shown in **FIG. 3**, each RNA guide tested induced indels within and/or adjacent to the HAO1 target sites. Indels were not induced with the non-targeting control. Therefore, HAO1-targeting RNA guides induced indels in primary hepatocytes. Indels were then correlated with mRNA levels for each target to determine whether indels lead to mRNA knockdown and subsequent protein knockdown. **FIG. 4** shows %mRNA knockdown of HAO1 in edited cells compared to unedited control cells. Although a higher percentage of NGS reads comprised indels using HAO1 E2T5 (SEQ ID NO: 989) compared to HAO1 E2T4 (SEQ ID NO: 988), HAO1 E2T4 resulted in a greater knockdown of HAO1 mRNA.

This Example thus shows that HAO1 can be targeted by Cas12i2 RNPs in mammalian cells such as primary human hepatocytes.

Example 4 - Editing of HAO1 Target Sites in HepG2 Cells with Cas12i2 Variants

This Example describes indel assessment on HAO1 targets using variants introduced into HepG2 cells by transient transfection.

5 The Cas12i2 variants of SEQ ID NO: 924 and SEQ ID NO: 927 were individually cloned into a pcda3.1 backbone (Invitrogen). Nucleic acids encoding RNA guides E1T2 (SEQ ID NO: 967), E1T3 (SEQ ID NO: 968), E2T4 (SEQ ID NO: 988), E2T5 (SEQ ID NO: 989), E2T10 (SEQ ID NO: 994) were cloned into a pUC19 backbone (New England Biolabs). The plasmids were then maxi-prepped and diluted.

10 HepG2 cells were harvested using TRYPLE™ (recombinant cell-dissociation enzymes; ThermoFisher) and counted. Cells were washed once with PBS and resuspended in SF buffer + supplement (SF CELL LINE 4D-NUCLEOFECTOR™ X KIT S; Lonza #V4XC-2032).

Approximately 16 hours prior to transfection, 25,000 HepG2 cells in EMEM/10%FBS were plated into each well of a 96-well plate. On the day of transfection, the cells were 70-90%
15 confluent. For each well to be transfected, a mixture of Lipofectamine™ 3000 and Opti-MEM® was prepared and then incubated at room temperature for 5 minutes (Solution 1). After incubation, the lipofectamine™:OptiMEM® mixture was added to a separate mixture containing nuclease plasmid and RNA guide plasmid and P3000 reagent (Solution 2). In the case of negative controls, the crRNA was not included in Solution 2. The Solution 1 and Solution 2 were
20 mixed by pipetting up and down and then incubated at room temperature for 15 minutes. Following incubation, the Solution 1 and Solution 2 mixture was added dropwise to each well of a 96 well plate containing the cells.

After 3 days, wells were harvested using TRYPLE™ (recombinant cell-dissociation enzymes; ThermoFisher) and transferred to 96-well TWIN.TEC® PCR plates (Eppendorf).

25 Media was flicked off and cells were resuspended in 20 µL QUICKEXTRACT™ (DNA extraction buffer; Lucigen). Samples were then cycled in a PCR machine at 65°C for 15 min, 68°C for 15 min, 98°C for 10 min. Samples were then frozen at -20°C and analyzed by NGS as described in Example 1.

As shown in **FIG. 5A**, comparable indel activity with the two Cas12i2 variants was
30 observed for E1T2, E1T3, E2T4, E2T5, E2T10. **FIG. 5B** shows the indel size frequency (left) and indel start position relative to the PAM for E1T3 and the variant Cas12i2 of SEQ ID NO: 924. As shown on the left, deletions ranged in size from 1 nucleotide to about 40 nucleotides.

The majority of the deletions were about 6 nucleotides to about 27 nucleotides in length. As shown on the right, the target sequence is represented as starting at position 0 and ending at position 20. Indels started within about 10 nucleotides and about 35 nucleotides downstream of the PAM sequence. The majority of indels started near the end of the target sequence, *e.g.*, about 18 nucleotides to about 25 nucleotides downstream of the PAM sequence.

Thus, this Example shows that HAO1 is capable of being targeted by multiple Cas12i2 polypeptides.

Example 5 - Editing of HAO1 in Primary Human Hepatocytes Using Cas12i2 mRNA Constructs

This Example describes indel assessment on HAO1 target sites *via* delivery of Cas12i2 mRNA and chemically modified HAO1-targeting RNA guides.

mRNA sequences corresponding to the variant Cas12i2 sequence of SEQ ID NO: 924 and the variant Cas12i2 sequence of SEQ ID NO: 927 were synthesized by Aldeveron with 1-pseudo-U modified nucleotides and using CleanCap® Reagent AG (TriLink Biotechnologies).

The Cas12i2 mRNA sequences, shown in **Table 8**, further comprised a C-terminal NLS.

Table 8. Cas12i2 mRNA Sequences

Description	mRNA Sequence
mRNA corresponding to variant Cas12i2 of SEQ ID NO: 924	AUGAGCUCCGCCAUAAGUCCUACAAGUCUGUGCUGCGGCCAAACGAGAGAAAGAAUCAGC UGCUGAAGUCCACCAUCCAGUGCCUGGAGGACGGCUCGCGCCUUCUUUUAAGAUUGCUGCA GGGCCUGUUUGGCGGCAUACCCCCGAGAUUCGUGAGAUUCAGCACAGAGCAGGAGAAGCAG CAGCAGGAUAUCGCCCUGUGGUGUGCCGUGAAUUGGUUCAGGCCUGUGAGCCAGGACUCCC UGACCCACACAAUCGCCUCCGAUAACUGGUGGAGAAGUUUGAGGAGUACUAUUGCGGGCAC AGCCAGCGAGCCCAUAAGCAGUACUUCAGCGCCUCCAUUGCGGAGUCCUACUAUUGGAU GACUGCCGCCAGCAGUACUAUGAUCUGUGCGGGAGCUGGGCGUGGAGGUGUCUGACCUGA CCCACGAUCUGGAGAUCUGUGCCGGGAGAAGUGUCUGGCCGUGGCCACAGAGAGCAACCA GAACAAUUCUAUACUACGCGUGCUGUUUGGCACCGGCGAGAAGGAGGAUAGGUCUGUGAAG CUGCGCAUCACAAAGAAGAUCCUGGAGGCCAUCAGCAACCUGAAGGAGAUCCCAAAGAAUG UGGCCCCCAUCCAGGAGAUCAUCCUGAAUGUGGCCAAGGCCACCAAGGAGACAUUCAGACA GGUGUACGCGAGGAAACCUGGGAGCACCACCCUGGAGAAGUUUAUCGCCAAGGACGGC CAGAAGGAGUUCGAUCUGAAGAAGCUGCAGACAGACCUGAAGAAAGUGAUCGGGGCAAGU CUAAGGAGAGAGAUUGGUGCUGUCAGGAGGAGCUGAGGAGCUACGUGGAGCAGAAUACCAU CCAGUAUGACCUGUGGGCCUGGGGCGAGAUGUUAACAAGGCCACACCGCCUGAAGAUC AAGUCCACAAGAAACUACAAUUUUGCCAAGCAGAGGCUGGAGCAGUUAAGGAGAUCCAGU CUCUGAACAAUCUGCUGUGGUGAAGAAGCUGAACGACUUUUUCGAUAGCGAGUUUUUCUC CGGCGAGGAGACCUACACAAUCUGCGUGCACCACCCUGGGCGGCAAGGACCUGUCCAAGCUG UAUAAGGCCUGGGAGGACGAUCCCGCCGAUCCUGAGAAUGCCAUCGUGGUGCUGGCGACG AUCUGAAGAACAUUUUUAGAAGGAGCCUAUCAGGAACAUCUGCGCUACAUUUCACCAU CCGCCAGGAGUGUAGCGCACAGGACAUCUGGCAGCAGCAAAGUACAAUCAGCAGCUGGAU CGGUUAUAAGAGCCAGAAGGCCAACCAUCCGUGCUGGGCAAUCAGGGCUUUUACCGGACAA ACGCCGUGAUCCUGCCAGAGAAGGCCAGCGGAACGACAGACCCAAUUCUCUGGAUCUGCG CAUCUGGCUGUACCUGAAGCUGCGGCACCCUGACGGCAGAUGGAAGAAGCACCACAUCCCA UUCUACGAUACCCGGUUUUUCCAGGAGAUCAUUGCCGCCGGCAAUAGCCUGUGGACACCU GUCAGUUUAGGACACCCCGCUUCGGCUAUCACCUGCCUAAGCUGACC GAUCAGACAGCCAU

	CCGCGUGAACAAAGAAGCACGUGAAGGCAGCAAAGACCGAGGCACGGAUCAGACUGGCCAUC CAGCAGGGGCACACUGCCAGUGUCCAAUCUGAAGAUACCCGAGAUUCUCCGCCACAAUACAUCU CUAAGGGCCAGGUGCGCAUCCCCGUGAAGUUUCGGGUGGGAAGGCAGAAGGGAACCCUGCA GAUCGGCGACCGGUUCUGCGGCUACGAUCAGAACCAGACAGCCUCUCACGCCUAUAGCCUG UGGGAGGUGGUGAAGGAGGGCCAGUACCACAAGGAGCUGGGCUGUUUUGUGCGCUUAUCU CUAGCGGGCACAUCGUGUCCAUCACCCGAGAACCAGGGGCAUACAGUUUGAUCAGCUGUCUUA UGAGGGCCUGGCCUACCCCCAGUAUGCCGACUGGAGAAAGAAGGCCUCCAAAGUUCGUGUCU CUGUGGCAGAUACCAAGAAGAACAAGAAGAAGGAGAUUCGUGACAGUGGAGGCCAAGGAGA AGUUUGACGCCAUCUGCAAGUACCAGCCUAGGCUGUAUAAGUUAACAAGGAGUACGCCUA UCUGCUGCGGGAUAUCGUGAGAGGCAAGAGCCUGGUGGAGCUGCAGCAGAUACGGCAGGAG AUCUUUCGUUAUCGAGCAGGACUGUGGAGUGACCCGCCUGGGAUUCUGAGCCUGUCCA CCCUGGAGACAGUGAAGGCCGUGAAGGGCAUACUACUCCUAUUUUUUAACAGCCUGAA UGCCUCUAAGAACAUAUCCCAUCAGCGACGAGCAGCGGAAGGAGUUUGAUCUGAGCUGUUC GCCUGCUGGAGAAGCUGGAGCUGAUCAGGACUCGGAAGAAGAAGCAGAAGGUGGAGAGAA UCGCCAAUAGCCUGAUCCAGACAUGCCUGGAGAACAUAUAAGUUCAUCAGGGGCGAGGG CGACCUGUCCACCACAAACAAUGCCACCAAGAAGAAGGCCAACUCUAGGAGCAUGGAUUGG CUGGCCAGAGGCGUGUUUAUAAGAUCGGGCAGCUGGGCCCCAAUGCACAACAUCACCCUGU UCGGCUGCGGCAGCCUGUACACAUCACCAGGACCCUCUGGUGCAGAGAAACCCAGAUAA GGCCAUAGAAGUGUAGAUGGGCAGCAAUCCAGUGAAGGACAUCGGCGAUUGGGUGCUGAGA AAGCUGUCCAGAACCCUGAGGGCCAAGAAUCGGGGCACCGGCGAGUACUAUACACAGGGCG UGAAGGAGUUCUGUCACUAUGAGCUGCAGGACCUGGAGGAGGAGCUGCUGAAGUGGCG GUCUGAUAGAAAGAGCAACAUCCUUGCUGGGUGCUGCAGAAUAGACUGGGCCGAGAAGCUG GGCAACAAGGAGGCGGUGGUGUACAUCCCAGUGAGGGGCGGGCCGAUCUAUUUGCAACCC ACAAGGUGGCAACAGGAGCCGUGAGCAUCGUGUUCGACCAGAAGCAAGUGUGGGUGUGUAA UGCAGAUACGUGGCAGCAGCAAAACAUCCGACUGACCGGCAAGGGCAUCGGCGAGCAGUCC UCUGACGAGGAGAACCCCGAUGGCCUCCAGGAUCAAGCUGCAGCUGACAUCUAAAAGGCCGG CGGCCACGAAAAAGGCCGCCAGGCAAAAAAGAAAAAGUAA (SEQ ID NO: 1089)
mRNA corresponding to variant Cas12i2 of SEQ ID NO: 927	AUGAGCUCGCCAUCAAGUCCUACAAGUCUGUGCUGCGGCCAAACGAGAGAAAGAAUCAGC UGCUGAAGUCCACCAUCCAGUGCCUGGAGGACGGCUCGCCUUCUUUUUCAAGAUGCUGCA GGGCCUGUUUGGCGGCAUACCCCCGAGAUCGUGAGAUUCAGCACAGAGCAGGAGAAGCAG CAGCAGGAUAUCGCCCUGUGGUGUGCCGUGAAUUGGUUCAGGCCUGUGAGCCAGGACUCCC UGACCCACACAACUGCCUCCGAUAACCCUGGUGGAGAAGUUUGAGGAGUACUAUGGCGGCAC AGCCAGCGACGCCAUCAAGCAGUACUUCAGCGCCUCCAUCGGCGAGUCCUACUAUUGGAU GACUGCCGCCAGCAGUACUAUGAUCUGUGCUGGGAGCUGGGCGUGGAGGUGUCUGACCUGA CCCACGAUCUGGAGAUCUGUGCCGGGAGAAGUGUCUGGCCUGGCCACAGAGAGCAACCA GAACAAUUCUAUACAGCGGUGUGUUUGGCACCGGCGAGAAGGAGGAUAGGUCUGUAGAG UGCGCAUACACAAAGAAGAUCCUGGAGGCCAUACGCAACCCUGAAGGAGAUCCAAAGAAG UGGCCCCCAUCCAGGAGAUCAUCCUGAAUGUGGCCAAGGCCACCAAGGAGACAUCAGACA GGUGUACGCAGGAACCCUGGGAGCACCAUCCACCCUGGAGAAGUUUAUCGCCAAGGACGGC CAGAAGGAGUUCGAUCUGAAGAAGCUGCAGACAGACCUGAAGAAGUGAUCGGGGCAAGU CUAAGGAGAGAGAUUGGUGCUGUCAGGAGGAGCUGAGGAGCUACUGGAGCAGAAUACCAU CCAGUAUGACCUGUGGGCCUGGGGCGAGAUGUUAACAAGGCCACACCGCCUGAAGAU AAGUCCACAAGAAACUACA AUUUUGCCAAGCAGAGGCUGGAGCAGUUAAGGAGAUCCAGU CUCUGAACAAUCUGCUGGUGGUGAAGAAGCUGAACGACUUUUUGCAUAGCGAGUUUUUC CGGCGAGGAGACCUACACAUCUGCGUGCACCACCUGGGCGGCAAGGACCUGUCCAAGCUG UAUAAGGCCUGGGAGGACGAUCCCGCCGAUCCUGAAGAUCCAUUGGUGUGCUGUGCGACG AUCUGAAGAACA AUUUUAAGAAGGAGCCUAUCAGGAACAUCUGCGCUACAUCUUCACCAU CCGCCAGGAGUGUAGCGCACAGGACAUCUCCUGGCAGCAGCAAGUAACAUAUCAGCAGCUGGAU CGGUUAUAAGAGCCAGAAGGCCAACCAUCCUGCUGGGCAAUACAGGGCUUUAACUGGACAA ACGCCUGAUCCUGCCAGAGAAGGCCAGCGGAACGACAGACCCAAUUCUCUGGAUCUGCG CAUCUGGCUGUACCUGAAGCUGCGGCACCCUGACGGCAGAUUGGAAGAAGCACCACAUCCCA UUCUACGAUACCCGGUUUUUCCAGGAGAUCAUAGCCCGCCGGCAAUAGCCUGUGGACACCU GUCAGUUUAGGACACCCCGCUUCGGCUAUCACCCUGCCUAAGCUGACCGAUCAGACAGCCAU CCGCGUGAACAAAGAAGCACGUGAAGGCAGCAAAGACCGAGGCACGGAUCAGACUGGCCAU CAGCAGGGCACACUGCCAGUGUCCAAUCUGAAGAUACCCGAGAUUCUCCGCCACAAUACAUCU CUAAGGGCCAGGUGCGCAUCCCCGUGAAGUUUCGGGUGGGAAGGCAGAAGGGAACCCUGCA GAUCGGCGACCGGUUCUGCGGCUACGAUCAGAACCAGACAGCCUCUCACGCCUAUAGCCUG UGGGAGGUGGUGAAGGAGGGCCAGUACCACAAGGAGCUGCGGUGUGCGGUGCGCUUAUCU CUAGCGGCACAUCGUGUCCAUCACCCGAGAACCAGGGGCAUACAGUUUGAUCAGCUGUCUUA

	UGAGGGCCUGGCCUACCCCCAGUAUGCCGACUGGAGAAAGAAGGCCUCCAAGUUCGUGUCU CUGUGGCAGAUACCAAGAAGAACAAGAAGAAGGAGAUUCGUGACAGUGGAGGCCAAGGAGA AGUUUGACGCCAUCUGCAAGUACCAGCCUAGGCUGUAUAAGUUAACAAGGAGUACGCCUA UCUGCUGCGGGAUAUCGUGAGAGGCAAGAGCCUGGUGGAGCUGCAGCAGAUACAGCAGGAG AUCUUUCGUUCAUCGAGCAGGACUGUGGAGUGACCCGCCUGGGAUUCUCUGAGCCUGUCCA CCCUGGAGACAGUGAAGGCCGUGAAGGGCAUCAUCUACUCCUAUUUUUCUACAGCCUGAA UGCCUCUAAGAACAUAUCCCAUCAGCGACGAGCAGCGGAAGGAGUUUGAUCCUGAGCUGUUC GCCCUGCUGGAGAAGCUGGAGCUGAUCAGGACUCGGAAGAAGAAGCAGAAGGUGGAGAGAA UCGCCAAUAGCCUGAUCCAGACAUGCCUGGAGAACAUAUCAAGUUAUCAGGGGCGAGGG CGACCUGUCCACCACAAACAAUGCCACCAAGAAGAAGGCCAACUCUAGGAGCAUGGAUUGG CUGGCCAGAGGCGUGUUUAAUAAGAUCGGCAGCUGGCCACCAUGCACAACAUCACCCUGU UCGGCUGCGGCAGCCUGUACACAUCCACCAGGACCCUCUGGUGCAGAGAAACCCAGAUAA GGCCAUAGAAGUGUAGAUGGGCAGCAUCCCAUGAAGGACAUCGGCGAUUGGGUGCUGAGA AAGCUGUCCAGAACCGAGGGCCAAGAAUCGGGGCACCGGCGAGUACUAUACACAGGGCG UGAAGGAGUUCUGUCACUAUGAGCUGCAGGACCGGAGGAGAGCUGCUGAAGUGGCG GUCUGAUAGAAAGAGCAACAUCUCCUUGCUGGGUGCUGCAGAAUAGACUGGCCGAGAAGCUG GGCAACAAGGAGGCGGUGUGUAUCAUCCAGUGAGGGGCGGCCGAUCUAUUUUGCAACCC ACAAGGUGGCAACAGGAGCCGUGAGCAUCGUGUUCGACCAGAAGCAAGUGGGGUGUGUAA UGCAGAUACGUGGCAGCAGCAACAUCCGACUGACCGGCAAGGGCAUCGGCCGCGAGUCC UCUGACGAGGAGAACCCCGAUGGCGGCAGGAUCAAGCUGCAGCUGACAUCUAAAAGGCCGG CGGCCACGAAAAAGGCCGCCAGGCAAAAAAGAAAAGUAA (SEQ ID NO: 1090)
--	--

Cas12i2 RNA guides were designed and ordered from Integrated DNA Technologies (IDT) as having 3' end modified phosphorothioated 2' O-methyl bases or 5' end and 3' end modified phosphorothioated 2' O-methyl bases guides, as specified in **Table 9**. Each variant Cas12i2 mRNA was mixed with a crRNA at a 1:1 (Cas12i2:crRNA) volume ratio (1050:1 crRNA:Cas12i2 molar ratio). The mRNA and crRNA were mixed immediately before electroporation. The primary human hepatocyte cells were cultured and electroporated as described in Example 3.

Table 9. Chemically modified RNA guide sequences

RNA guide	Sequence
3' end modified E2T5	AGAAAUCCGUCUUUCAUUGACGGCGGAGCAUCCUUGGAUA*mC*mA *mG (SEQ ID NO: 1091)
5' and 3' end modified E2T5	mA*mG*mA*AAUCCGUCUUUCAUUGACGGCGGAGCAUCCUUGGAUA *mC*mA*mG (SEQ ID NO: 1092)

10

FIG. 6 shows editing of an HAO1 target site by a variant Cas12i2 mRNA and 3' end modified E2T5 (SEQ ID NO: 1091) or 5' and 3' end modified E2T5 (SEQ ID NO: 1092). Indels in the HAO1 target site were introduced following electroporation of the Cas12i2 mRNA of SEQ ID NO: 1089 or SEQ ID NO: 1090 and either the RNA guide of SEQ ID NO: 1091 or SEQ ID NO: 1092. Approximately 50% NGS reads comprised an indel following electroporation of the Cas12i2 mRNA of SEQ ID NO: 1090 and the RNA guide of SEQ ID NO: 1091 or SEQ ID NO:

15

1092. Statistically significant higher % indels were observed using variant Cas12i2 mRNA of SEQ ID NO: 1090 compared to variant Cas12i2 mRNA of SEQ ID NO: 1089. No statistical difference was observed using 5' and 3' versus 3' only modifications to RNA guide E2T5.

This Example thus shows that HAO1 can be targeted by Cas12i2 mRNA constructs and chemically modified RNA guides in mammalian cells.

Example 6 - Off-Target Analysis of Cas12i2 and HAO1-Targeting RNA Guides

This Example describes on-target versus off-target assessment of a Cas12i2 variant and an HAO1-targeting RNA guide.

HEK293T cells were transfected with a plasmid encoding the variant Cas12i2 of SEQ ID NO: 924 or the variant Cas12i2 of SEQ ID NO: 927 and a plasmid encoding E2T5 (SEQ ID NO: 989), E1T2 (SEQ ID NO: 967), E1T3 (SEQ ID NO: 968), and E2T10 (SEQ ID NO: 994) according to the method described in Example 16 of PCT/US21/25257. The tagmentation-based tag integration site sequencing (TTISS) method described in Example 16 of PCT/US21/25257 was then carried out.

FIG. 7A and **FIG. 7B** show plots depicting on-target and off-target TTISS reads. The black wedge and centered number represent the fraction of on-target TTISS reads. Each grey wedge represents a unique off-target site identified by TTISS. The size of each grey wedge represents the fraction of TTISS reads mapping to a given off-target site. **FIG. 7A** shows TTISS reads for variant Cas12i2 of SEQ ID NO: 924, and **FIG. 7B** shows TTISS reads for variant Cas12i2 of SEQ ID NO: 927.

As shown in **FIG. 7A**, variant Cas12i2 of SEQ ID NO: 924 paired with E2T5 demonstrated a low likelihood of off-target editing, as 100% of TTISS reads mapped to the on-target. No TTISS reads mapped to potential off-target sites. E1T2 also showed a low likelihood of off-target editing. For E1T2, 98% of TTISS reads mapped to the on-target, and two potential off-target sites represented a combined 2% of TTISS reads. For E5T10, 95% of TTISS reads mapped to the on-target, and two potential off-target sites represented a combined 5% of TTISS reads. E2T10 demonstrated a higher likelihood of off-target editing using the TTISS method. For E2T10, only 65% of TTISS reads mapped to the on-target and 4 potential off-target sites represented the remaining combined 35% of TTISS reads. One potential off-target represented the majority of potential off-target TTISS reads for E2T10.

As shown in **FIG. 7B**, variant Cas12i2 of SEQ ID NO: 927 paired with E2T5 demonstrated a low likelihood of off-target editing, as 100% of TTISS reads mapped to the on-target. No TTISS reads mapped to potential off-target sites. Variant Cas12i2 of SEQ ID NO: 927 paired with the E1T2 or E1T3 also demonstrated a low likelihood of off-target editing. For E1T2, 100% of TTISS reads in replicate 1 and 96% of TTISS reads in replicate 2 mapped to the on-target; two potential off-target sites represented the remaining 4% of TTISS reads in replicate 2. For E1T3, 100% of TTISS reads in replicate 1 and 92% of TTISS reads in replicate 2 mapped to the on-target; two potential off-target sites represented the remaining 8% of TTISS reads in replicate 2.

Therefore, this Example shows that compositions comprising Cas12i2 and HAO1-targeting RNA guides comprise different off-target activity profiles.

Example 7 - HAO1 Protein Knockdown with Cas12i2 and HAO1-Targeting RNA Guides

This Example describes use of a Western Blot to identify knockdown of HAO1 protein using variant Cas12i2 of SEQ ID NO: 924 and HAO1-targeting RNA guides.

Primary hepatocyte cells from human donors were thawed from liquid nitrogen very quickly in a 37°C water bath. The cells were added to pre-warmed hepatocyte recovery media (Thermo Fisher, CM7000) and centrifuged at 100g for 10 minutes. The cell pellet was resuspended in appropriate volume of hepatocyte plating Medium (Williams' Medium E, Thermo Fisher A1217601 supplemented with Hepatocyte Plating Supplement Pack (serum-containing), Thermo Fisher CM3000). The cells were subjected to trypan blue viability count with an Inucyte disposable hemocytometer (Fisher scientific, 22-600-100). The cells were then washed in PBS and resuspended in P3 buffer + supplement (Lonza, VXP-3032) at a concentration of ~5000 cells/μL. Resuspended cells were dispensed at 500,000 cells/reaction into Lonza electroporation cuvettes

For the RNP reactions, E2T5 (SEQ ID NO: 989) was used as the HAO1-targeting RNA guides. RNPs were added to each reaction at a final concentration of 20 μM (Cas12i2), and transfection enhancer oligos were then added at a final concentration of 4 μM. Unelectroporated cells and cells electroporated without cargo were used as negative controls.

The strips were electroporated using an electroporation device (program CA137, Lonza 4D-nucleofector). Immediately following electroporation, pre-warmed Hepatocyte plating medium was added to each well and mixed very gently by pipetting. For each technical replicate

plate, 500,000 cells of diluted nucleofected cells were plated into a pre-warmed collagen-coated 24-well plate (Thermo Fisher) with wells containing Hepatocyte plating medium. The cells were then incubated at 37°C. The media was changed to hepatocyte maintenance media (Williams' Medium E, Thermo Fisher A1217601 supplemented with William's E medium Cell Maintenance Cocktail, Thermo Fisher CM 4000) after the cells attached after 24 hours. Fresh hepatocyte maintenance media was replaced every 48 hours.

16 days post RNP electroporation, the media was aspirated, and the cells were washed gently with PBS. Cells were then lysed with RIPA Lysis and Extraction buffer (Thermo Fisher 89901) + 1X protease inhibitors (Thermo Fisher 78440) for 30 minutes on ice, mixing the samples every 5 minutes. Cell lysate was quantified *via* Pierce BCA Protein Assay Kit (Thermo Fisher 23227). 15 µg of total protein per sample was prepared for SDS-PAGE in 1X Laemmli Sample buffer (BioRad 1610747) and 100mM DTT, then heated at 95C for 10 minutes. Samples were run on a 4-15% TGX gel (BioRad 5671084) at 200V for 45 minutes. Samples were transferred to a 0.2µm nitrocellulose membrane (BioRad 1704159) using the Trans Blot Turbo System. The membrane was blocked in Intercept TBS Blocking Buffer (Li-cor 927-60001) for 30 minutes at room temperature. The blot was then incubated in a 1:1000 dilution of primary anti-HAO1 antibody (Genetex GTX81144) and 1:2500 dilution of primary anti-vinculin antibody (Sigma V9131) in blocking buffer at 4C overnight. The blot was washed three times with TBST (ThermoFisher 28360) for 5 minutes each, then incubated with a 1:12500 dilution of IR680 anti-mouse (ThermoFisher PI35518) and IR800 anti-rabbit secondary antibodies (ThermoFisher PISA535571) in TBST for 1 hour at room temperature. The blot was then washed three times with TBST for 5 minutes each and visualized on the Li-cor Odyssey CLX.

Knockdown of HAO1 protein was observed in primary human hepatocytes at Day 7 post editing by Cas12i2 RNPs targeting the HAO1 gene with E2T5 (lanes 1-3 of **FIG. 8**). HAO1 knockdown was not observed for the buffer only controls (lanes 4-7).

This Example thus shows that HAO1 protein levels were decreased following editing with Cas12i2 and HAO1-targeting RNA guides.

OTHER EMBODIMENTS

All of the features disclosed in this specification may be combined in any combination. Each feature disclosed in this specification may be replaced by an alternative feature serving the same, equivalent, or similar purpose. Thus, unless expressly stated otherwise, each feature disclosed is only an example of a generic series of equivalent or similar features.

From the above description, one skilled in the art can easily ascertain the essential characteristics of the present invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions. Thus, other embodiments are also within the claims.

5

EQUIVALENTS

While several inventive embodiments have been described and illustrated herein, those of ordinary skill in the art will readily envision a variety of other means and/or structures for performing the function and/or obtaining the results and/or one or more of the advantages described herein, and each of such variations and/or modifications is deemed to be within the scope of the inventive embodiments described herein. More generally, those skilled in the art will readily appreciate that all parameters, dimensions, materials, and configurations described herein are meant to be exemplary and that the actual parameters, dimensions, materials, and/or configurations will depend upon the specific application or applications for which the inventive teachings is/are used. Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific inventive embodiments described herein. It is, therefore, to be understood that the foregoing embodiments are presented by way of example only and that, within the scope of the appended claims and equivalents thereto, inventive embodiments may be practiced otherwise than as specifically described and claimed. Inventive embodiments of the present disclosure are directed to each individual feature, system, article, material, kit, and/or method described herein. In addition, any combination of two or more such features, systems, articles, materials, kits, and/or methods, if such features, systems, articles, materials, kits, and/or methods are not mutually inconsistent, is included within the inventive scope of the present disclosure.

25 All definitions, as defined and used herein, should be understood to control over dictionary definitions, definitions in documents incorporated by reference, and/or ordinary meanings of the defined terms.

All references, patents and patent applications disclosed herein are incorporated by reference with respect to the subject matter for which each is cited, which in some cases may encompass the entirety of the document.

30 The indefinite articles “a” and “an,” as used herein in the specification and in the claims, unless clearly indicated to the contrary, should be understood to mean “at least one.”

The phrase “and/or,” as used herein in the specification and in the claims, should be understood to mean “either or both” of the elements so conjoined, *i.e.*, elements that are conjunctively present in some cases and disjunctively present in other cases. Multiple elements listed with “and/or” should be construed in the same fashion, *i.e.*, “one or more” of the elements so conjoined. Other elements may optionally be present other than the elements specifically identified by the “and/or” clause, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, a reference to “A and/or B”, when used in conjunction with open-ended language such as “comprising” can refer, in one embodiment, to A only (optionally including elements other than B); in another embodiment, to B only (optionally including elements other than A); in yet another embodiment, to both A and B (optionally including other elements); etc.

As used herein in the specification and in the claims, “or” should be understood to have the same meaning as “and/or” as defined above. For example, when separating items in a list, “or” or “and/or” shall be interpreted as being inclusive, *i.e.*, the inclusion of at least one, but also including more than one, of a number or list of elements, and, optionally, additional unlisted items. Only terms clearly indicated to the contrary, such as “only one of” or “exactly one of,” or, when used in the claims, “consisting of,” will refer to the inclusion of exactly one element of a number or list of elements. In general, the term “or” as used herein shall only be interpreted as indicating exclusive alternatives (*i.e.*, “one or the other but not both”) when preceded by terms of exclusivity, such as “either,” “one of,” “only one of,” or “exactly one of.” “Consisting essentially of,” when used in the claims, shall have its ordinary meaning as used in the field of patent law.

As used herein in the specification and in the claims, the phrase “at least one,” in reference to a list of one or more elements, should be understood to mean at least one element selected from any one or more of the elements in the list of elements, but not necessarily including at least one of each and every element specifically listed within the list of elements and not excluding any combinations of elements in the list of elements. This definition also allows that elements may optionally be present other than the elements specifically identified within the list of elements to which the phrase “at least one” refers, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, “at least one of A and B” (or, equivalently, “at least one of A or B,” or, equivalently “at least one of A and/or B”) can refer, in one embodiment, to at least one, optionally including more than one, A, with no B present (and

optionally including elements other than B); in another embodiment, to at least one, optionally including more than one, B, with no A present (and optionally including elements other than A); in yet another embodiment, to at least one, optionally including more than one, A, and at least one, optionally including more than one, B (and optionally including other elements); etc.

- 5 It should also be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one step or act, the order of the steps or acts of the method is not necessarily limited to the order in which the steps or acts of the method are recited.

WHAT IS CLAIMED IS:

1. A gene editing system for genetic editing of a hydroxyacid oxidase 1 (HAO1) gene, comprising
 - (i) a Cas12i2 polypeptide or a first nucleic acid encoding the Cas12i2 polypeptide, wherein the Cas12i2 polypeptide comprises an amino acid sequence at least 95% identical to SEQ ID NO: 922 and comprises one or more mutations relative to SEQ ID NO: 922;
 - (ii) an RNA guide or a second nucleic acid encoding the RNA guide, wherein the RNA guide comprises a spacer sequence specific to a target sequence within an HAO1 gene, the target sequence being adjacent to a protospacer adjacent motif (PAM) comprising the motif of 5'-TTN-3', which is located 5' to the target sequence.
2. The gene editing system of claim 1, wherein the one or more mutations in the Cas12i2 polypeptide are at positions D581, G624, F626, P868, I926, V1030, E1035, and/or S1046 of SEQ ID NO: 922.
3. The gene editing system of claim 1 or claim 2, wherein the one or more mutations are amino acid substitutions, which optionally is D581R, G624R, F626R, P868T, I926R, V1030G, E1035R, S1046G, or a combination thereof.
4. The gene editing gene editing system of claim 3, wherein the Cas12i2 polypeptide comprises:
 - (i) mutations at positions D581, D911, I926, and V1030, which optionally are amino acid substitutions of D581R, D911R, I926R, and V1030G;
 - (ii) mutations at positions D581, I926, and V1030, which optionally are amino acid substitutions of D581R, I926R, and V1030G;
 - (iii) mutations at positions D581, I926, V1030, and S1046, which optionally are amino acid substitutions of D581R, I926R, V1030G, and S1046G;
 - (iv) mutations at positions D581, G624, F626, I926, V1030, E1035, and S1046, which optionally are amino acid substitutions of D581R, G624R, F626R, I926R, V1030G, E1035R, and S1046G; or

(v) mutations at positions D581, G624, F626, P868, I926, V1030, E1035, and S1046, which optionally are amino acid substitutions of D581R, G624R, F626R, P868T, I926R, V1030G, E1035R, and S1046G.

5 5. The gene editing system of claim 1, wherein the Cas12i2 polypeptide comprises the amino acid sequence of SEQ ID NO: 923, 924, 925, 926, or 927, optionally wherein the Cas12i2 polypeptide comprises the amino acid sequence of SEQ ID NO: 924 or 927.

10 6. The gene editing system of any one of claims 1-5, which comprises the first nucleic acid encoding the Cas12i2 polypeptide.

7. The gene editing system of claim 6, wherein the first nucleic acid is a messenger RNA (mRNA).

15 8. The gene editing system of claim 7, wherein the first nucleic acid is included in a viral vector, which optionally is an adeno-associated viral (AAV) vector.

20 9. The gene editing system of any one of claims 1-8, wherein the target sequence is within exon 1 or exon 2 of the HAO1 gene.

25 10. The gene editing system of claim 9, wherein the target sequence comprises:
 (i) 5'-CAAAGTCTATATATGACTAT-3' (SEQ ID NO:1025);
 (ii) 5'-GGAAGTACTGATTTAGCATG-3' (SEQ ID NO: 1026);
 (iii) 5'-TAGATGGAAGCTGTATCCAA-3' (SEQ ID NO: 1046);
 (iv) 5'-CGGAGCATCCTTGGATACAG-3' (SEQ ID NO: 1047); or
 (v) 5'-AGGACAGAGGGTCAGCATGC-3 (SEQ ID NO: 1052).

30 11. The system of claim 10, wherein the spacer sequence comprises:
 (i) 5'-CAAAGUCUAUAUAUGACUAU-3' (SEQ ID NO: 1093);
 (ii) 5'-GGAAGUACUGAUUUAGCAUG-3' (SEQ ID NO: 1094);
 (iii) 5'-UAGAUGGAAGCUGUAUCCAA-3' (SEQ ID NO: 1095);
 (iv) 5'-CGGAGCAUCCUUGGAUACAG-3' (SEQ ID NO: 1096); or

(v) 5'-AGGACAGAGGGUCAGCAUGC-3 (SEQ ID NO: 1097).

12. The gene editing system of any one of claims 1-11, wherein the spacer sequence is 20-30-nucleotide in length, optionally wherein the spacer is 20-nucleotide in length.

5

13. The gene editing system of any one of claims 1-12, wherein the RNA guide comprises the spacer and a direct repeat sequence.

14. The gene editing system of claim 13, wherein the direct repeat sequence is 23-36-nucleotide in length.

10

15. The gene editing system of claim 14, wherein the direct repeat sequence is at least 90% identical to any one of SEQ ID NOs: 1-10 or a fragment thereof that is at least 23-nucleotide in length.

15

16. The gene editing system of claim 15, wherein the direct repeat sequence is any one of SEQ ID NOs: 1-10, or a fragment thereof that is at least 23-nucleotide in length.

17. The gene editing system of claim 16, wherein the direct repeat sequence is 5'-AGAAAUCCGUCUUUCAUUGACGG-3' (SEQ ID NO: 10).

20

18. The gene editing system of claim 1, wherein the RNA guide comprises the nucleotide sequence of:

(i) 5'-AGAAAUCCGUCUUUCAUUGACGGCAAAGUCUAUAUAUGACUAU-3' (SEQ ID NO: 967);

25

(ii) 5'-AGAAAUCCGUCUUUCAUUGACGGGGAAGUACUGAUUUAGCAUG-3' (SEQ ID NO: 968);

(iii) 5'-AGAAAUCCGUCUUUCAUUGACGGUAGAUGGAAGCUGUAUCCAA-3' (SEQ ID NO: 988);

30

(iv) 5'-AGAAAUCCGUCUUUCAUUGACGGCGGAGCAUCCUUGGAUACAG-3' (SEQ ID NO: 989); or

(v) 5'-AGAAAUCCGUCUUUCAUUGACGGAGGACAGAGGGUCAGCAUGC-3'
(SEQ ID NO: 994).

19. The gene editing system of any one of claims 1-18, wherein the system comprises
5 the second nucleic acid encoding the RNA guide.

20. The gene editing system of claim 19, wherein the nucleic acid encoding the RNA
guide is located in a viral vector.

10 21. The gene editing system of any one of claims 7-20, wherein the viral vector
comprises the both the first nucleic acid encoding the Cas12i2 polypeptide and the second
nucleic acid encoding the RNA guide.

22. The gene editing system of any one of claims 1-20, wherein the system comprises
15 the first nucleic acid encoding the Cas12i2 polypeptide, which is located in a first vector, and
wherein the system comprises the second nucleic acid encoding the RNA guide, which is located
in a second vector; optionally wherein the first and/or the second vector is a viral vector.

23. The gene editing system of claim 22, wherein the first and second vector are the
20 same vector.

24. The gene editing system of any one of claims 1-23, wherein the system comprises
one or more lipid nanoparticles (LNPs), which encompass (i), (ii), or both.

25 25. The gene editing system of claim 24, wherein the system comprises the LNP,
which encompass (i), and wherein the system comprises a viral vector comprising the second
nucleic acid encoding the RNA guide; optionally wherein the viral vector is an AAV vector.

26. The gene editing system of claim 24, wherein the system comprises the LNP,
30 which encompass (ii), and wherein the system comprises a viral vector comprising the first
nucleic acid encoding Cas12i2 polypeptide; optionally wherein the viral vector is an AAV
vector.

27. A gene editing system for genetic editing of a hydroxyacid oxidase 1 (HAO1) gene, comprising

- 5 (i) a Cas12i polypeptide or a first nucleic acid encoding the Cas12i polypeptide, optionally wherein the Cas12i polypeptide is a Cas12i2 polypeptide;
- (ii) an RNA guide or a second nucleic acid encoding the RNA guide, wherein the RNA guide comprises a spacer sequence specific to a target sequence within exon 1 or exon 2 of an HAO1 gene, the target sequence being adjacent to a protospacer adjacent motif (PAM) comprising the motif of 5'-TTN-3', which is located 5' to the target sequence.

10

28. The gene editing system of claim 27, wherein the target sequence comprises:

- (i) 5'-CAAAGTCTATATATGACTAT-3' (SEQ ID NO: 1025);
- (ii) 5'-GGAAGTACTGATTTAGCATG-3' (SEQ ID NO: 1026);
- (iii) 5'-TAGATGGAAGCTGTATCCAA-3' (SEQ ID NO: 1046);
- 15 (iv) 5'-**CGGAGCATCCTTGGATACAG**-3' (SEQ ID NO: 1047); or
- (v) 5'-AGGACAGAGGGTCAGCATGC-3 (SEQ ID NO: 1052).

29. The gene editing system of claim 27, wherein the spacer sequence comprises:

- (i) 5'-CAAAGUCUAUAUAUGACUAU-3' (SEQ ID NO: 1093);
- 20 (ii) 5'-GGAAGUACUGAUUUAGCAUG-3' (SEQ ID NO: 1094);
- (iii) 5'-UAGAUGGAAGCUGUAUCCAA-3' (SEQ ID NO: 1095);
- (iv) 5'-**CGGAGCAUCCUUGGAUACAG**-3' (SEQ ID NO: 1096); or
- (v) 5'-AGGACAGAGGGUCAGCAUGC-3 (SEQ ID NO: 1097).

25 30. The gene editing system of any one of claims 27-29, which comprises the first nucleic acid encoding the Cas12i polypeptide.

31. The gene editing system of claim 30, wherein the first nucleic acid is a messenger RNA (mRNA).

30

32. The gene editing system of claim 30, wherein the first nucleic acid is included in a viral vector, which optionally is an adeno-associated viral (AAV) vector.

33. The gene editing system of any one of claims 27-32, wherein the spacer is 20-30-nucleotide in length, optionally wherein the spacer is 20-nucleotide in length.

5 34. The gene editing system of any one of claims 27-33, wherein the RNA guide comprises the spacer sequence and a direct repeat sequence.

35. The gene editing system of claim 34, wherein the direct repeat sequence is 23-36-nucleotide in length.

10

36. The gene editing system of claim 35, wherein the direct repeat sequence is at least 90% identical to any one of SEQ ID NOs: 1-10 or a fragment thereof that is at least 23-nucleotide in length.

15 37. The gene editing system of claim 36, wherein the direct repeat sequence is any one of SEQ ID NOs: 1-10, or a fragment thereof that is at least 23-nucleotide in length.

38. The gene editing system of claim 37, wherein the direct repeat sequence is 5'-AGAAAUCCGUCUUUCAUUGACGG-3' (SEQ ID NO: 10).

20

39. The gene editing system of claim 34, wherein the RNA guide comprises the nucleotide sequence of:

- (i) 5'-AGAAAUCCGUCUUUCAUUGACGGCAAAGUCUAUAUAUGACUAU-3' (SEQ ID NO: 967);
- 25 (ii) 5'-AGAAAUCCGUCUUUCAUUGACGGGGAAGUACUGAUUUAGCAUG-3' (SEQ ID NO: 968);
- (iii) 5'-AGAAAUCCGUCUUUCAUUGACGGUAGAUGGAAGCUGUAUCCAA-3' (SEQ ID NO: 988);
- (iv) 5'-AGAAAUCCGUCUUUCAUUGACGGCGGAGCAUCCUUGGAUACAG-3' (SEQ ID NO: 989); or
- 30 (v) 5'-AGAAAUCCGUCUUUCAUUGACGGAGGACAGAGGGUCAGCAUGC-3' (SEQ ID NO: 994).

40. The gene editing system of any one of claims 27-39, wherein the system comprises the second nucleic acid encoding the RNA guide.

5 41. The gene editing system of claim 40, wherein the nucleic acid encoding the RNA guide is located in a viral vector.

42. The gene editing system of any one of claims 32-41, wherein the viral vector comprises the both the first nucleic acid encoding the Cas12i2 polypeptide and the second
10 nucleic acid encoding the RNA guide.

43. The gene editing system of any one of claims 27-42, wherein the system comprises the first nucleic acid encoding the Cas12i2 polypeptide, which is located in a first vector, and wherein the system comprises the second nucleic acid encoding the RNA guide,
15 which is located in a second vector.

44. The gene editing system of any one of claims 27-43, wherein the system comprises one or more lipid nanoparticles (LNPs), which encompass (i), (ii), or both.

20 45. The gene editing system of claim 44, wherein the system comprises the LNP, which encompass (i), and wherein the system comprises a viral vector comprising the second nucleic acid encoding the RNA guide; optionally wherein the viral vector is an AAV vector.

46. The gene editing system of claim 44, wherein the system comprises the LNP,
25 which encompass (ii), and wherein the system comprises a viral vector comprising the first nucleic acid encoding Cas12i2 polypeptide; optionally wherein the viral vector is an AAV vector.

47. A pharmaceutical composition comprising the gene editing system set forth in any
30 one of claims 1-46.

48. A kit comprising the elements (i) and (ii) of the gene editing system set forth in any one of claims 1-46.

49. A method for editing a hydroxyacid oxidase 1 (HAO1) gene in a cell, the method
5 comprising contacting a host cell with the gene editing system for editing the HAO1 gene set forth in any one of claims 1-46 to genetically edit the HAO1 gene in the host cell.

50. The method of claim 49, wherein the host cell is cultured *in vitro*.

10 51. The method of claim 49, wherein contacting step is performed by administering the system for editing the HAO1 gene to a subject comprising the host cell.

52. A cell comprising a disrupted a hydroxyacid oxidase 1 (HAO1) gene, wherein the cell optionally is produced by contacting a host cell with the gene editing system of any one of
15 claims 1-46 to genetically edit the HAO1 gene in the host cell, thereby disrupting the HAO1 gene.

53. A method for treating primary hyperoxaluria (PH) in a subject, comprising administering to a subject in need thereof a gene editing system for editing a hydroxyacid
20 oxidase 1 (HAO1) gene set forth in any one of claims 1-46 or the cell of claim 52.

54. The method of claim 53, wherein the subject is a human patient having the PH, which optionally is PH1, PH2, or PH3.

25 55. The method of claim 54, wherein the PH is PH1.

56. An RNA guide, comprising (i) a spacer sequence that is specific to a target sequence in a hydroxyacid oxidase 1 (HAO1) gene, wherein the target sequence is adjacent to a protospacer adjacent motif (PAM) comprising the motif of 5'-TTN-3', which is located 5' to the
30 target sequence; and (ii) a direct repeat sequence.

57. The RNA guide of claim 56, wherein the spacer is 20-30-nucleotide in length, optionally 20-nucleotide in length.

58. The RNA guide of claim 56 or claim 57, wherein the direct repeat sequence is 23-36-nucleotide in length, optionally 23-nucleotide in length.

59. The RNA guide of any one of claims 56-58, wherein the target sequence is within exon 1 or exon 2 of the HAO1 gene.

60. The RNA guide of claim 59, wherein the target sequence comprises:
(i) 5'-CAAAGTCTATATATGACTAT-3' (SEQ ID NO: 1025);
(ii) 5'-GGAAGTACTGATTAGCATG-3' (SEQ ID NO: 1026);
(iii) 5'-TAGATGGAAGCTGTATCCAA-3' (SEQ ID NO: 1046);
(iv) 5'-**CGGAGCATCCTTGGATACAG**-3' (SEQ ID NO: 1047); or
(v) 5'-AGGACAGAGGGTCAGCATGC-3 (SEQ ID NO: 1052).

61. The RNA guide of claim 60, wherein the spacer sequence comprises:
(i) 5'-CAAAGUCUAUAUAUGACUAU-3' (SEQ ID NO: 1093);
(ii) 5'-GGAAGUACUGAUUUAGCAUG-3' (SEQ ID NO: 1094);
(iii) 5'-UAGAUGGAAGCUGUAUCCAA-3' (SEQ ID NO: 1095);
(iv) 5'-**CGGAGCAUCCUUGGAUACAG**-3' (SEQ ID NO: 1096); or
(v) 5'-AGGACAGAGGGUCAGCAUGC-3 (SEQ ID NO: 1097).

62. The RNA guide of any one of claims 56-61, wherein the direct repeat sequence is at least 90% identical to any one of SEQ ID NOs: 1-10 or a fragment thereof that is at least 23-nucleotide in length.

63. The RNA guide of claim 62, wherein the direct repeat sequence is any one of SEQ ID NOs: 1-10, or a fragment thereof that is at least 23-nucleotide in length.

64. The RNA guide of claim 63, wherein the direct repeat sequence is 5'-AGAAAUCCGUCUUUCAUUGACGG-3' (SEQ ID NO: 10).

65. The RNA guide of claim 56, which comprises the nucleotide sequence of:

(i) 5'-AGAAAUCCGUCUUUCAUUGACGGCAAAGUCUAUAUAUGACUAU-3'
(SEQ ID NO: 967);

5 (ii) 5'-AGAAAUCCGUCUUUCAUUGACGGGGAAGUACUGAUUUAGCAUG-3'
(SEQ ID NO: 968);

(iii) 5'-AGAAAUCCGUCUUUCAUUGACGGUAGAUGGAAGCUGUAUCCAA-3'
(SEQ ID NO: 988);

10 (iv) 5'-AGAAAUCCGUCUUUCAUUGACGGCGGAGCAUCCUUGGAUACAG-3'
(SEQ ID NO: 989); or

(v) 5'-AGAAAUCCGUCUUUCAUUGACGGAGGACAGAGGGUCAGCAUGC-3'
(SEQ ID NO: 994).