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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 January 2019 (24.01.2019)

WIPO | PCT

(10) International Publication Number
WO 2019/017988 A1

(51) International Patent Classification:
C07H 21/02 (2006.01) *C12N 15/63* (2006.01)
C12N 15/00 (2006.01) *C12N 15/85* (2006.01)
C12N 15/113 (2010.01) *C12Q 1/02* (2006.01)

(21) International Application Number:
PCT/US2017/054271

(22) International Filing Date:
29 September 2017 (29.09.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/535,499 21 July 2017 (21.07.2017) US

(71) Applicant: **ARIZONA BOARD OF REGENTS ON BEHALF OF ARIZONA STATE UNIVERSITY** [US/US]; 1475 North Scottsdale Road, Suite 200, Scottsdale, AZ 85257 (US).

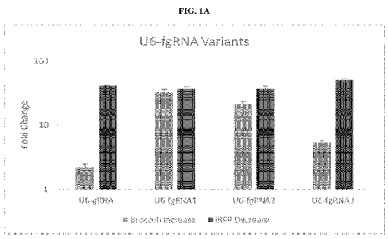
(72) Inventors: **KIANI, Samira**; 10240 N. 77th Street, Scottsdale, AZ 85258 (US). **WANG, Xiao**; 1900 W. Pelican Dr., Chandler, AZ 85286 (US). **MENN, David**; P.o.box 879809, Tempe, AZ 85287-9709 (US). **EBRAHIMKHANI, Mohammad, Reza**; 10240 N. 77th Street, Scottsdale, AZ 85258 (US).

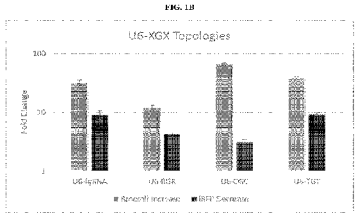
(74) Agent: **LEWIS, Jessica, L.**; Quarles & Brady LLP, 33 East Main Sreet, Suite 900, Madison, WI 53703-3095 (US).

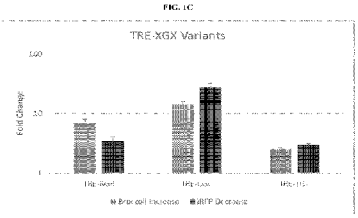
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS,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

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

(54) Title: CRISPR FLUORESCENT GUIDE RNA (fgRNA) FOR UNDERSTANDING gRNAs EXPRESSED FROM POL II PROMOTERS







(57) Abstract: Provided herein are tools for understanding and engineering dynamics of synthetic genetic circuits which utilize CRISPR components. More particularly, methods, systems, and compositions for directing Cas9 activity using a fluorescent guide RNA (fgRNA) which fluoresces in the presence of small molecules (e.g., DFHBI-IT) are described and illustrated in the present provisional application.

WO 2019/017988 A1

[Continued on next page]

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

CRISPR Fluorescent Guide RNA (fgRNA) For Understanding gRNAs Expressed From Pol II Promoters

BACKGROUND

[0001] Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) technology has become a staple of the field of synthetic biology, with such diverse applications as induced mutagenesis, insertion/deletion for silencing genes, gene activation and repression, and genome labeling. This is due in a large part to the technology's specificity, ease of design, and modularity.

[0002] The CRISPR system consists of a guide RNA (gRNA) and the protein Cas9. When coexpressed, the gRNA and Cas9 complex with each other, and the gRNA targets Cas9 to a specific sequence of DNA through complementary base-pairing with a ~20 nucleotide region on the 5' end of the gRNA. There are few reliable, non-invasive methods by which gRNA expression levels can be measured. Those that do exist are indirect, such as activating or repressing a downstream fluorescent gene, and are obfuscated by complex gRNA/Cas9 and gRNA/DNA interactions. Accordingly, there remains a need for direct, non-invasive methods for understanding the dynamics of genetic circuits which utilize CRISPR components.

SUMMARY

[0003] Provided herein are tools for understanding and engineering dynamics of synthetic genetic circuits which utilize CRISPR components. More particularly, systems, methods, and compositions for directing Cas9 activity using a fluorescent guide RNA (fgRNA) which fluoresces in the presence of small molecules (e.g., DFHBI-1T).

[0004] In a first aspect, provided herein is a method of directly detecting and measuring gRNA expression in a cell. The method can comprise or consist essentially of (a) introducing a synthetic regulatory system into the cell, the synthetic regulatory system comprising: (i) a nucleotide sequence encoding an inactivated Cas nuclease; and (ii) a fluorescent guide RNA (fgRNA) operably linked to a CRISPR-responsive promoter, the fgRNA comprising a fluorophore-binding RNA aptamer sequence; and a 20-nucleotide (nt) spacer sequence that hybridizes with a target sequence of a DNA molecule in a eukaryotic cell that contains the DNA molecule; and (b) contacting the cell comprising the synthetic regulatory system to a small molecule inducer of the fluorophore binding RNA aptamer; wherein gRNA expression is measured

by fluorescence monitoring. The CRISPR-responsive promoter can be a RNA Pol II promoter. The RNA Pol II promoter can be doxycycline inducible. The fgRNA can be configured for ribozyme-mediated editing, Csy4-mediated editing, or tRNA-mediated editing. The fluorophore binding RNA aptamer sequence can be selected from the group consisting of a Broccoli aptamer, Spinach aptamer, and Spinach2 aptamer. The small molecule inducer can be DFHBI-1T. The fluorophore RNA aptamer sequence can be selected from the group consisting of SEQ ID NO:1. SEQ ID NO:2, and SEQ ID NO:3.

[0005] In another aspect, provided herein is a method of monitoring efficiency of CRISPR-mediated gene repression. The method can comprise or consist essentially of (a) introducing a synthetic regulatory system into the cell, the synthetic regulatory system comprising: (i) a nucleotide sequence encoding an inactivated Cas nuclease; and (ii) a fluorescent guide RNA (fgRNA) operably linked to a CRISPR-responsive promoter, the fgRNA comprising a fluorophore-binding RNA aptamer sequence; and a 20-nucleotide (nt) spacer sequence that hybridizes with a target sequence of a DNA molecule in a eukaryotic cell that contains the DNA molecule; and (b) contacting the cell comprising the synthetic regulatory system to a small molecule inducer of the fluorophore binding RNA aptamer; wherein gRNA expression is measured by fluorescence monitoring. The CRISPR-repressible promoter can be a doxycycline inducible RNA Pol II promoter. The fgRNA can be configured for ribozyme-mediated editing, Csy4-mediated editing, or tRNA-mediated editing. The fluorophore binding RNA aptamer sequence can be selected from the group consisting of a Broccoli aptamer, Spinach aptamer, and Spinach2 aptamer. The small molecule inducer can be DFHBI-1T. The fluorophore RNA aptamer sequence can be selected from the group consisting of SEQ ID NO:1. SEQ ID NO:2, and SEQ ID NO:3.

[0006] In a further aspect, provided herein is a synthetic regulatory system where the system comprises or consists essentially of (i) a nucleotide sequence encoding a Cas nuclease; and (ii) a fluorescent guide RNA (fgRNA) operably linked to a CRISPR-responsive promoter. The fgRNA can comprise a fluorophore-binding RNA aptamer sequence, and a 20-nucleotide (nt) spacer sequence that hybridizes with a target gene sequence of a DNA molecule in a eukaryotic cell that contains the DNA molecule. The CRISPR-repressible promoter can be a doxycycline inducible RNA Pol II promoter. The RNA Pol II promoter can be doxycycline inducible. The fgRNA can be configured for ribozyme-mediated editing, Csy4-mediated editing, or tRNA-mediated editing. The fluorophore-binding RNA aptamer sequence can be selected from the group

consisting of a Broccoli aptamer, Spinach aptamer, and Spinach2 aptamer. The small molecule inducer can be DFHBI-1T. The fluorophore RNA aptamer sequence can be selected from the group consisting of SEQ ID NO:1. SEQ ID NO:2, and SEQ ID NO:3.

[0007] In another aspect, provided herein is a fluorescent guide RNA (fgRNA) operably linked to a CRISPR-responsive promoter, the fgRNA comprising a fluorophore-binding RNA aptamer sequence; and a 20-nucleotide (nt) spacer sequence that hybridizes with a target sequence of a DNA molecule in a eukaryotic cell that contains the DNA molecule. The fgRNA can be configured for ribozyme-mediated editing, Csy4-mediated editing, or tRNA-mediated editing. The fluorophore-binding RNA aptamer sequence can be selected from the group consisting of a Broccoli aptamer, Spinach aptamer, and Spinach2 aptamer. The fluorophore RNA aptamer sequence can be selected from the group consisting of SEQ ID NO:1. SEQ ID NO:2, and SEQ ID NO:3.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] FIGS. 1A-1C presents flow cytometry data. (A) Pol III driven fgRNA functionality with the Broccoli sequence inserted into varying parts of the gRNA sequence. (B) Pol III driven fgRNA functionality when paired with 3 different post-transcriptional editing methods: ribozyme-guide-ribozyme (RGR), Csy4-guide-Csy4 (CGC), and tRNA-guide-tRNA (TGT). (C) fgRNA functionality as assayed by fluorescence expression when fgRNA is driven by a doxycycline-inducible Pol II (TRE) promoter and post-transcriptionally edited by different methods. All data are the mean of geometric means of 4 replicates $\pm$ standard deviation.

[0009] FIGS. 2A-2C. (A) Fluorescent RNA apatamer Broccoli and its three potential insertions sites within the gRNA (top). Flow cytometry data (bottom) indicate that insertion into the "1" position has no effect on gRNA repression strength and is detectably green. (B) A comparison of three common methods of editing a long RNA transcript into a gRNA using a U6 (pol III) promoter: topologies (top) and their cleavage locations (arrows) and flow cytometry results (bottom). Addition of flanking editor sequences has a measureable impact on gRNA repression. (C) Initial pol II gRNA expression efficiency in a simple repressor circuit. Topology (top) and flow cytometry results (bottom). All flow cytometry is the mean of four replicates $\pm$ SD.

[0010] While the present invention is susceptible to various modifications and alternative forms, exemplary embodiments thereof are shown by way of example in the drawings and are

herein described in detail. It should be understood, however, that the description of exemplary embodiments is not intended to limit the invention to the particular forms disclosed, but on the contrary, the intention is to cover all modifications, equivalents and alternatives falling within the spirit and scope of the invention as defined by the appended claims.

DETAILED DESCRIPTION

[0011] All publications, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference as though set forth in their entirety in the present application.

[0012] The compositions and methods described herein are based at least in part on the inventors' development of a nucleic acid construct capable of selectively directing Cas9 activity and fluorescing in the presence of a small molecule, and which permits direct, non-invasive measurement of CRISPR gRNA expressed from various promoters including RNA Pol II promoters.

[0013] In a first aspect, therefore, provided herein is a guide RNA construct comprising a fluorophore-binding nucleic acid aptamer. Useful nucleic acid aptamers have the capacity to bind to a small molecule dye whose fluorescence is switched on upon binding the RNA. In this manner, the nucleic acid aptamer binds selectively to conditionally fluorescent molecules ("fluorophores").

[0014] Any appropriate fluorophore-binding nucleic acid aptamer can be used according to the methods and systems described herein. Preferably, the fluorophore binding nucleic acid aptamer is the broccoli aptamer. When the small molecule DFHBI-1T is introduced to cells cultured in a growth medium, it permeates the cell membranes, binds to the broccoli aptamer, and fluoresces green. As described in U.S. Patent No. 9,664,676 (which is hereby incorporated by reference in its entirety), the core sequence element of an aptamer that induces the fluorescence of the fluorophore DFHBI-1T has been identified to be, as follows: GAGANGGUCGGGUCCAGN-N-GCUGUNGAGUAGAGUGUGGGCUC (SEQ ID NO:4), where N at each of positions 5, 18, and 25 can be any single nucleotide base (A, U, G, or C), and N at position 19 can be any single nucleotide base (A, U, G, or C) or an insertion of any length of various nucleotide bases. Numerous sequences of different lengths have been placed at position 19 and the resulting aptamer retained fluorescence. The core can also be preceded or followed by any arbitrary sequence, but the core is needed for fluorescence.

[0015] Other fluorophore binding nucleic acid aptamers that can be used include, without limitation, Spinach and Spinach2 (Paige et al., "RNA mimics of green fluorescent protein," Science 333: 642-646 (2011), which is hereby incorporated by reference in its entirety), Baby Spinach, and "Bunch of Baby Spinach" (BoBS) aptamers (Shechner et al., *Nature Methods* 12, 664–670 (2015)). The small molecule fluorophore ligand ("inducer") for these aptamers is DFHBI–1T. In certain embodiments, the guide RNA construct is configured to direct Cas9 activity and to fluoresce in the presence of the small molecule DFHBI-1T. As used herein, the term "aptamer" refers to nucleic acid molecules and encompasses nucleic acid molecules that bind specifically to small molecule fluorophores and are useful for *in vitro* or *in vivo* monitoring of the activity, trafficking or localization, degradation, or quantification of various molecules. In addition to the methods described and exemplified herein, any conventional method for imaging, visualizing, or quantifying fluorescent molecules can be used in accordance with this disclosure.

[0016] Nucleotide sequences of three exemplary fgRNA variants are listed below:

fgRNA1:

NNNNNNNNNNNNNNNNNNNNGTTTGAGAGCTAGCGCAGACGGTCGGGTCCAGATATTCGTATCTGTCGAGTAGAGTGTGGGCTGCGCTAGCAAGTTCAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO:1)

fgRNA2:

NNNNNNNNNNNNNNNNNNNNGTTTGAGAGCTAGAAATAGCAAGTTCAAATAAGGCTAGTCCGTTATCAACTTGCGCAGACGGTCGGGTCCAGATATTCGTATCTGTCGAGTAGAGTGTGGGCTGCGCAAGTGGCACCGAGTCGGTGC (SEQ ID NO:2)

fgRNA3:

NNNNNNNNNNNNNNNNNNNNGTTTGAGAGCTAGAAATAGCAAGTTCAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGCGCAGACGGTCGGGTCCAGATATTCGTATCTGTCGAGTAGAGTGTGGGCTGCGC (SEQ ID NO:3)

[0017] The sequence in **bold** font is the 20nt gRNA target sequence. This sequence will vary depending on the particular application. The underlined sequence is a broccoli aptamer sequence. Editing sequences can be placed directly on either side of the full fgRNA sequence. If no editing is needed, the transcriptional start is preferably placed as close to the beginning of the fgRNA sequence as possible, and the terminator is preferably located immediately after the fgRNA sequence.

[0018] Eukaryotes utilize three nuclear enzymes, Pol I, II, and III, to synthesize different classes of RNA, of which the Pol II transcription machinery is the most complex. RNA polymerase II promoters, which constitute the majority of the characterized promoters, cannot be directly used for gRNA production *in vivo* because the 5' CAP, possible 5'/3' UTR, and the poly A tail may hinder guide RNA function. Additionally, RNAs transcribed by RNA polymerase II are rapidly exported from the nuclei into the cytosol while nuclear localization is required for the CAS9/gRNA duplex to access the genomic DNA. Thus, the RNAs transcribed by RNA polymerase II are physically separated from the intended targets that are located in the nucleus. Consequently, guide RNAs for *in vivo* assays are usually expressed under promoters transcribed by RNA polymerase III (e.g., U6 promoter).

[0019] As described herein, however, fluorescent guide RNAs (fgRNAs) can be expressed under a RNA Pol II promoter such as a tetracycline response element (TRE) promoter using any number of different mRNA editing strategies. While gRNA is typically produced by RNA Polymerase III (Pol III) in eukaryotic cells, editing mRNA into fgRNA allows the production of fgRNA from RNA Polymerase II (Pol II) promoters as well. This greatly expands the contexts in which this technology can be utilized.

[0020] In some cases, CRISPR-responsive promoters are used. As used herein, the term "CRISPR-responsive promoter" encompasses eukaryotic promoters as well as synthetic gene regulatory devices and circuits for regulated gene expression. In some embodiments, a CRISPR-responsive promoter comprises a RNA Pol II promoter or an RNA Pol II promoter. The CRISPR-responsive promoter can be a CRISPR-repressible promoter (CRP) or a CRISPR-activatable promoter (CAP). A promoter, generally, is a region of nucleic acid that initiates transcription of a nucleic acid encoding a product. A promoter may be located upstream (e.g., 0 bp to −100 bp, −30 bp, −75 bp, or −90 bp) from the transcriptional start site of a nucleic acid encoding a product, or a transcription start site may be located within a promoter. A promoter may have a length of 100-1000 nucleotide base pairs, or 50-2000 nucleotide base pairs. In some embodiments, promoters have a length of at least 2 kilobases (e.g., 2-5 kb, 2-4 kb, or 2-3 kb).

[0021] In certain embodiments, it may be advantageous to use a cell type-specific promoter to drive expression of a fgRNA provided herein. For example, a CRISPR-based genetic circuit may be configured to comprise a cell type-specific promoter in place of a synthetic promoter to drive fgRNA expression in a spatially controlled manner.

[0022] In some cases, the self-cleaving ribozyme is Csy4, a CRISPR-associated endonuclease which recognizes and cleaves a unique 28-nt RNA sequence. When this sequence is placed on either side of a gRNA within an mRNA transcript and expressed, Csy4 efficiently cleaves gRNAs sandwiched between 28-nt Csy4 recognition sites to produce a functional gRNA. If Csy4 is not expressed, the gRNA is not released, adding temporal and/or spatial control to the system. In some cases, the Csy4 endonuclease is derived from *Pseudomonas aeruginosa*.

[0023] In other cases, the RGR motif is composed of a gRNA flanked by a 5' Hammerhead (HH) type ribozyme (Pley et al., *Nature* 372, 1994), which cleaves directly before the first nucleotide of the gRNA, and a 3' Herpes Delta Virus (HDV) ribozyme (Ferre-D'Amare et al., *Nature* 395: 567-74, 1998), which cleaves directly after the last gRNA nucleotide. When a gRNA sequence is flanked by these two self-cleaving ribozymes, the resulting gRNA is nearly indistinguishable from those produced from a Pol III promoter. By way of example, an RNA molecule can be designed to comprise a HH type ribozyme at the 5′-end, a fgRNA as described herein, and a HDV ribozyme at the 3′-end.

[0024] tRNA-mediated editing is achieved by inserting tRNA sequences into the mRNA transcript on either side of the gRNA. Endogenous tRNA cleaving enzymes RNase P and RNase Z cleave on the 5' and 3' sides of the tRNA, respectively. In some cases, flanking editing sequences are included on either side of a fgRNA.

[0025] Referring to FIG. 1b, three different RNA editing strategies were assayed: editing by the Csy4 protein (CGC), ribozyme-guide-ribozyme (RGR) motif, and endogenous tRNA machinery (TGT). Using fgRNA1 (see FIG. 1A), we compared the efficiency of repression by U6 driven fgRNAs edited by the Csy4 protein (CGC), ribozyme-guide-ribozyme (RGR) motif, or endogenous tRNA machinery (TGT). Our data demonstrate that gRNAs can be efficiently produced *in vitro* and *in vivo* from essentially any promoter when the primary transcripts are flanked by self-cleaving ribozymes. Referring to FIG. 1c, repression of Infrared Fluorescent Protein (iRFP) by TRE driven RGR-edited, Csy4-edited, and tRNA-edited fgRNAs was detected, but the level of expression and efficiency of repression was lower than that of a U6-driven fgRNA construct.

[1] fgRNAs described herein can be used with any CRISPR-based system. CRISPR systems belong to different classes, with different repeat patterns, sets of genes, and species ranges. A CRISPR enzyme is typically a type I or III CRISPR enzyme. In some cases, a fgRNA

is used with a CRISPR system derived from a type II CRISPR system. The type II CRISPR enzyme may be any Cas enzyme. The terms "Cas" and "CRISPR-associated Cas" are used interchangeably herein. The Cas enzyme can be any naturally-occurring nuclease as well as any chimeras, mutants, homologs, or orthologs. In some embodiments, one or more elements of a CRISPR system is derived from a particular organism comprising an endogenous CRISPR system, such as *Streptococcus pyogenes* (SP) CRISPR systems or *Staphylococcus aureus* (SA) CRISPR systems. The CRISPR system is a type II CRISPR system and the Cas enzyme is Cas9 or a catalytically inactive Cas9 (dCas9). Other non-limiting examples of Cas proteins include Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (also known as Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, homologues thereof, or modified versions thereof. A comprehensive review of the Cas protein family is presented in Haft et al. (2005) Computational Biology, *PLoS Comput. Biol.* 1:e60. At least 41 CRISPR-associated (Cas) gene families have been described to date.

[0026] It will be understood that the CRISPR-Cas system for use with a fgRNA described herein is non-naturally occurring in a cell, i.e. engineered or exogenous to the cell. The CRISPR-Cas system as referred to herein has been introduced in a cell. Methods for introducing the CRISPR-Cas system in a cell are known in the art. In some cases, the cell comprises one or more engineered nucleic acid molecules encoding individual components of the CRISPR-Cas system, which are expressed in the cell to establish a functional CRISPR complex, capable of modifying (such as cleaving) a target DNA sequence.

[0027] So that the methods and systems provided herein may more readily be understood, certain terms are defined:

[0028] As used in this specification and the appended claims, the singular forms "a," "an," and "the" include plural references unless the context clearly dictates otherwise. Any reference to "or" herein is intended to encompass "and/or" unless otherwise stated.

[0029] The terms "comprising", "comprises" and "comprised of as used herein are synonymous with "including", "includes" or "containing", "contains", and are inclusive or open-ended and do not exclude additional, non-recited members, elements, or method steps. The phraseology and terminology used herein is for the purpose of description and should not be regarded

as limiting. The use of "including," "comprising," "having," "containing," "involving," and variations thereof, is meant to encompass the items listed thereafter and additional items. Use of ordinal terms such as "first," "second," "third," etc., in the claims to modify a claim element does not by itself connote any priority, precedence, or order of one claim element over another or the temporal order in which acts of a method are performed. Ordinal terms are used merely as labels to distinguish one claim element having a certain name from another element having a same name (but for use of the ordinal term), to distinguish the claim elements.

[0030] As used herein, the terms “synthetic" and "engineered" are used interchangeably and refer to the aspect of having been manipulated by the hand of man.

[0031] The terms “nucleic acid” and “nucleic acid molecule,” as used herein, refer to a compound comprising a nucleobase and an acidic moiety, e.g., a nucleoside, a nucleotide, or a polymer of nucleotides. Typically, polymeric nucleic acids, e.g., nucleic acid molecules comprising three or more nucleotides are linear molecules, in which adjacent nucleotides are linked to each other via a phosphodiester linkage. In some embodiments, “nucleic acid” refers to individual nucleic acid residues (e.g. nucleotides and/or nucleosides). In some embodiments, “nucleic acid” refers to an oligonucleotide chain comprising three or more individual nucleotide residues. As used herein, the terms “oligonucleotide” and “polynucleotide” can be used interchangeably to refer to a polymer of nucleotides (e.g., a string of at least three nucleotides). In some embodiments, “nucleic acid” encompasses RNA as well as single and/or double-stranded DNA. Nucleic acids may be naturally occurring, for example, in the context of a genome, a transcript, an mRNA, tRNA, rRNA, siRNA, snRNA, a plasmid, cosmid, chromosome, chromatid, or other naturally occurring nucleic acid molecule. On the other hand, a nucleic acid molecule may be a non-naturally occurring molecule, e.g., a recombinant DNA or RNA, an artificial chromosome, an engineered genome, or fragment thereof, or a synthetic DNA, RNA, DNA/RNA hybrid, or include non-naturally occurring nucleotides or nucleosides. Furthermore, the terms “nucleic acid,” “DNA,” “RNA,” and/or similar terms include nucleic acid analogs, i.e. analogs having other than a phosphodiester backbone. Nucleic acids can be purified from natural sources, produced using recombinant expression systems and optionally purified, chemically synthesized, etc. Where appropriate, e.g., in the case of chemically synthesized molecules, nucleic acids can comprise nucleoside analogs such as analogs having chemically modified bases or sugars, and backbone modifications. A nucleic acid sequence is presented in the 5′ to 3′ direction unless otherwise indicated. In some embodiments, a nucleic acid

is or comprises natural nucleosides (e.g. adenosine, thymidine, guanosine, cytidine, uridine, deoxyadenosine, deoxythymidine, deoxyguanosine, and deoxycytidine); nucleoside analogs (e.g., 2-aminoadenosine, 2-thiothymidine, inosine, pyrrolo-pyrimidine, 3-methyl adenosine, 5-methylcytidine, 2-aminoadenosine, C5-bromouridine, C5-fluorouridine, C5-iodouridine, C5-propynyl-uridine, C5-propynyl-cytidine, C5-methylcytidine, 2-aminoadeno sine, 7-deazaadenosine, 7-deazaguanosine, 8-oxoadenosine, 8-oxoguanosine, O(6)-methylguanine, and 2-thiocytidine); chemically modified bases; biologically modified bases (e.g., methylated bases); intercalated bases; modified sugars (e.g., 2′-fluororibose, ribose, 2′-deoxyribose, arabinose, and hexose); and/or modified phosphate groups (e.g., phosphorothioates and 5′-N-phosphoramidite linkages).

[0032] The terms "protein," "peptide," and "polypeptide" are used interchangeably herein and refer to a polymer of amino acid residues linked together by peptide (amide) bonds. The terms refer to a protein, peptide, or polypeptide of any size, structure, or function. Typically, a protein, peptide, or polypeptide will be at least three amino acids long. A protein, peptide, or polypeptide may refer to an individual protein or a collection of proteins. One or more of the amino acids in a protein, peptide, or polypeptide may be modified, for example, by the addition of a chemical entity such as a carbohydrate group, a hydroxyl group, a phosphate group, a farnesyl group, an isofarnesyl group, a fatty acid group, a linker for conjugation, functionalization, or other modification, etc. A protein, peptide, or polypeptide may also be a single molecule or may be a multi-molecular complex. A protein, peptide, or polypeptide may be just a fragment of a naturally occurring protein or peptide. A protein, peptide, or polypeptide may be naturally occurring, recombinant, or synthetic, or any combination thereof. A protein may comprise different domains, for example, a nucleic acid binding domain and a nucleic acid cleavage domain. In some embodiments, a protein comprises a proteinaceous part, e.g., an amino acid sequence constituting a nucleic acid binding domain, and an organic compound, e.g., a compound that can act as a nucleic acid cleavage agent.

[0033] As used herein, "modifying" ("modify") one or more target nucleic acid sequences refers to changing all or a portion of a (one or more) target nucleic acid sequence and includes the cleavage, introduction (insertion), replacement, and/or deletion (removal) of all or a portion of a target nucleic acid sequence. All or a portion of a target nucleic acid sequence can be completely or partially modified using the methods provided herein. For example, modifying a target nucleic acid sequence includes replacing all or a portion of a target nucleic acid sequence with one or more nucleotides (e.g., an exogenous nucleic acid sequence) or removing or deleting all or a portion (e.g.,

one or more nucleotides) of a target nucleic acid sequence. Modifying the one or more target nucleic acid sequences also includes introducing or inserting one or more nucleotides (e.g., an exogenous sequence) into (within) one or more target nucleic acid sequences.

[0034] Unless otherwise defined, all terms used in disclosing the invention, including technical and scientific terms, have the meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. By means of further guidance, term definitions are included to better appreciate the teaching of the present invention.

[0035] The present invention is further illustrated by the following Examples, which in no way should be construed as further limiting. The entire contents of all of the references (including literature references, issued patents, published patent applications, and co-pending patent applications) cited throughout this application are hereby expressly incorporated herein by reference.

EXAMPLES

[0036] The section below describes a RNA construct capable of (i) directing Cas9 activity and (ii) fluorescing in the presence of the small molecule DFHBI-1T. Here, we demonstrate that this construct allows direct, non-invasive measurement of the CRISPR gRNA, which was not previously possible. In addition, we demonstrate expression of the fgRNA construct from RNA Pol III and Pol II promoters.

[0037] We engineered a fluorophore-binding RNA aptamer within gRNAs. We chose "Broccoli" aptamer due to its brightness and short sequence. The small molecule DFHBI-1T permeates cell membranes, binds to Broccoli aptamer, and fluoresces green. A gRNA sequence has regions into which additional sequences can be inserted while retaining its ability to complex with Cas9 and target DNA (e.g., stem loop 1, stem loop 2, and 3' end). We inserted Broccoli into these regions of a gRNA driven by a Pol III (U6) promoter and examined whether this modification altered gRNA functionality by indirectly measuring Cas9/gRNA-mediated repression of an infra-red fluorescent protein (iRFP) driven by our previously reported CRISPR-responsive promoter (CRP).

[0038] After inserting Broccoli into the indicated regions of a gRNA driven by Pol III (U6) promoter, we examined whether these modifications altered gRNA functionality using indirect measurement of Cas9/gRNA-mediated repression of an infra-red fluorescent protein (iRFP) driven

by a CRISPR-responsive promoter (CRP). Flow cytometry data demonstrate that Broccoli containing fluorescent gRNAs are functional to a level comparable with unmodified gRNAs (**FIG. 1a**).

[0039] After observing that the fluorescent gRNA (fgRNA) with Broccoli in stem loop 1 (fgRNA1) functioned best, we further modified this gRNA design with three previously published strategies to edit gRNA post-transcriptionally. We generated fgRNAs flanked by ribozymes (RGR), flanked by the RNA endonuclease Csy4 target sites (CGC), and flanked by tRNA sequence (TGT).

[0040] ***fgRNA Sequences:*** The nucleotide sequences of the three variants of the fgRNA are as follows:

[0041] fgRNA1:

NNNNNNNNNNNNNNNNNNNGTTTGAGAGCTAGCGCAGACGGTCGGGTCCAGATATTCGTATCTGTCGAGTAGAGTGTGGGCTGCGCTAGCAAGTTCAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO:1)

[0042] fgRNA2:

NNNNNNNNNNNNNNNNNNNGTTTGAGAGCTAGAAATAGCAAGTTCAAATAAGGCTAGTCCGTTATCAACTTGCGCAGACGGTCGGGTCCAGATATTCGTATCTGTCGAGTAGAGTGTGGGCTGCGCAAGTGGCACCGAGTCGGTGC (SEQ ID NO:2)

[0043] fgRNA3:

NNNNNNNNNNNNNNNNNNNGTTTGAGAGCTAGAAATAGCAAGTTCAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGCGCAGACGGTCGGGTCCAGATATTCGTATCTGTCGAGTAGAGTGTGGGCTGCGC (SEQ ID NO:3)

[0044] The sequence in **bold** font is the 20nt gRNA target sequence. This sequence will vary depending on the particular application. The underlined sequence is the *broccoli* aptamer sequence. Editing sequences can be placed directly on either side of the full fgRNA sequence. If no editing is needed, the transcriptional start is preferably placed as close to the beginning of the fgRNA sequence as possible, and the terminator is preferably located immediately after the fgRNA sequence.

[0045] To study the functionality of these designs, we repeated flow cytometry experiments, demonstrating that modified fgRNAs are still functional (FIG. 1b). To show the feasibility of our approach in direct assesment of gRNAs expression from Pol II promoters, we

engineered CGC, RGR, and TGT under TRE promoter and examined fgRNA expression (as indicated by Broccoli fluorescence level), and functionality (as indicated by iRFP decrease) by flow cytometry in the presence/absence of doxicycline (FIG. 1c). Our data demontrate that fgRNAs enable direct measurement of gRNAs expression and functionality from Pol II promoters.

[0046] *Data Collection Methods*: HEK293ft cells were transfected with plasmids comprising a simple repressor circuit in amounts shown in Table 1. dCas9 and guide RNA (gRNA) complex to repress the expression of infrared fluorescent protein (iRFP). dCas9 was fused with BFP, functioning as a transfection marker, allowing us to disregard untransfected cells in the analysis. An activator construct co-expressed reverse tetracycline trans-activator (rtTA) and Gal4-VP16. rtTA complexes with Doxycycline to drive TRE (pol II) guide expression, while Gal4-VP16 drives iRFP expression in the absence of repression. Csy4 was added only to conditions which required editing by the Csy4 protein. Total mass was normalized to 600ng using a non-coding, empty plasmid.

Table 1: Mass of plasmids used in transfections.

dCas9-BFP	75
Activator	20
CRP-iRFP	20
Csy4	50*
Repressor	100
Empty	to 600ng

[0047] All transfections were performed with PEI, using standard procedures and a 2:1 PEI:DNA ratio. Cell medium was changed every 24 hours. Flow cytometry was performed after 72 hours using a BD FACSCelesta (Becton Dickson) flow cytometer. Flow cytometry data were analyzed using FlowJo (FlowJo, LLC) software. To analyze the data, cells were first gated by Front Scatter (FSC) and Side Scatter (SSC) to eliminate non-cell events. Identified cells were then gated by BFP in order to remove non-transfected cells. From the remaining population of transfected cells, the geometric mean of BB515 (green fluorescence) and APC-Cy7 (infrared fluorescence) was calculated for 4 replicates of all conditions. These geometric mean values were then used to calculate fold changes between guide/no guide (for Pol III, U6 conditions) or between doxicycline /no doxicycline (for Pol II, TRE conditions). The results are presented in FIGS. 1A-1C and FIGS. 2A-2C.

CLAIMS

What is claimed is:

1. A method of directly detecting and measuring guide RNA (gRNA) expression in a cell, the method comprising

(a) introducing a synthetic regulatory system into the cell, the synthetic regulatory system comprising:

(i) a nucleotide sequence encoding a Cas nuclease; and

(ii) a fluorescent guide RNA (fgRNA) operably linked to a CRISPR-responsive promoter, the fgRNA comprising:

a fluorophore-binding RNA aptamer sequence, and

a 20-nucleotide (nt) spacer sequence that hybridizes with a target gene sequence of a DNA molecule in a eukaryotic cell that contains the DNA molecule; and

(b) contacting the cell comprising the synthetic regulatory system to a small molecule inducer of the fluorophore-binding RNA aptamer sequence; wherein gRNA expression is measured by fluorescence monitoring.

2. The method of claim 1, wherein the CRISPR-responsive promoter is a RNA Pol II promoter.

3. The method of claim 2, wherein the RNA Pol II promoter is doxycycline inducible.

4. The method of claim 2, wherein the fgRNA is configured for ribozyme-mediated editing, Csy4-mediated editing, or tRNA-mediated editing.

5. The method of claim 1, wherein the fluorophore-binding RNA aptamer sequence is selected from the group consisting of a Broccoli aptamer, Spinach aptamer, and Spinach2 aptamer.

6. The method of claim 1, wherein the small molecule inducer is DFHBI-1T.

7. The method of claim 1, wherein the fluorophore-binding RNA aptamer sequence is selected from the group consisting of SEQ ID NO:1. SEQ ID NO:2, and SEQ ID NO:3.

8. A method of monitoring efficiency of CRISPR-mediated gene repression, the method comprising

(a) introducing a synthetic regulatory system into the cell, the synthetic regulatory system comprising:

(i) a nucleotide sequence encoding an inactivated Cas nuclease; and

(ii) a fluorescent guide RNA (fgRNA) operably linked to a CRISPR-responsive promoter, the fgRNA comprising:

a fluorophore-binding RNA aptamer sequence; and

a 20-nucleotide (nt) spacer sequence that hybridizes with a target sequence of a DNA molecule in a eukaryotic cell that contains the DNA molecule; and

(b) contacting the cell comprising the synthetic regulatory system to a small molecule inducer of the fluorophore binding RNA aptamer; wherein gRNA expression is measured by fluorescence monitoring.

9. The method of claim 8, wherein the CRISPR-responsive promoter is a doxycycline inducible RNA Pol II promoter.

10. The method of claim 8, wherein the fgRNA is configured for ribozyme-mediated editing, Csy4-mediated editing, or tRNA-mediated editing.

11. The method of claim 8, wherein the fluorophore-binding RNA aptamer sequence is selected from the group consisting of a Broccoli aptamer, Spinach aptamer, and Spinach2 aptamer.

12. The method of claim 8, wherein the small molecule inducer is DFHBI-1T.

13. The method of claim 8, wherein the fluorophore-binding RNA aptamer sequence is selected from the group consisting of SEQ ID NO:1. SEQ ID NO:2, and SEQ ID NO:3.

14. A synthetic regulatory system comprising:

(i) a nucleotide sequence encoding a Cas nuclease; and

(ii) a fluorescent guide RNA (fgRNA) operably linked to a CRISPR-responsive promoter, the fgRNA comprising:

a fluorophore-binding RNA aptamer sequence, and

a 20-nucleotide (nt) spacer sequence that hybridizes with a target gene sequence of a DNA molecule in a eukaryotic cell that contains the DNA molecule.

15. The system of claim 14, wherein the CRISPR-responsive promoter is a doxycycline inducible RNA Pol II promoter.

16. The system of claim 15, wherein the RNA Pol II promoter is doxycycline inducible.

17. The system of claim 14, wherein the fgRNA is configured for ribozyme-mediated editing, Csy4-mediated editing, or tRNA-mediated editing.

18. The system of claim 14, wherein the fluorophore-binding RNA aptamer sequence is selected from the group consisting of a Broccoli aptamer, Spinach aptamer, and Spinach2 aptamer.

19. The system of claim 14, wherein the small molecule inducer is DFHBI-1T.

20. The system of claim 14, wherein the fluorophore-binding RNA aptamer sequence is selected from the group consisting of SEQ ID NO:1. SEQ ID NO:2, and SEQ ID NO:3.

21. A fluorescent guide RNA (fgRNA) operably linked to a CRISPR-responsive promoter, the fgRNA comprising:

a fluorophore-binding RNA aptamer sequence; and

a 20-nucleotide (nt) spacer sequence that hybridizes with a target sequence of a DNA molecule in a eukaryotic cell that contains the DNA molecule.

22. The fgRNA of claim 21, wherein the fgRNA is configured for ribozyme-mediated editing, Csy4-mediated editing, or tRNA-mediated editing.

23. The fgRNA of claim 21, wherein the fluorophore-binding RNA aptamer sequence is selected from the group consisting of a Broccoli aptamer, Spinach aptamer, and Spinach2 aptamer.

24. The fgRNA of claim 21, wherein the fluorophore-binding RNA aptamer sequence is selected from the group consisting of SEQ ID NO:1. SEQ ID NO:2, and SEQ ID NO:3.

FIG. 1A

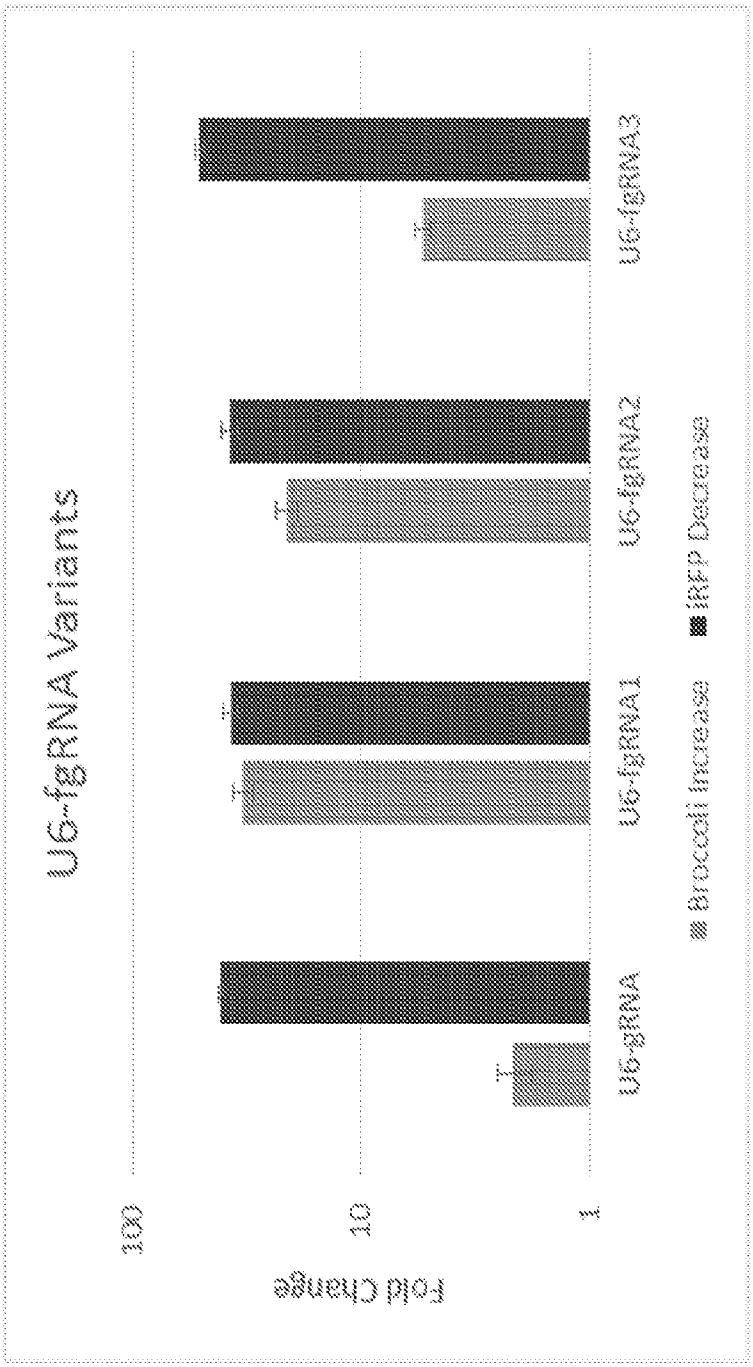

U6-fgRNA Variants
Fold Change
100
10
1
U6-gRNA
U6-fgRNA1
U6-fgRNA2
U6-fgRNA3
Broccoli Increase
iREP Decrease

FIG. 1B

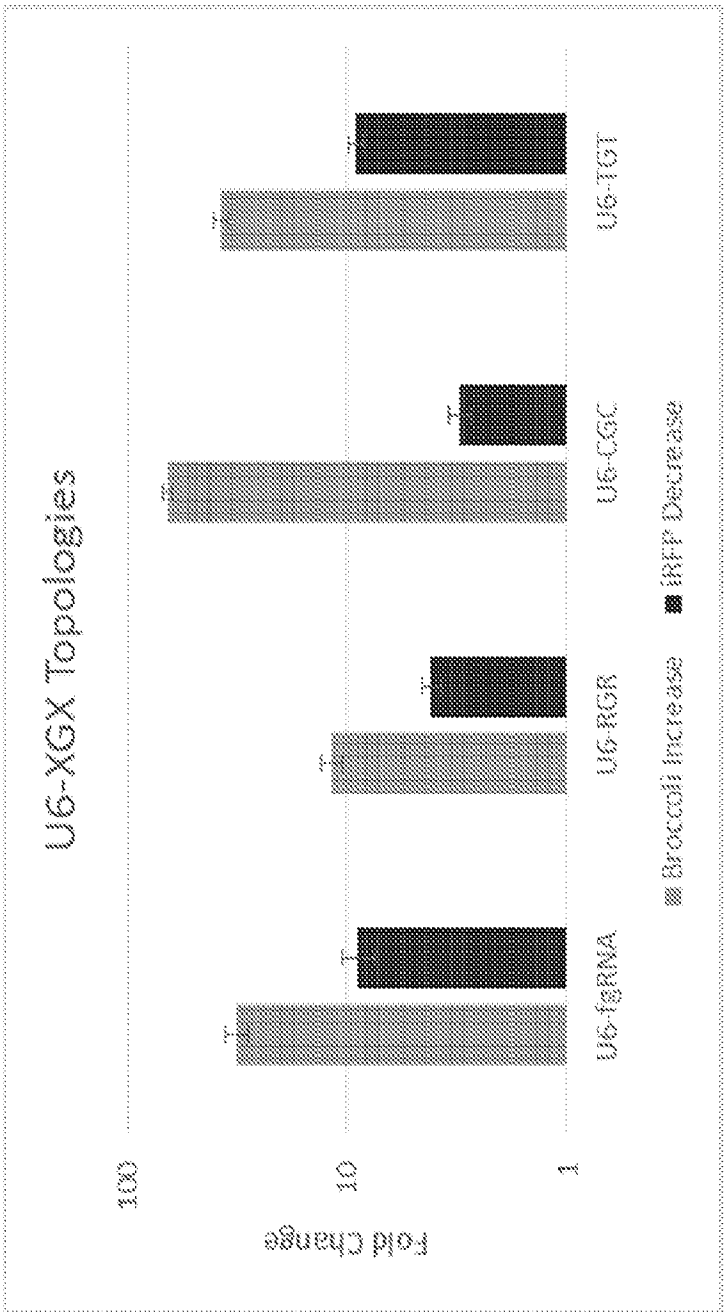

U6-XGX Topologies
Fold change
100
10
1
U6-fgRNA
U6-ROR
U6-CGC
U6-TGT
Broccoli Increase
tRFP Decrease

FIG. 1C

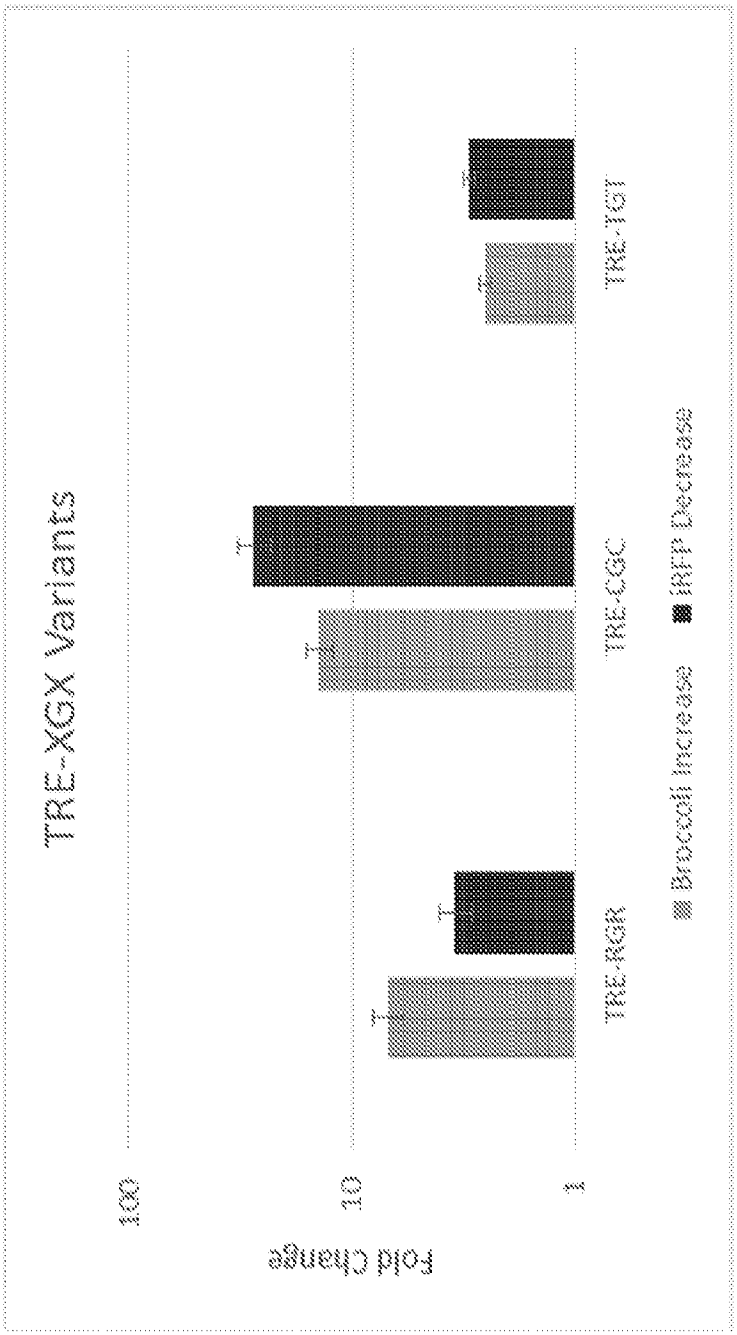

TRE-XGX Variants
fold change
100
10
1
TRE-RGR
TRE-CGC
TRE-TGT
Broccoli Increase
iRFP Decrease

FIG. 2A

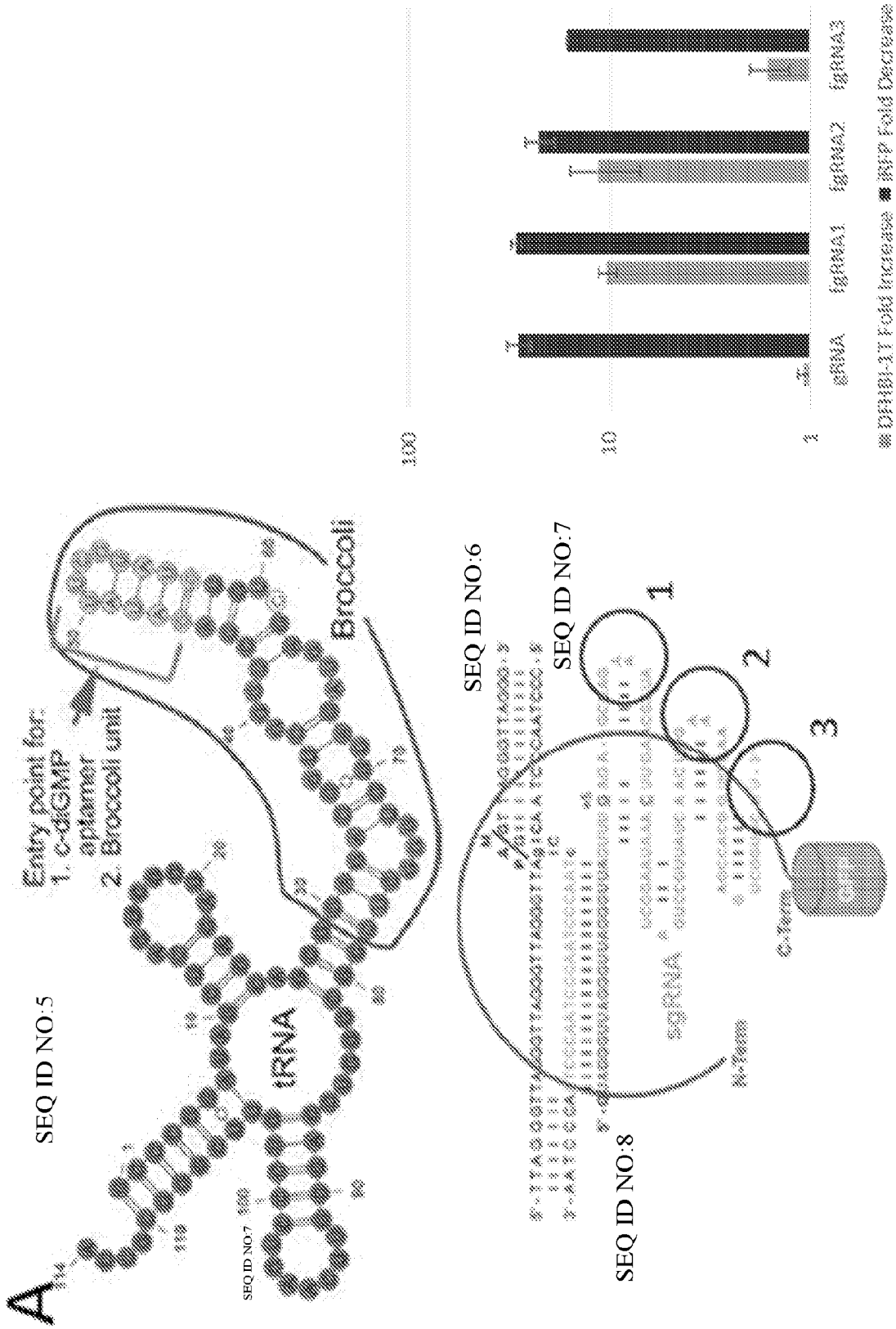

A
SEQ ID NO:5
Entry point for:
1. c-diGMP aptamer
2. Broccoli unit
tRNA
Broccoli
SEQ ID NO:7
SEQ ID NO:6
SEQ ID NO:7
SEQ ID NO:8
sgRNA
N-Term
C-Term
1
2
3
100
10
1
gRNA
tgRNA1
tgRNA2
tgRNA3
DHB1-3T Fold Increase
BFP Fold Decrease

FIG. 2B

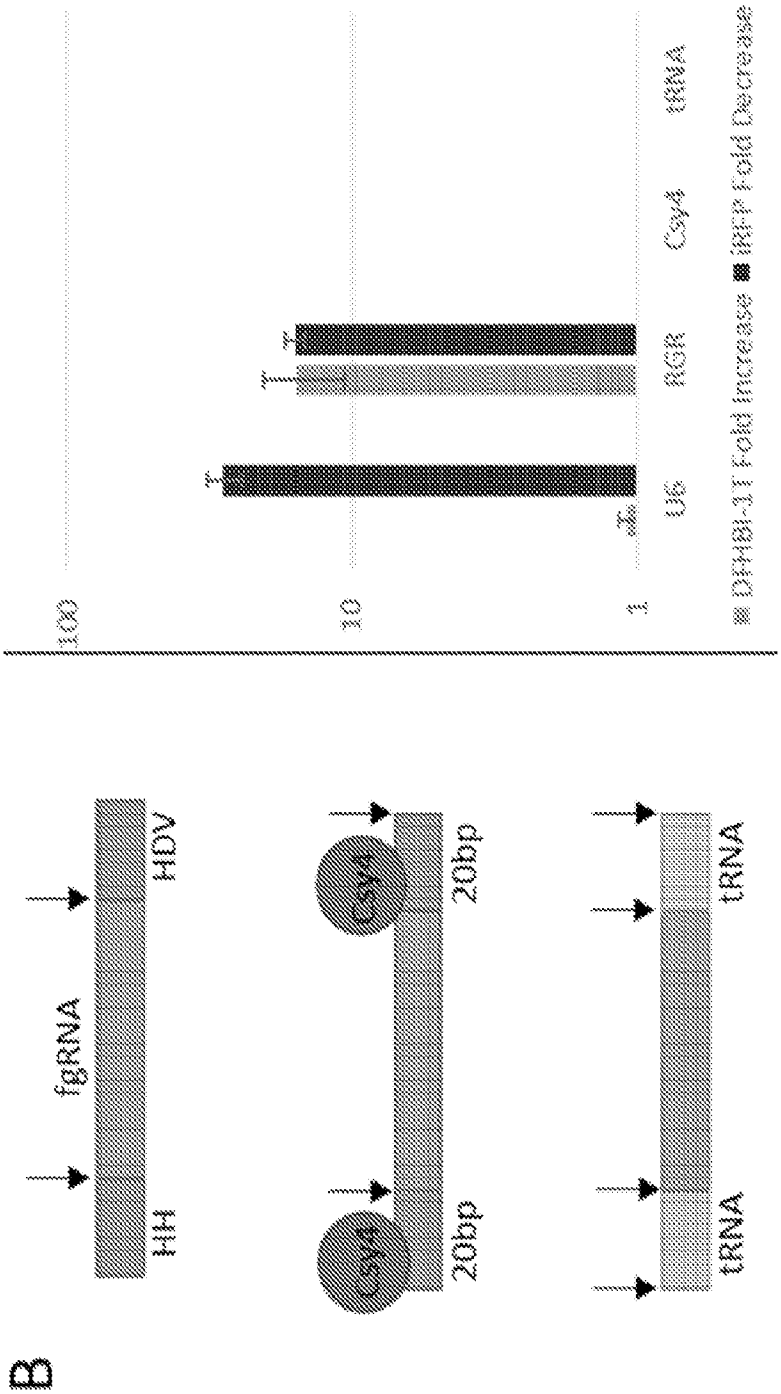
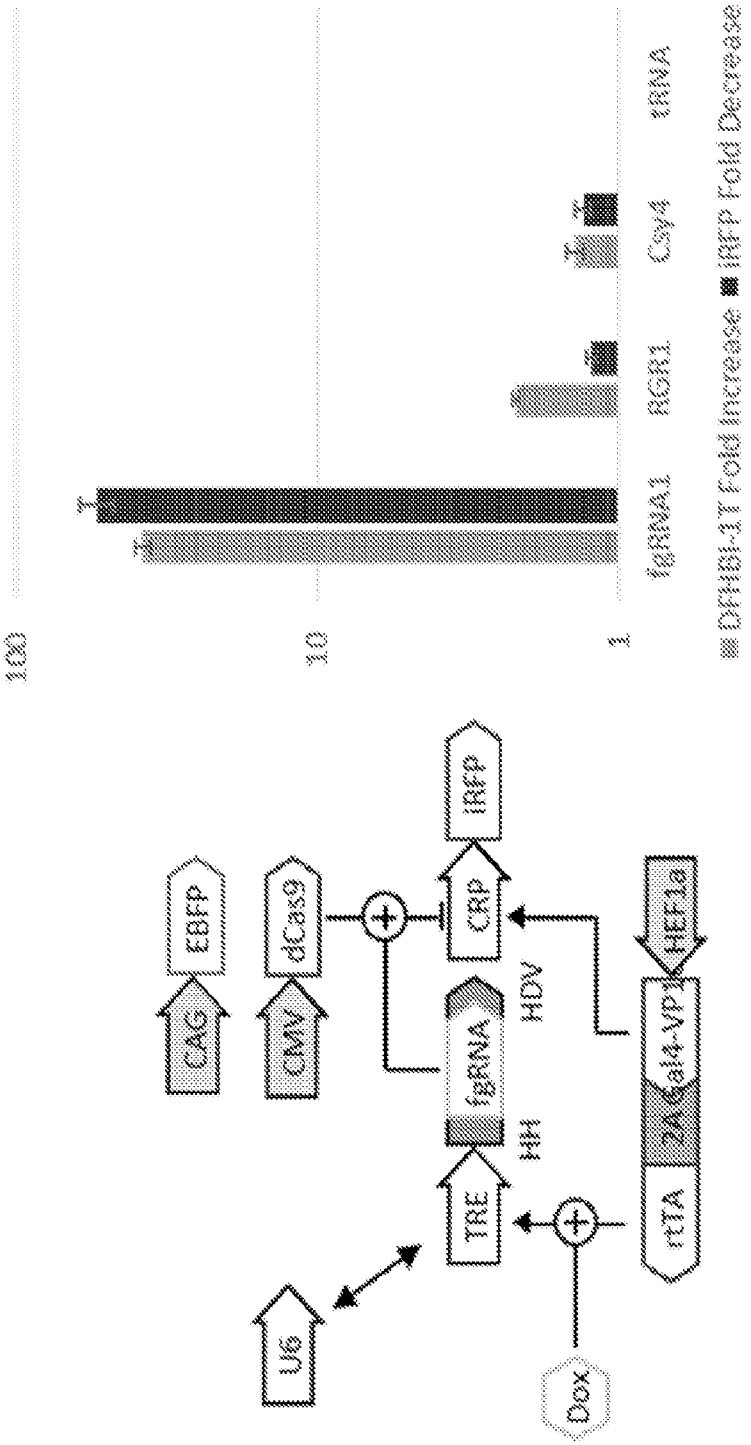

B
HH
fgRNA
HDV
Csy4
20bp
Csy4
20bp
tRNA
tRNA
100
10
1
DHBI-1T Fold Increase
tRNA
Csy4
RGR
U6
tRNA
Csy4
RGR
U6
iRFP Fold Decrease

FIG. 2C


100
10
1
fgRNA1
RGR1
Csy4
tRNA
DHBI-1T Fold Increase
iRFP Fold Decrease
U6
TRE
HH
fgRNA
HDV
CAG
EBFP
CMV
dCas9
CRP
iRFP
rtTA
2A
Gal4-VP1
HEF1a
Dox

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2017/054271

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C07H 21/02; C12N 15/00; C12N 15/113; C12N 15/63; C12N 15/85; C12Q 1/02 (2017.01)
CPC - C07H 1/00; C07H 21/02; C12N 15/113; C12N 15/63; C12N 15/85; C12N 2740/16043 (2017.08)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC - 435/252.33; 435/320.1; 436/501; 506/9; 506/16; 536/23.1 (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2016/183402 A2 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 17 November 2016 (17.11.2016) entire document	1, 2, 4-6, 8, 10-12, 14, 17-19, 21-23
--		
Y		3, 9, 15, 16
Y	US 2016/0208243 A1 (THE BROAD INSTITUTE, INC. et al) 21 July 2016 (21.07.2016) entire document	3, 9, 15, 16
A	US 2015/0141282 A1 (CORNELL UNIVERSITY) 21 May 2015 (21.05.2015) entire document	1-24
A	WO 2016/054106 A1 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 07 April 2016 (07.04. 2016) entire document	1-24
A	US 2017/0022499 A1 (MASSACHUSETTS INSTITUTE OF TECHNOLOGY), 26 January 2017 (26.01.2017) entire document	1-24
A	SHECHNER et al. "CRISPR Display: A Modular Method for Locus-Specific Targeting of Long Noncoding RNAs and Synthetic RNA Devices in Vivo," Nature Methods, 01 July 2015 (01.07.2015), Vol. 12, Pgs. 664-670. entire document	1-24

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
01 December 2017	27 DEC 2017
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300	Authorized officer Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (January 2015)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2017/054271

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. ☒ forming part of the international application as filed:

      ☒ in the form of an Annex C/ST.25 text file.

      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☐ furnished subsequent to the international filing date for the purposes of international search only:

      ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

SEQ ID NOs: 1-8 were searched.