



(51) International Patent Classification:

C12N 15/11 (2006.01) G06F 17/16 (2006.01)
C12N 15/50 (2006.01) G06F 17/40 (2006.01)
C12Q 1/6809 (2018.01)

(21) International Application Number:

PCT/US2021/045790

(22) International Filing Date:

12 August 2021 (12.08.2021)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/064,446 12 August 2020 (12.08.2020) US

(71) Applicant: SPERATUM BIOPHARMA, INC. [US/US];

9 E Lockerman Street, Suite 331, Dover, DE 19901 (US).

(72) Inventors: MARIN-MULLER, Christian, Roberto; 9

E Lockerman Street, Suite 331, Dover, DE 19901 (US).

MARTINEZ, Osvaldo, Vega; 9 E Lockerman Street, Suite 331, Dover, DE 19901 (US). VALVERDE-HERNANDEZ, Juan, Carlos; 9 E Lockerman Street, Suite 331, Dover, DE 19901 (US).

(74) Agent: BOCHNER, Andrew, D.; Bochner IP PLLC, 295 Madison Ave., 12th Fl., New York, NY 10017 (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, IT, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(54) Title: SUBSTRATE SEQUENCES FOR THE RNAi-MEDIATED REGULATION OF GENOMIC AND SUB-GENOMIC VIRAL RNAs

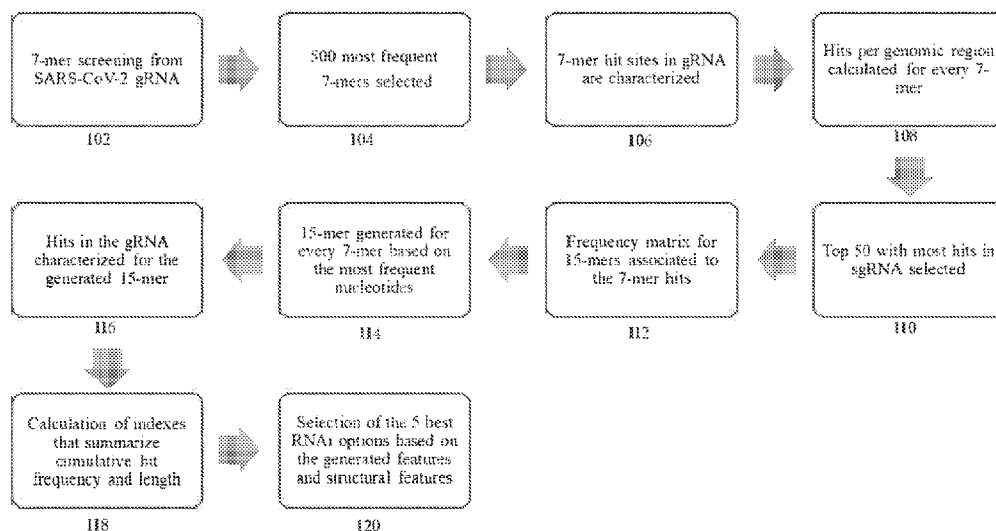


FIG.1

(57) Abstract: Provided is a method for identifying an RNAi sequence, comprising the steps of screening for one or more 7-mers from a gRNA; selecting a first number of the one or more 7-mers; characterizing one or more hit sites for each of the one or more 7-mers; calculating a hit per genomic region for each of the 7-mers; selecting a second number of the 7-mers; creating a frequency matrix for one or more 15-mers associated to each of the 7-mers; generating one or more generated 15-mers for each of the 7-mers; characterizing hits in the gRNA for each of the one or more generated 15-mers; calculating a summary index based on a cumulative hit frequency and a length; and selecting the RNAi sequences based on one or more generated features and one or more structural features.



(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).

Published:

— *with international search report (Art. 21(3))*

Substrate Sequences for the RNAi-Mediated Regulation of Genomic and Sub-genomic Viral RNAs

CLAIM OF PRIORITY

[0001] This application claims priority from U.S. Provisional Patent Application No. 63/064,446, filed on August 12, 2020, the contents of which are incorporated herein by reference.

FIELD OF THE INVENTION

[0001] The invention is related generally to the field of RNAi sequences. More particularly, the invention relates to apparatuses, methods, and systems for RNAi sequence identification.

INTRODUCTION

[0002] Coronaviruses, and in particular SARS-CoV-2, have caused a viral outbreak leading to a worldwide pandemic of Covid-19 illness. Current circumstances globally have resulted in a worldwide public health emergency.

[0003] As a human pathogen, according to the World Health Organization (WHO), as of August 2020, over eighteen (18) million cases of Covid-19 have been confirmed globally, with at least 687,000 deaths. By August 2021, the number of global cases and deaths has risen to over two hundred (200) million cases, with at least four (4) million deaths.

[0004] It is now known and understood that SARS-CoV-2 is readily transmitted from human to human, spreading to multiple continents and leading to the WHO's declaration of a Public Health Emergency of International Concern (PHEIC) on 30 January 2020.

[0005] While scientists have made progress in developing possible new treatments, efficacy and safety of such highly-experimental treatments remain unproven, and many questions

remain unanswered. Though vaccines or other antiviral treatments for SARS-CoV-2 are of the greatest necessity, still many lack effectiveness.

[0006] While efficacy is critical to treating the SARS-CoV-2 pandemic, such treatments must also be shown to be safe. Indeed, smaller safety trials must be run to demonstrate safety before engaging in widespread efficacy trials, resulting in additional time.

[0007] It would be desirable, therefore, to provide treatments with increased chance of efficacy and safety, in order to maximize the likelihood of success and potential time to approval for treatment of SARS-CoV-2.

SUMMARY

[0008] In an embodiment, the invention of the present disclosure may be a method for identifying an RNAi sequence, comprising the steps of identifying a potential 7-mer sequence; and registering one or more hit sites, where the one or more hit sites are one or more positions in a genomic RNA sequence where the potential 7-mer sequence is present. In an embodiment, the method may further comprise the steps of registering, alongside with the one or more hit sites, a name of the genomic region where the 7-mer sequence was found and a registered 15-mer sequence starting at a 7-mer sequence start position; generating, using the registered 15-mer sequence, a nucleotide position frequency matrix and a generated 15-mer for each 7-mer sequence based on the most frequent nucleotide for each position; and predicting, using the generated 15-mer sequence, one or more hits for a putative RNAi molecule. In an embodiment, the method may also comprise the steps of calculating a hit feasibility index for the one or more registered hit sites; generating a general index; and developing a summary matrix.

[0009] In an embodiment, the potential 7-mer sequence may be identified by screening for the most 7-mers in a target RNA. In an embodiment, the one or more hits for the putative RNAi molecule may be predicted based on perfect matching. In an embodiment, the hit feasibility index may assign a weight to each hit site based on a match sequence ΔG value and an AU content proportion 30 nucleotides upstream and downstream from the hit site. The hit feasibility index may be calculated via the following equation: $H_{nm} = \log_3(p(AU)_{nm} \times -\Delta G_{nm})$, where n is the generated 15-mer used to predict a m hit site, H is the hit feasibility, ΔG is the free

energy required for the match to happen considering only base pairing, and $p(\text{AU})$ is the proportion 30 nt upstream and downstream from the hit site.

[0010] In an embodiment, the general index may be calculated by the following equation: $IG_m = \sum_{k=1} (\sum_{m=1} H_{nm})$, where IG is the general index which summarizes the effect of the hits, and A is the amount of hits for the n generated 15-mer of the match. The summary matrix may comprise the 7-mer, the 7-mer's generated 15-mer, a sense strand, the RNAi proposed sequence, the hit feasibility index, and the general index. In a further embodiment, the summary matrix may comprise a proposed antisense strand. In yet another embodiment, the summary matrix may further comprise a seed sequence GC content and a guide strand GC content.

[0011] In an embodiment, the invention of the present disclosure may be a method for identifying an RNAi sequence, comprising the steps of screening for one or more 7-mers from a gRNA; selecting a first number of the one or more 7-mers, the first number based on the frequency of the one or more 7-mers; and characterizing one or more hit sites for each of the one or more 7-mers. The method may further comprise the steps of calculating a hit per genomic region for each of the one or more 7-mers; selecting a second number of the one or more 7-mers, the second number based on the most hits in a sgRNA; and creating a frequency matrix for one or more 15-mers associated to each of the one or more 7-mers. In a further embodiment, the method also includes the steps of generating, based on the most frequent nucleotides, one or more generated 15-mers for each of the one or more 7-mers; characterizing hits in the gRNA for each of the one or more generated 15-mers; calculating a summary index based on a cumulative hit frequency and a length; and selecting a third number of the RNAi sequences based on one or more generated features and one or more structural features.

[0012] In an embodiment, the first number, the second number, and the third number may be any appropriate quantities. In an embodiment, the first number may be 500. In an embodiment, the second number may be 50. In an embodiment, the third number may be 5. In an embodiment, the one or more generated features and the one or more structural features may be a function of the summary index.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] FIG. 1 illustrates an embodiment of an RNAi design workflow, in accordance with the invention.

[0014] FIG. 2A is a table depicting five RNAi SARS-CoV-2 silencing sequences.

[0015] FIG. 2B is a table depicting the number of matching sites of RNAi SARS-CoV-2 silencing sequences.

[0016] FIG. 3A is a table of sequences that may attack SARS-CoV-2 RNA.

[0017] FIG. 3B is a depiction of 3'UTR (untranslated region) of the SARS-CoV-2 RNA.

[0018] FIG. 4 is a graph of NPM-SC2 sequences 3'UTR regulation.

[0019] FIG. 5 is a graph of NPM-SC2 sequences cytotoxicity.

DETAILED DESCRIPTION

[0020] Provided herein are novel treatments for SARS-CoV-2. In an embodiment, ribonucleic acid (RNA), such as RNA-interference (RNAi) molecules are utilized. The RNAi molecules may be used for the downregulation of viral RNA translation.

[0021] Small-interfering RNAs (siRNAs) may be designed to bind to specific target sequences. Off-target effects may occur through non-specific binding, and present a limitation to siRNA design. MicroRNAs (miRNAs) are naturally occurring RNAi molecules. They differ from siRNAs in that a miRNA may have evolved to target multiple different targets simultaneously, through virtue of imperfect pairing. This imperfect pairing leads to differential regulatory responses that differ from the canonical siRNA regulation. In accordance with an embodiment, the RNAi molecules described here are designed as synthetic molecules in a manner similar to siRNAs, but taking into account the permissibility of imperfect pairing as with a miRNA. Thus, the antiviral oligonucleotides generated provide broader, more flexible targeting, and can anticipate off-target effects.

[0022] SARS-CoV-2 is found in the order *Nidoviridae*, due to its unique, nested replication strategy of 30kb or similarly-sized single-stranded viral RNA is processed into many smaller messenger RNAs (mRNAs), each encoding multiple proteins. The SARS-CoV-2 mRNAs contain a 3'UTR, the targeting sites remain intact even once the nested RNA replication

process has begun. Thus, in accordance with the invention, an RNAi-based attack against viral RNA is proposed, to target such nested sequences.

[0023] In accordance with the invention, proposed are siRNA and miRNA sequences including, but not limited to, one or more of: strand-specific RNA sequences (ssRNA); non-canonical processing; no passenger strands; target minimization in human genome in order to maximize targets in viral genome; administrable with any suitable nucleic acid delivery vehicle; elevated silencing rates, short processing routes to bioactive states; multiple targeting sites to reduce genetic variability between strains; simplified production and isolation/purification in comparison to standard duplexes, and insertable into alternative siRNA construct design.

[0024] In accordance with principles of the invention, a genomic understanding of SARS-CoV-2 may be used to develop drugs and treatments tailored to treating the novel coronavirus associated with Covid-19. Thus, viral RNA of SARS-CoV-2 may be targeted with an anti-SARS-CoV-2 treatment: siRNA.

[0025] In an embodiment, the siRNA is specifically designed with the ability and potential to act on multiple target sites on the RNA of SARS-CoV-2. In an embodiment, in particular, the siRNA is designed to act on a plurality of sites on (1) the 3'UTR region of mRNA, and (2) the coding region (CDS) of SARS-CoV-2 RNA. Due to its location immediately after the translation termination codon, 3'UTR mRNA may be responsible for regulating gene expression, and are therefore well suited for receiving a therapeutic. By targeting such sites with siRNA, Ago2 mediated-cleavage may occur, which may result in the destruction of viral RNA prior to ribosome entry, thereby resulting in reduction of viral replication of SARS-CoV-2. Alternatively, inhibition of viral RNA translation may occur, through repression of ribosome drop-off or sequestering of the viral RNA, all through interaction with the target sequences.

[0026] RNA viruses are likely to mutate, and the use of RNAi-inducing molecules against such viruses may result in imperfect matches, due to the specific complementarity needed for siRNA activity. Moreover, siRNAs tend to require integrity of a relatively long sequence, which may make mutations even more cumbersome, in order to appropriately match with an intended target RNA.

[0027] Accordingly, the invention contemplates addressing likely mutations and requirements for high complementarity of siRNA through the generation of a synthetic oligonucleotide that behaves like a miRNA. Often, miRNA is less specific, and more suitably, may match with a greater number of match sites per target RNA. As such, due to increased promiscuity for regulatory target sites, miRNA regulation to specific RNA targets may be better suited for genetic variability and mutation, as compared to siRNA targeting. Due to the presence of only one match site, siRNA complementary site deletion eliminates the siRNA's silencing effect on the target RNA, increasing the likelihood of mutation destroying its regulation. In contrast, miRNAs, with multiple complementary sites on target, are capable of maintaining regulation even when one match site is deleted or changed. In an embodiment, however, a single target site with the right sequence may be sufficient for effective regulation.

[0028] Therefore, in an embodiment, a proprietary RNAi is provided, specifically formulated to increase silencing specificity while maximizing the number of complementary sites in the selected RNA target. As such, RNA-based viruses are able to be treated effectively, even with their high mutation rates. Accordingly, in accordance with numerous embodiments, the invention provides systems and methods identifying potential targeting sites, and provides systems and methods for designing target sequences, in order to maximize antiviral targeting, while decreasing the risk of mutation and target destruction. In some embodiments, combinations of RNAi molecules can be used, with either miRNA or siRNA-designed molecules, or may be used alone. The RNAi molecules may target any or multiple portions of the viral genome.

[0029] Referring to FIG. 1, FIG. 1 illustrates a non-limiting example of an RNAi design workflow, in accordance with the invention. At a first step 102, identification of potential RNAi sequences is performed (for example, the 7-mer screening from SARS-CoV-2 gRNA). In a second step 104, a potential 7-nucleotide seed sequence may be searched for, by searching the most frequent 7-mers in the target RNA (for example, the 500 most frequent 7-mers selected). As a non-limiting example, 5'UTR and CDS may be evaluated, due to the RNA-induced silencing process causing a silencing of transcripts through regulation in 5'UTR and CDS. In an embodiment, after compiling the 7' mers, in step 104, the 500 most frequent ones may be selected for further analysis.

[0030] In a next step 106, every position in the genomic RNA sequence where the selected 7-mers are present, known as the hit site, may be registered alongside with the position, the name of the genomic region where the 7-mer was found, and the 15-mer starting at the 7-mer start position. For example, the 7-mer hit sites in gRNA may be characterized. Fifty 7-mers with the most hits regions may be selected. In step 108, the hits per genomic region may be calculated for every 7-mer. In step 110, the top fifty with the most hits in sgRNA may be selected.

[0031] In step 112, using the registered 15-mers, a nucleotide position frequency matrix may be generated, and a 15-mer may be generated. In step 114, each 7-mer is based on the most frequent nucleotide for each position, with the sequence corresponding to the proposed sense strand for the RNAi. Accordingly, non-canonical base pairing and beyond seed sequence pairing may be promoted.

[0032] In step 116, hits of generated 15-mer sequences may be predicted for the presumed RNAi molecule based on perfect matching. Thus, perfect seed sequence and beyond seed sequence matches may be searched. For example, matches from 6 to 22 nucleotides long where registered. Although a large body of research supports and validates the effect of 6,7 and 8 nucleotide long matches, shorter RNAi matches with transcripts such as 5 nucleotide long match and larger matches such as 10 and 15 matches may be shown to be involved in the silencing of RNA molecules.

[0033] In step 118, for every registered hit site, a *hit feasibility* index may be calculated. A weight may be assigned to every hit, based on its match sequence ΔG value and the AU content proportion 30 nt upstream and downstream from the hit site. A calculation may be performed of indexes that summarize cumulative hit frequency and length. In step 120, a selection of the 5 best RNAi options may be made based on the generated features and structural features.

[0034] Thus, in those embodiments where the structure of the seed sequence is highly responsible for its target silencing capacities, lower ΔG seed sequence pairing values tend to correlate with higher suppression levels.

[0035] Where higher AU contents are found near the RNAi binding sites, this may be

correlated with more transcript and protein silencing, alluding to usually a less stable secondary structure.

$$H_{n_m} = (p(AU)_{n_m} \times -\Delta G_{n_m})$$

[0036] Disclosed above is an exemplary equation where n is the generated 15-mer used to predict the m hit site, H is the *hit feasibility*, ΔG is the free energy required for the match to happen considering only base pairing and is the $p(AU)$ proportion 30 nt upstream and downstream from the hit site. Next, two indexes may be calculated for the generated 15-mers:

$$IG_m = \sum_{k=1}^{15} \left(\sum_{m=1}^A H_{n_m} \right)$$

[0037] In an embodiment, IG is the general index which summarizes the effect of the hits, A is the amount of hits for the n generated 15-mer of the match.

[0038] Referring to step 112, a matrix may be developed featuring: the 7-mer, it's generated 15-mer or proposed antisense strand, the sense or guide strand, the RNAi proposed sequence, the seed sequence GC content, the guide strand GC content, as well as all aforementioned indexes. In an embodiment, human transcripts that may be silenced by seed off-target effects may be searched using the Custom microRNA Prediction functionality of the miRDB platform.

[0039] Referring to FIG. 2A, the table depicts five RNAi SARS-CoV-2 silencing sequences, in accordance with an embodiment

[0040] Referring to FIG. 2B, the table depicts the number of matching sites of RNAi SARS-CoV-2 silencing sequences.

[0041] As disclosed herein, the invention may be modified by chemical modifications, as contemplated. Alternatively, molecular presentation may be modified to administer treatment as an RNAi mimic or as an expression vector, such as a virus, plasmid, or others.

[0042] The various embodiments may be modified in combination with various delivery methods, such as LNP, polymeric nanoparticles, aptamer associated delivery, antibody delivery, affimer associated delivery or metal nanoparticles.

[0043] Further, treatment may be administered either as linear or circular RNA, or as part of a longer non-coding RNA. Treatment may also be administered as a DNA counterpart, including via a plasmid or other system vector, or shRNA vector system.

[0044] The embodiments disclosed herein combine the benefits of siRNA and miRNA to add capacity to have multiple complementarity sites, as in miRNA, while promoting a high specificity similar to siRNA. Additionally, this may address issues of mutation, and may be used to design RNAi molecules to target either single or multiple selected RNA.

[0045] Therefore, the various embodiments may be designed to silence coding and non-coding RNA or silence single-selected RNA or combinations of RNA. Further, regions inside the RNA to be targeted may be enriched for target sites. These embodiments may be implemented clinically or non-clinically, and administered alone or in combination with other compounds.

[0046] In an embodiment, there may be a number of SARS-CoV-2 *in vitro* surrogate expression methods.

[0047] Regarding a cell culture, in an embodiment, ACE2 and TMPRSS2-expressing human colon carcinoma cell-line Caco-2 (ATCC® HTB-37™), human embryonic kidney cell-line HEK 293T (ATCC® CRL-3216™) and grivet kidney cell-line Vero E6 (ATCC® CRL-1586™) may be maintained in DMEM medium (D6429, Sigma-Aldrich) supplemented with 10% fetal bovine serum (F2442, Sigma-Aldrich), antibiotic-antimycotic solution (A5955, Sigma-Aldrich), and 1x of non-essential amino acid solution (M714, Sigma-Aldrich).

[0048] Regarding, plasmid construct design, in an embodiment, for a first approach of the therapeutic potential of the anti SARS-CoV-2 designed miRNA mimics, a Firefly-Renilla luciferases reporter vector and a S-protein/N-protein expression vector may be designed as follows: (a) For the Firefly-Renilla luciferases reporter vector, the SARS-CoV-2 universal 3'UTR sequence may be synthesized and cloned in the 3' end of the Firefly luciferase transcript of the pmirGLO Vector (E1330, Promega) by external services (GenScript). The Renilla luciferase transcript may not include any SARS-CoV-2 elements, being the endogenous control; and (b) For the S-protein/N-protein expression vector, the SARS-CoV-2 S-protein and

N-protein (1) mRNA complete sequences and (2) 3'UTR paring sequences mutated mRNA may be synthesized and cloned in the pHMGFP vector (E6421, Promega) by external services (GenScript). These designs may allow the discrimination between 3'UTR and 5'UTR-CDS regulation capabilities of the anti-SARS-CoV-2 designed miRNA mimics.

[0049] Regarding transfection, in an embodiment, Lipofectamine3000 (L3000015, ThermoFisher) may be used for transient expression of the plasmid constructs on the Caco-2, HEK 293T and Vero E6 cell-lines following manufacturer instructions.

[0050] Regarding a luciferase assay, in an embodiment, the assays may be performed using the Dual-Glo® Luciferase Assay System (E2940, Promega) following manufacturer instructions.

[0051] Regarding eukaryotic expression assays, in an embodiment, specifically RT-qPCR, the total RNA may be extracted from cell-lines by using the mirVana™ miRNA Isolation Kit (AM1561, ThermoFisher) per manufacturer instructions and, RT-qPCR may be carried out by using the SuperScript™ III One-Step RT-PCR System kit (12574026, ThermoFisher) per manufacturer instructions in a Q qPCR machine (Quantabio).

[0052] Regarding a western blot, in an embodiment, the total protein may be extracted from cell-lines by using RIPA Buffer (R0278, Sigma-Aldrich), Protease Inhibitor Tablets (S8820, Sigma-Aldrich) and Protease Inhibitor Cocktail (P8340, Sigma-Aldrich) following the recommended protocol. Total protein quantification may be done by using Pierce™ BCA Protein Assay Kit (23227, Thermo Fisher) per manufacturer instructions.

[0053] Regarding evaluation of viral titer in SARS-CoV-2 *in vitro* model using live virus, in an embodiment, SARS-CoV-2 (accession: NC_045512.2) stock may be cultured following conditions previously described. Virus stocks will be titrated on Vero E6 cells (ATCC® CRL-1586™), in a biosafety level-4 (BSL-4) facility. Virus titration assay may be done following the protocol previously described. Virus titer may be calculated using the formula: $\text{PFU} / \text{mL} = \text{number of plates counted} / (\text{inoculation volume (50 uL in this case)} * \text{sample dilution})$.

[0054] For example, this work may be performed through external services or collaborations. RT-qPCR may be used to confirm viral RNA loads and the specific identity of SARS-CoV-2. Standard and approved methodologies for amplification of viral RNA may be used.

[0055] Regarding the dose-range finding safety study in an outbred murine model, CD-1 mice (CD-1® IGS Mice, Charles River) may be used as experimental subjects. Mice may be grouped according to their weight. The experimental subjects may be treated intravenously every other day with a determined dosing amount in a final injection volume of 200 μ L for one month. Once the treatment regimen is completed, blood parameters including, but not limited to, reference factors for hepatic and kidney functions, may be evaluated as the safety reference. Pathology analysis for various body tissues may be performed by a certified veterinary pathologist.

[0056] Regarding the efficacy study in a transgenic murine or other animal model, in an embodiment, a novel transgenic mouse model may be developed for studying SARS-CoV-2. In such a case, specific pathogen-free, 6-11-month-old, female wild type (WT) C57BL/6J (000664) and, transgenic (TG) K18-hACE2 (034860) mice may be obtained. In an embodiment, two experiments may be run: (a) Nanoparticle formulation delivery: SARS-CoV-2 free WT and TG mice may be dosed intravenously in a q.o.d regimen (M, W, F) with the nanoparticle formulation containing an Anti-SARS-CoV-2 miRNA mimic. After completion of treatment, mice may be sacrificed, and Anti-SARS-CoV-2 miRNA mimic may be quantified on the respiratory system by RT-qPCR; and (b) Efficacy study: TG mice may be inoculated intranasally with SARS-CoV-2 at a dosage of 10^5 TCID₅₀. In an embodiment, after confirmed infection, mice will be separated in two groups: (1) Intravenously dosing of the nanoparticle formulation containing an Anti-SARS-CoV-2 miRNA mimic in a q.o.d regimen (M, W, F) and, (2) Intravenously dosing of the nanoparticle formulation containing a scramble miRNA mimic in a q.o.d regimen (M, W, F). Animals may be continuously observed daily to record body weights, clinical symptoms, responsiveness to external stimuli and death.

[0057] However, if the aforementioned model is not available or an alternative animal model is deemed more appropriate, an adequate alternative research methodology may be proposed.

[0058] Regarding human blood immune response *ex vivo* assay, in an embodiment, to determine whether the formulated nanoparticle induces a human cytokine response, it may be incubated for 24 hours in whole blood collected from healthy volunteers. Anticoagulant may be added to the whole blood to prevent coagulation and the formulated nanoparticle may be tested at three different concentrations. After the incubation, the plasma may be separated by centrifugation and IL-1 β , IL-6, IL-12, INF α , and TNF α may be analyzed by ELISA (KHC0011, BMS223HS, KHC0121, BMS213HS, ThermoFisher).

[0059] Thus, a series of RNAi-inducing siRNA/miRNA molecules may be designed to specifically target the SARS-CoV-2 viral RNA at different locations. The sequences shown in FIG. 3A may have been synthesized, and may have obtained results for the first candidate, labeled as “D”. The scrambled, non-effective sequence “NC” may be used as a “negative control.”

[0060] Sequence “D” may be tested for efficacy against a luciferase reporter construct, designed to include the 3’UTR (untranslated region) of the SARS-CoV-2 RNA, as shown in FIG. 3B.

[0061] 293T cells may be transfected with both the luciferase reporter construct and the antiviral sequence “D,” as described above. Following forty-eight hours, the luciferase expression of the reporter construct may be examined using a plate reader. There may be a significant reduction in luciferase activity following transfection of sequence “D,” demonstrating the regulatory potential and efficacy of the designed sequence. A non-limiting example of the results is shown in FIG. 4.

[0062] Potential cytotoxic effects of sequence “D” in fibroblasts may also be tested. The cells may be transfected with either a positive cytotoxic control, the NC, or sequence “D.” In an embodiment, no cytopathic effects may be observed through MTS assay at 48 hours post

transfection, indicating the relative safety of the sequence. A non-limiting example of the results is shown in FIG. 5.

[0063] In an embodiment, the invention of the present disclosure may be a method for identifying an RNAi sequence, comprising the steps of identifying a potential 7-mer sequence; and registering one or more hit sites, where the one or more hit sites are one or more positions in a genomic RNA sequence where the potential 7-mer sequence is present. In an embodiment, the method may further comprise the steps of registering, alongside with the one or more hit sites, a name of the genomic region where the 7-mer sequence was found and a registered 15-mer sequence starting at a 7-mer sequence start position; generating, using the registered 15-mer sequence, a nucleotide position frequency matrix and a generated 15-mer for each 7-mer sequence based on the most frequent nucleotide for each position; and predicting, using the generated 15-mer sequence, one or more hits for a putative RNAi molecule. In an embodiment, the method may also comprise the steps of calculating a hit feasibility index for the one or more registered hit sites; generating a general index; and developing a summary matrix.

[0064] In an embodiment, the potential 7-mer sequence may be identified by screening for the most 7-mers in a target RNA. In an embodiment, the one or more hits for the putative RNAi molecule may be predicted based on perfect matching. In an embodiment, the hit feasibility index may assign a weight to each hit site based on a match sequence ΔG value and an AU content proportion 30 nucleotides upstream and downstream from the hit site. The hit feasibility index may be calculated via the following equation: $H_{nm} = \log_3(p(AU)_{nm} \times -\Delta G_{nm})$, where n is the generated 15-mer used to predict a m hit site, H is the hit feasibility, ΔG is the free energy required for the match to happen considering only base pairing, and $p(AU)$ is the proportion 30 nt upstream and downstream from the hit site.

[0065] In an embodiment, the general index may be calculated by the following equation: $IG_m = \sum_{k=1}^A (\sum_{m=1}^n H_{nm})$, where IG is the general index which summarizes the effect of the hits, and A is the amount of hits for the n generated 15-mer of the match. The summary matrix may comprise the 7-mer, the 7-mer's generated 15-mer, a sense strand, the RNAi proposed sequence, the hit feasibility index, and the general index. In a further embodiment, the summary matrix

may comprise a proposed antisense strand. In yet another embodiment, the summary matrix may further comprise a seed sequence GC content and a guide strand GC content.

[0066] In an embodiment, the invention of the present disclosure may be a method for identifying an RNAi sequence, comprising the steps of screening for one or more 7-mers from a gRNA; selecting a first number of the one or more 7-mers, the first number based on the frequency of the one or more 7-mers; and characterizing one or more hit sites for each of the one or more 7-mers. The method may further comprise the steps of calculating a hit per genomic region for each of the one or more 7-mers; selecting a second number of the one or more 7-mers, the second number based on the most hits in a sgRNA; and creating a frequency matrix for one or more 15-mers associated to each of the one or more 7-mers. In a further embodiment, the method also includes the steps of generating, based on the most frequent nucleotides, one or more generated 15-mers for each of the one or more 7-mers; characterizing hits in the gRNA for each of the one or more generated 15-mers; calculating a summary index based on a cumulative hit frequency and a length; and selecting a third number of the RNAi sequences based on one or more generated features and one or more structural features.

[0067] In an embodiment, the first number, the second number, and the third number may be any appropriate quantities. In an embodiment, the first number may be 500. In an embodiment, the second number may be 50. In an embodiment, the third number may be 5. In an embodiment, the one or more generated features and the one or more structural features may be a function of the summary index.

[0068] While this invention has been described in conjunction with the embodiments outlined above, many alternatives, modifications and variations will be apparent to those skilled in the art upon reading the foregoing disclosure. For example, the aforementioned embodiments may include specific quantities, chemicals, compounds, genetic material, sequences, time periods, or other quantities. In various embodiments, any of the aforementioned quantities may be non-limiting. Accordingly, the embodiments of the invention, as set forth above, are intended to be illustrative, not limiting. Various changes may be made without departing from the spirit and scope of the invention.

What is claimed is:

1. A method for identifying an RNAi sequence, comprising the steps of:
 - a. identifying a potential 7-mer sequence;
 - b. registering one or more hit sites, wherein the one or more hit sites are one or more positions in a genomic RNA sequence where the potential 7-mer sequence is present;
 - c. registering, alongside with the one or more hit sites, a name of the genomic region where the 7-mer sequence was found and a registered 15-mer sequence starting at a 7-mer sequence start position;
 - d. generating, using the registered 15-mer sequence, a nucleotide position frequency matrix and a generated 15-mer for each 7-mer sequence based on the most frequent nucleotide for each position;
 - e. predicting, using the generated 15-mer sequence, one or more hits for a putative RNAi molecule;
 - f. calculating a hit feasibility index for the one or more registered hit sites;
 - g. generating a general index; and
 - h. developing a summary matrix.
2. The method of claim 1, wherein the potential 7-mer sequence is identified by screening for the most 7-mers in a target RNA.
3. The method of claim 1, wherein the one or more hits for the putative RNAi molecule are predicted based on perfect matching.

4. The method of claim 1, wherein the hit feasibility index assigns a weight to each hit site based on a match sequence ΔG value and an AU content proportion 30 nucleotides upstream and downstream from the hit site.
5. The method of claim 4, wherein the hit feasibility index is calculated via the following equation:

$$H_{n_m} = \log_3(p(AU)_{n_m} \times -\Delta G_{n_m}),$$

wherein n is the generated 15-mer used to predict a m hit site, H is the hit feasibility, ΔG is the free energy required for the match to happen considering only base pairing, and $p(AU)$ is the proportion 30 nt upstream and downstream from the hit site.

6. The method of claim 5, wherein the general index is calculated by the following equation:

$$IG_m = \sum_{k=1} (\sum_{m=1} H_{n_m}),$$

wherein IG is the general index which summarizes the effect of the hits, and A is the amount of hits for the n generated 15-mer of the match.

7. The method of claim 6, wherein the summary matrix comprises the 7-mer, the 7-mer's generated 15-mer, a sense strand, the RNAi proposed sequence, the hit feasibility index, and the general index.
8. The method of claim 7, wherein the summary matrix further comprises a proposed antisense strand.
9. The method of claim 7, wherein the summary matrix further comprises a seed sequence GC content and a guide strand GC content.
10. A method for identifying an RNAi sequence, comprising the steps of:

- a. screening for one or more 7-mers from a gRNA;
 - b. selecting a first number of the one or more 7-mers, the first number based on the frequency of the one or more 7-mers;
 - c. characterizing one or more hit sites for each of the one or more 7-mers;
 - d. calculating a hit per genomic region for each of the one or more 7-mers;
 - e. selecting a second number of the one or more 7-mers, the second number based on the most hits in a sgRNA;
 - f. creating a frequency matrix for one or more 15-mers associated to each of the one or more 7-mers;
 - g. generating, based on the most frequent nucleotides, one or more generated 15-mers for each of the one or more 7-mers;
 - h. characterizing hits in the gRNA for each of the one or more generated 15-mers;
 - i. calculating a summary index based on a cumulative hit frequency and a length;
and
 - j. selecting a third number of the RNAi sequences based on one or more generated features and one or more structural features.
11. The method of claim 10, wherein the first number is 500.
 12. The method of claim 10, wherein the second number is 50.
 13. The method of claim 10, wherein the third number is 5.
 14. The method of claim 10, wherein the one or more generated features and the one or more structural features are a function of the summary index.

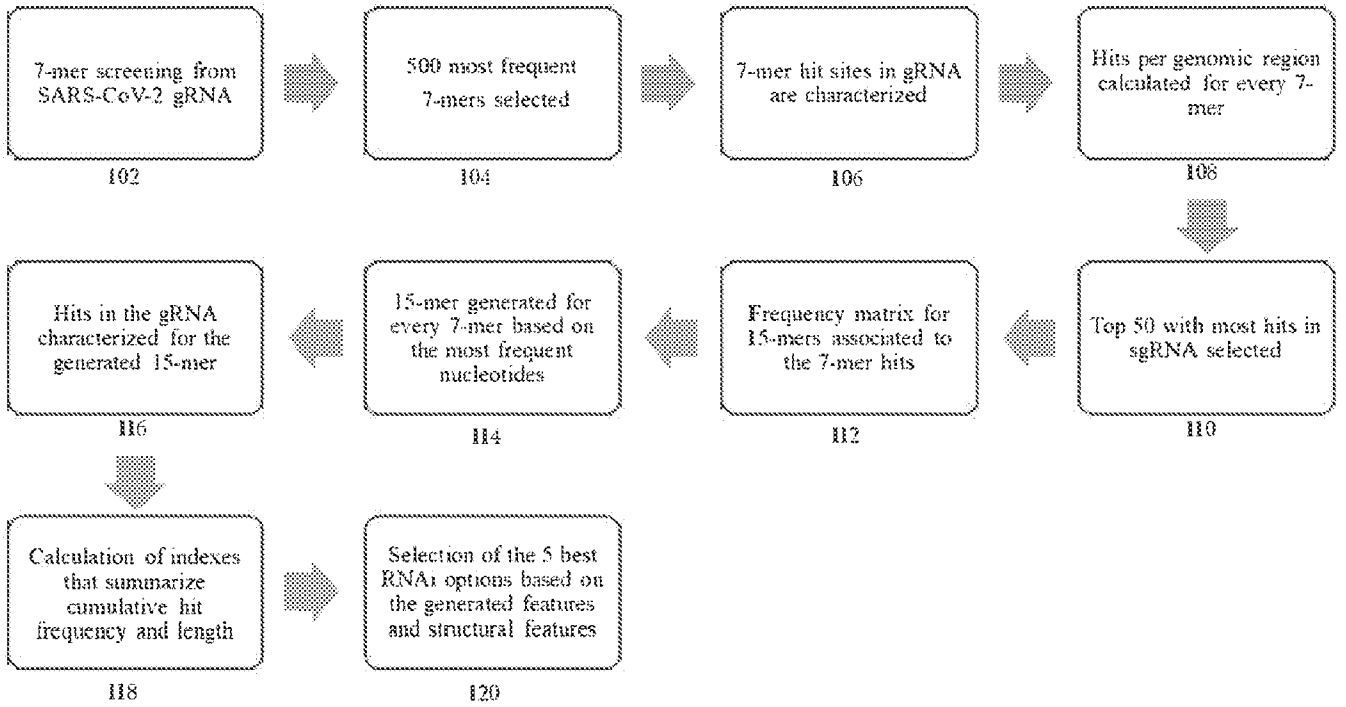


FIG.1

RNAi Identifier	Sequence
Anti-SC2-A	UCACCAUUACUAGGGUUGUGAUUAAACCUAGUAAUGGUGCUC
Anti-SC2-B	UGUUGAGUCAGAGCUAUGUGAUUAUAGCUCUGACUCAACCUC
Anti-SC2-C	UCAUUAGUAGGGUUGAUGUGAUUAUCAACCCUACUAAUGCUC
Anti-SC2-D	UUCACUGUACACUCGAUGUGAUUAUCGAGUGUACAGUGACUC
Anti-SC2-E	UUUACACAUUAGGGCUUGUGAUUAAGCCCUAAUGUGUAAACUC

FIG. 2A

RNAi Identifier	On 3'UTR	On CDS
Anti-SC2-A	0	17
Anti-SC2-B	1	23
Anti-SC2-C	1	8
Anti-SC2-D	1	0
Anti-SC2-E	1	0

FIG. 2B

NPM-SC2 ID	Sequence
A	UCACCAUUACUAGGUUUUGUGAUUAAACCUAGUAAUGGUGCUC
B	UGUUGAGUCAGAGCUAUGUGAUUAUAGCUCUGACUCAACCUC
C	UCAUUAGUAGGGUUGAUGUGAUUAUCAACCCUACUAAUGCUC
D	UUCACUGUACACUCGAUGUGAUUAUCGAGUGUACAGUGACUC
E	UUUACACAUUAGGGCUUGUGAUUAAGCCCUAAUGUGUAACUC
NC	UACGUGACACGUUCGGUGUGAUUACCGAACGUGUCACGUCUC

FIG. 3A

➤ SARS-CoV-2 3'UTR

CAATCTTTAATCAGTGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTTTCACCGA
 GGCCACGCGGAGTACGATCGAGTGTACAGTGAACAATGCTAGGGAGAGCTGCCTATATGG
 AAGAGCCCTAATGTGTAATAATTTTAGTAGTGCTATCCCATGTGATTTTAATAGCTTCT
 TAGGAGAATGACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIG. 3B

NPM-SC2 sequences 3'UTR regulation-

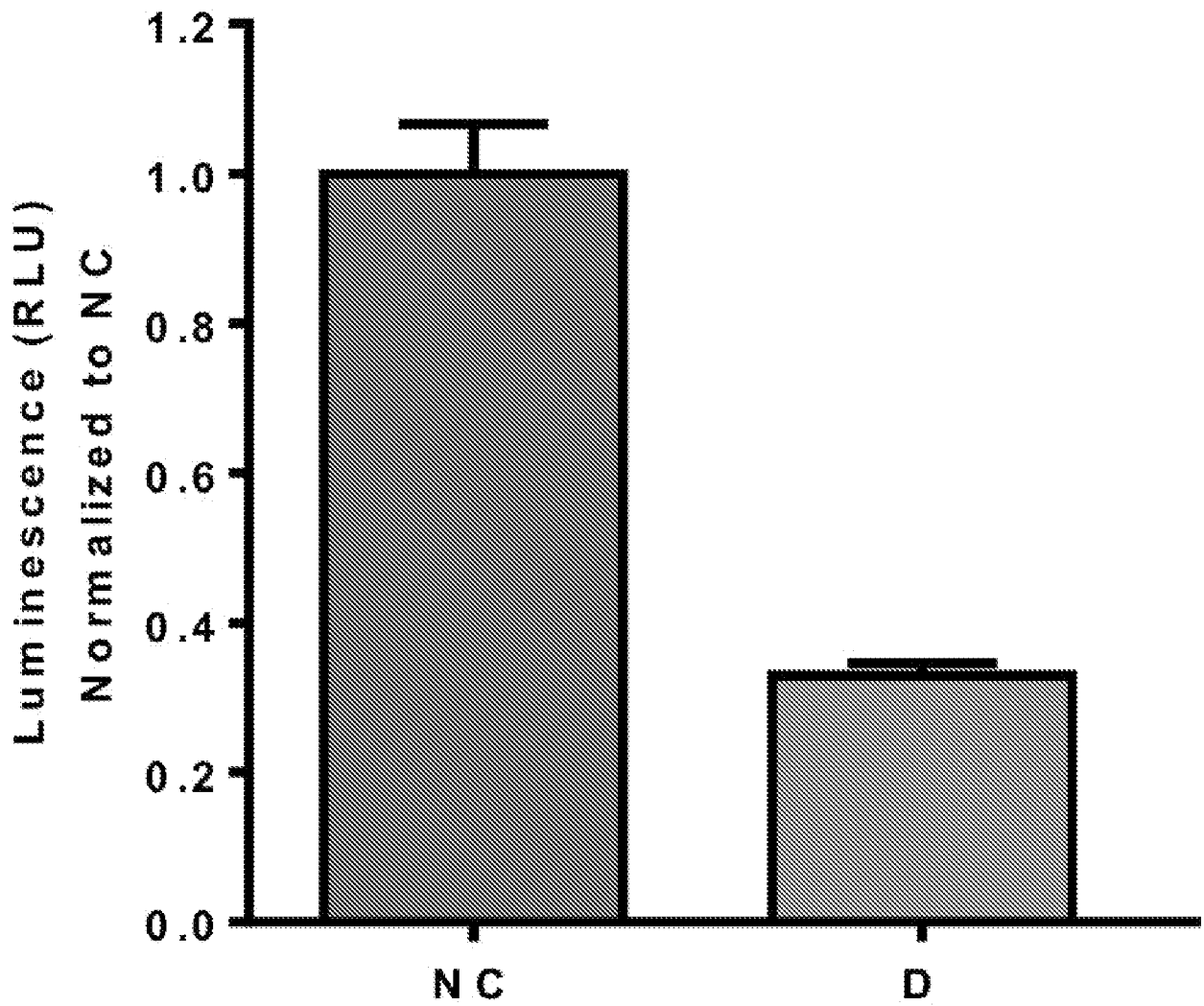


FIG. 4

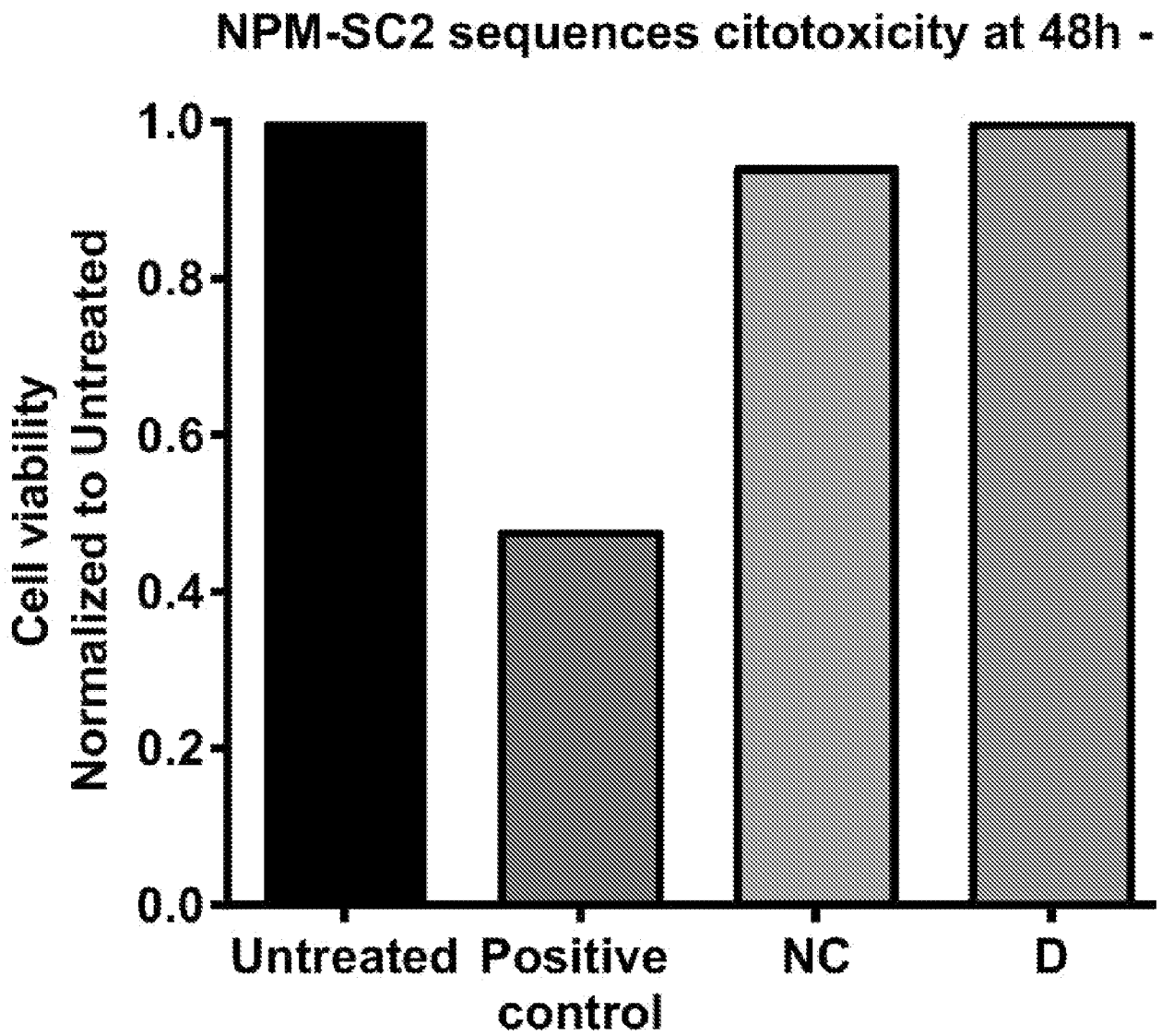


FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 2021/045790

A. CLASSIFICATION OF SUBJECT MATTER		<i>C12N 15/11 (2006.01)</i> <i>C12N 15/50 (2006.01)</i> <i>C12Q 1/6809 (2018.01)</i> <i>G06F 17/16 (2006.01)</i> <i>G06F 17/40 (2006.01)</i>
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)		
C12N 15/11, 15/50, C12Q 1/6809, G06F 17/16, 17/40		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
PatSearch (RUPTO Internal), USPTO, PAJ, Espacenet, Information Retrieval System of FIPS		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	WO 2010/138755 A1 (MERCK SHARP & DOHME CORP. et.al.) 02.12.2010, paragraphs [0025], [0075], [0135], [0315], [0336]	1-3, 7-14 4-6
Y	ADAMI Christoph et.al. From Entropy to Information: Biased Typewriters and the Origin of Life, 2017 pp.1-19 DOI: arXiv:1506.06988v2 [q-bio.PE], page 6, paragraph 3	1-3, 7-14
Y	ARAKAWA Hiroshi. A method to convert mRNA into a gRNA library for CRISPR/Cas9 editing of any organism. Research Article, 2016, pp.1-10, Results: gRNA library construction via molecular biology techniques, figure 1	10-14
<input type="checkbox"/> Further documents are listed in the continuation of Box C.		<input type="checkbox"/> See patent family annex.
* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“D” document cited by the applicant in the international application	“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
“E” earlier document but published on or after the international filing date	“&”	document member of the same patent family
“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		
Date of the actual completion of the international search	Date of mailing of the international search report	
01 November 2021 (01.11.2021)	18 November 2021 (18.11.2021)	
Name and mailing address of the ISA/RU: Federal Institute of Industrial Property, Berezhkovskaya nab., 30-1, Moscow, G-59, GSP-3, Russia, 125993 Facsimile No: (8-495) 531-63-18, (8-499) 243-33-37	Authorized officer G. Gromova Telephone No. 8(495)531-64-81	