



US 20250263699A1

(19) United States

(12) Patent Application Publication (10) Pub. No.: US 2025/0263699 A1

Zhu et al.

(43) Pub. Date: Aug. 21, 2025

(54) COMPOSITIONS AND METHODS FOR TREATING LIVER DISEASES WITH SIRNAS TARGETING TBX3

(71) Applicant: The Board of Regents of The University of Texas System, Austin, TX (US)

(72) Inventors: Hao Zhu, Dallas, TX (US); Zixi Wang, Dallas, TX (US); Lin Li, Dallas, TX (US)

(21) Appl. No.: 18/854,299

(22) PCT Filed: Apr. 7, 2023

(86) PCT No.: PCT/US2023/065554

§ 371 (c)(1),

(2) Date: Oct. 4, 2024

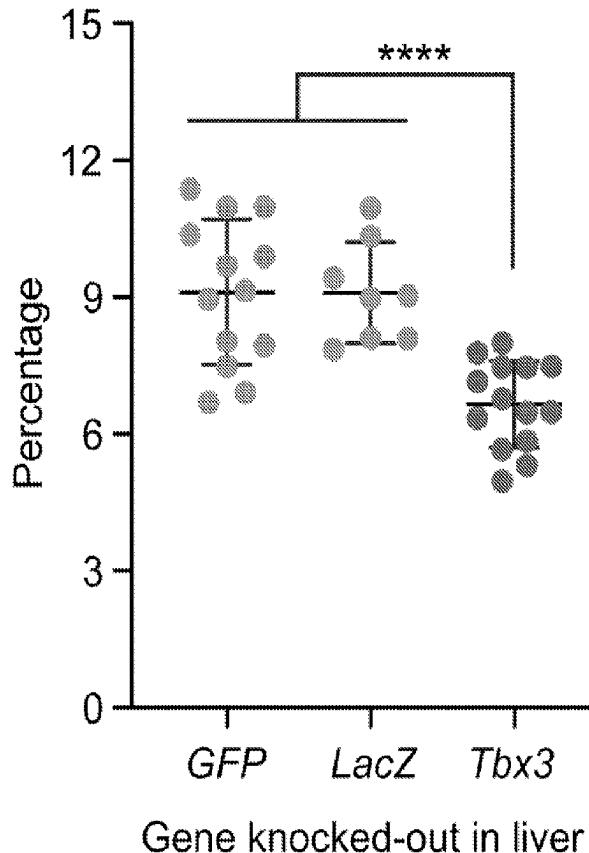
**Related U.S. Application Data**

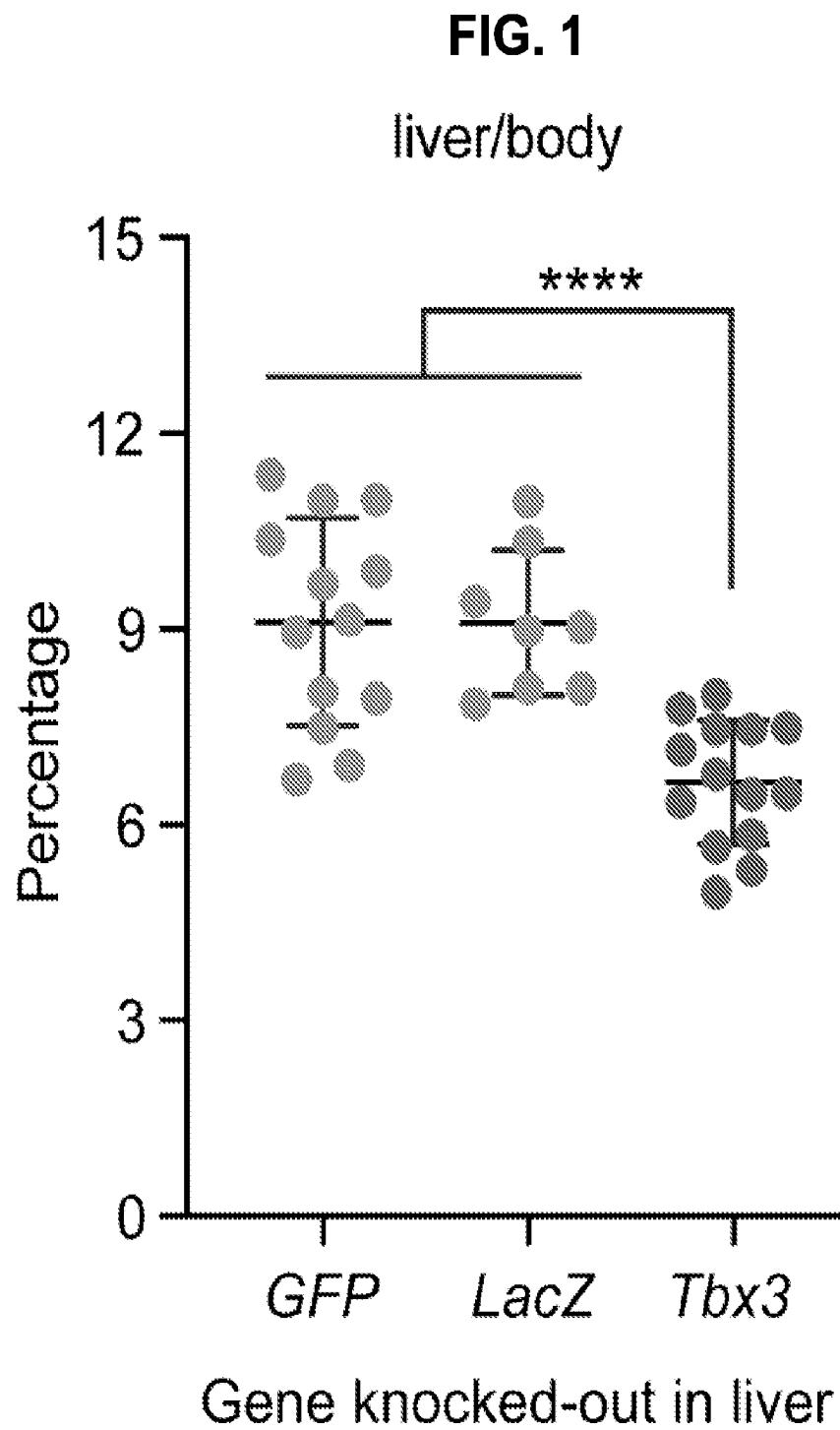
(60) Provisional application No. 63/491,264, filed on Mar. 20, 2023, provisional application No. 63/328,549, filed on Apr. 7, 2022.

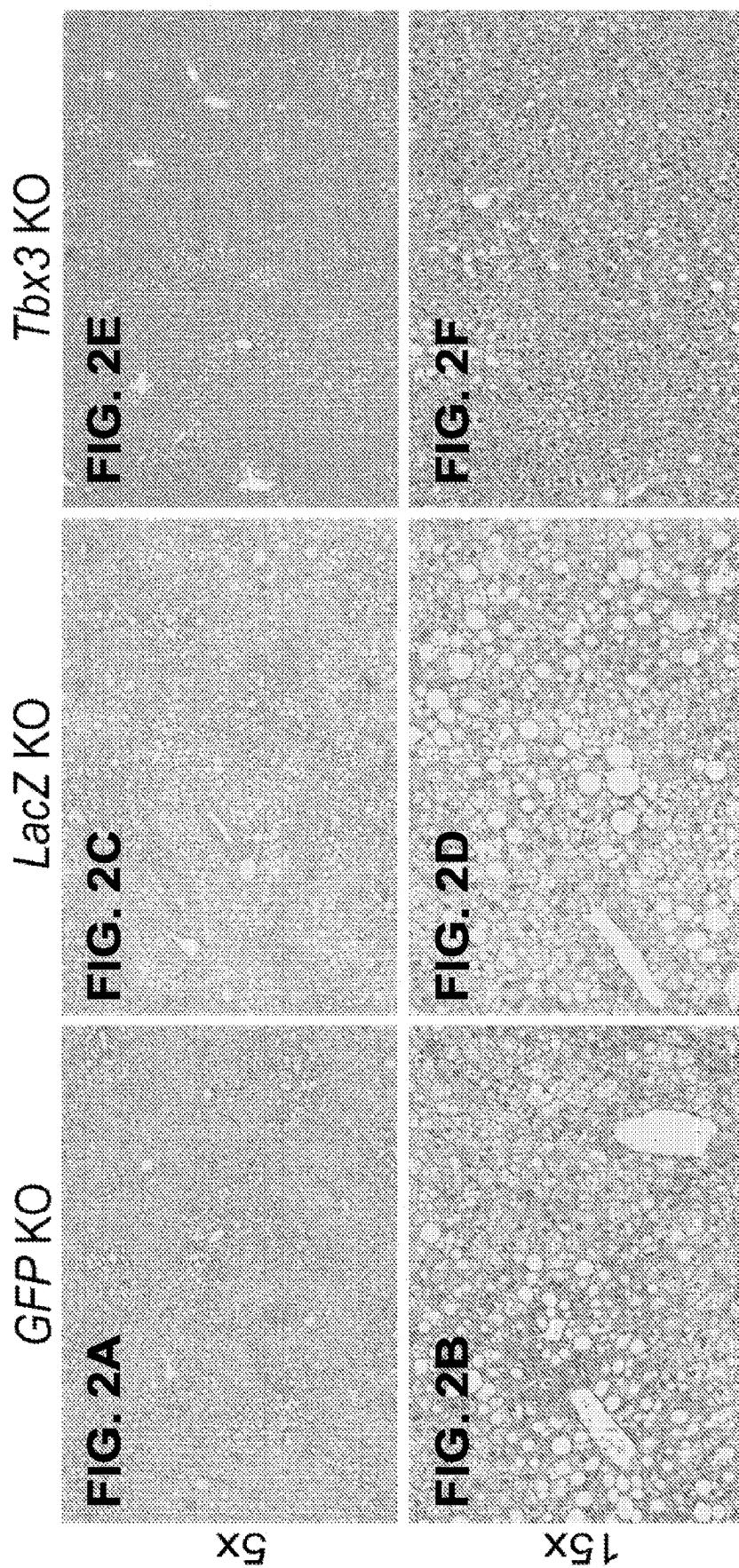
**Publication Classification**(51) Int. Cl.  
*C12N 15/113* (2010.01)  
*A61P 1/16* (2006.01)  
*C12N 9/22* (2006.01)  
*C12N 15/11* (2006.01)  
*C12N 15/86* (2006.01)  
*C12N 15/88* (2006.01)(52) U.S. Cl.  
CPC ..... *C12N 15/113* (2013.01); *A61P 1/16* (2018.01); *C12N 9/222* (2025.05); *C12N 15/111* (2013.01); *C12N 15/86* (2013.01); *C12N 15/88* (2013.01); *C12N 2310/11* (2013.01); *C12N 2310/14* (2013.01); *C12N 2310/20* (2017.05); *C12N 2310/321* (2013.01); *C12N 2310/322* (2013.01); *C12N 2310/351* (2013.01); *C12N 2320/32* (2013.01); *C12N 2750/14143* (2013.01)**(57) ABSTRACT**

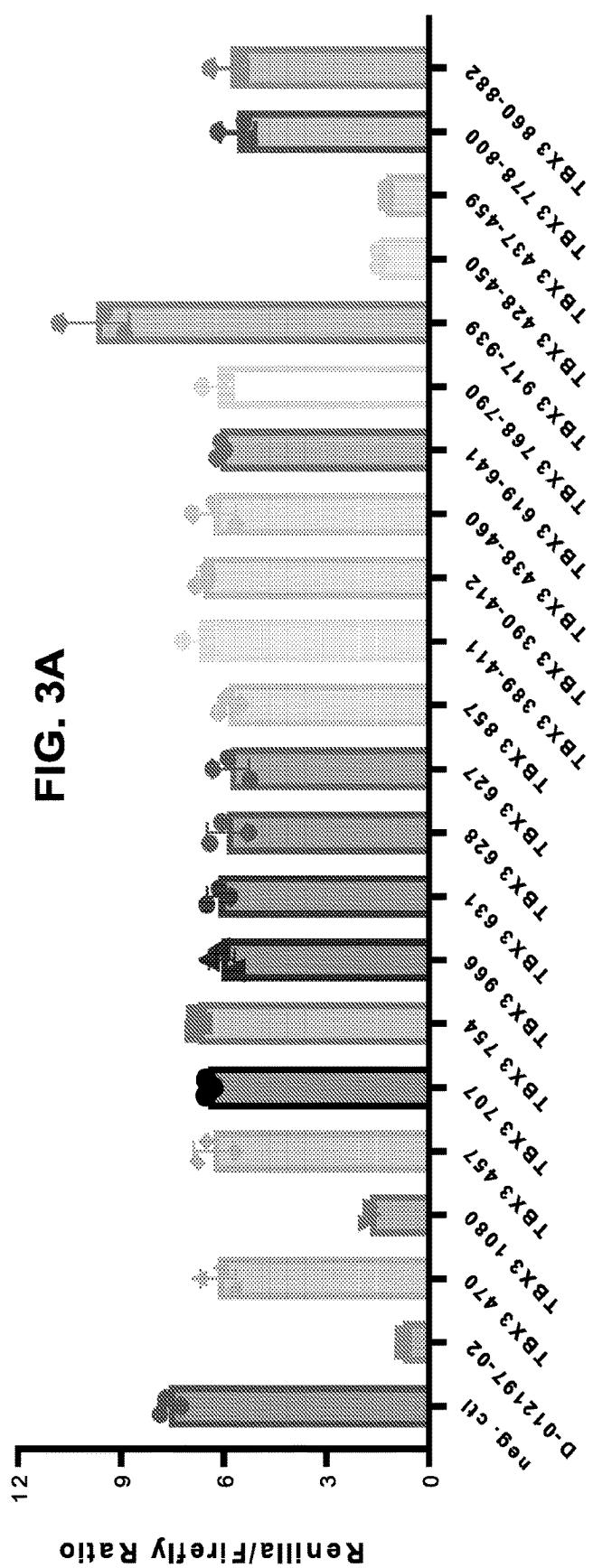
Disclosed herein are compositions comprising siRNAs capable of downregulating T-box transcription factor 3 (TBX3) gene expression or a variant thereof. Also disclosed herein are methods of using such compositions in the treatment of a liver disease or injury, such as fatty liver disease (FLD), non-alcoholic fatty liver disease (NAFLD) or non-alcoholic steatohepatitis (NASH).

Specification includes a Sequence Listing.

**liver/body**







**FIG. 3A**

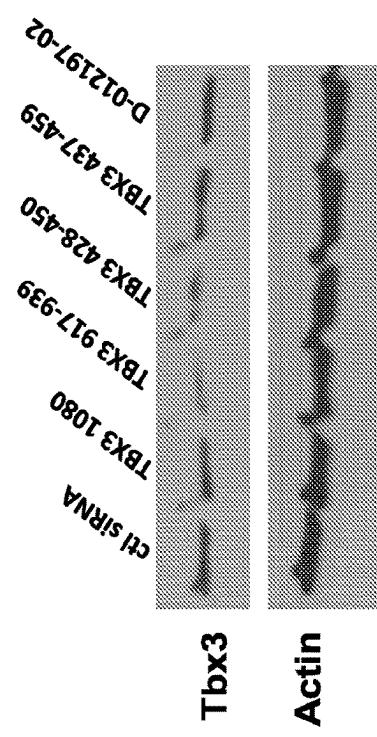
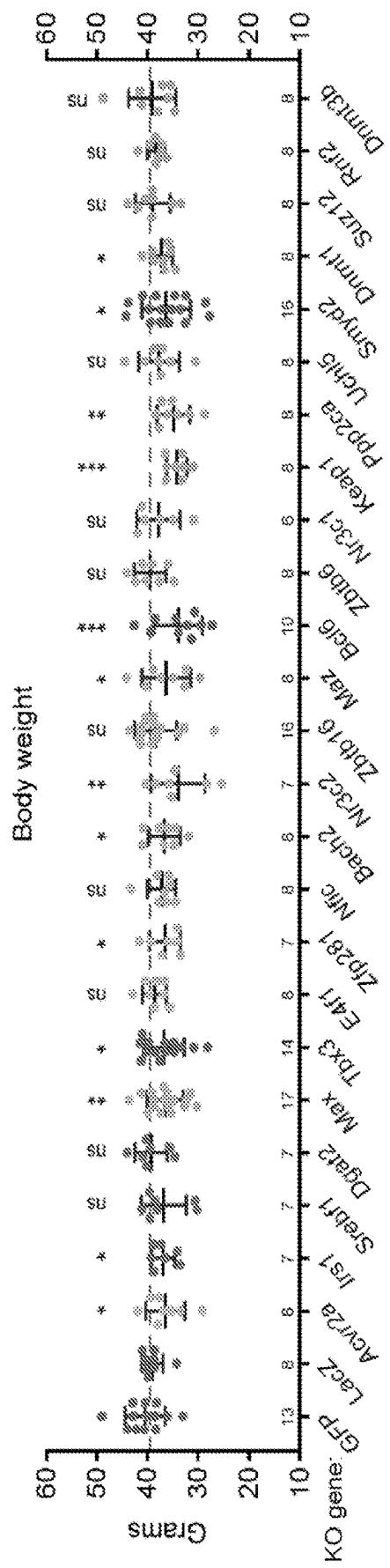
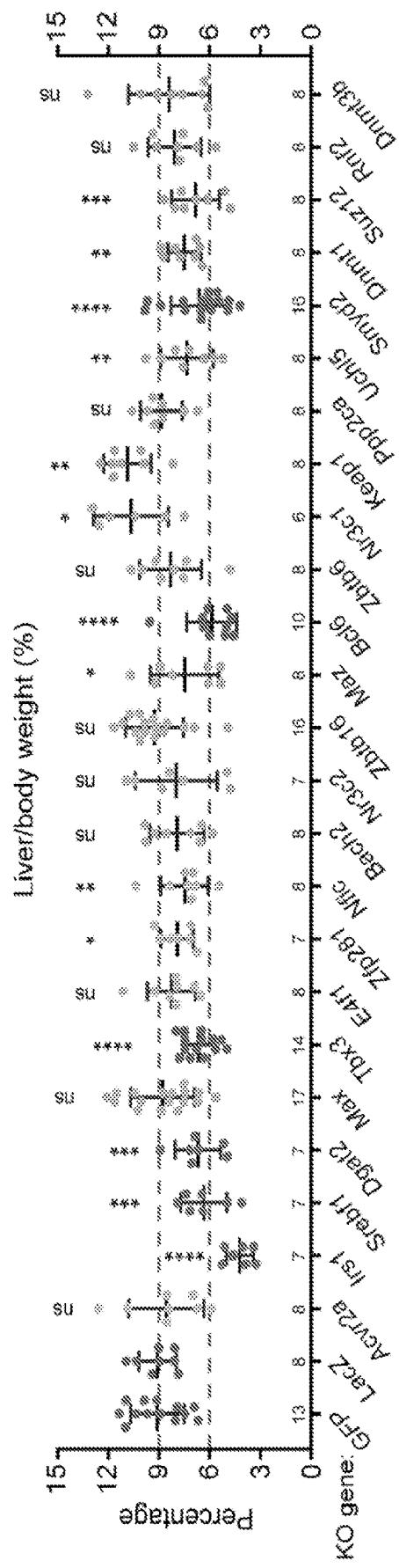


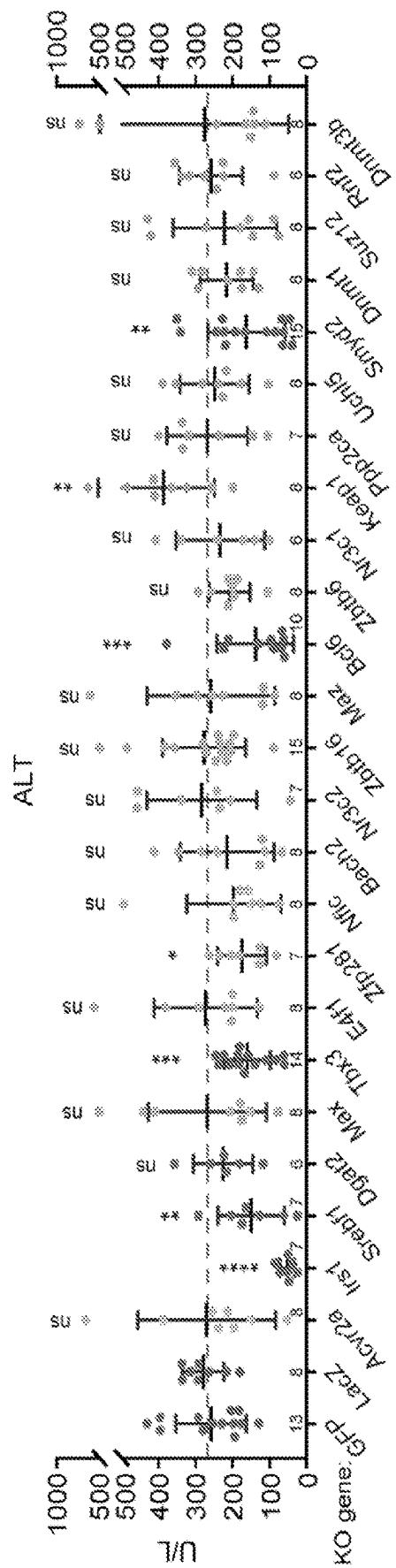
FIG. 3B



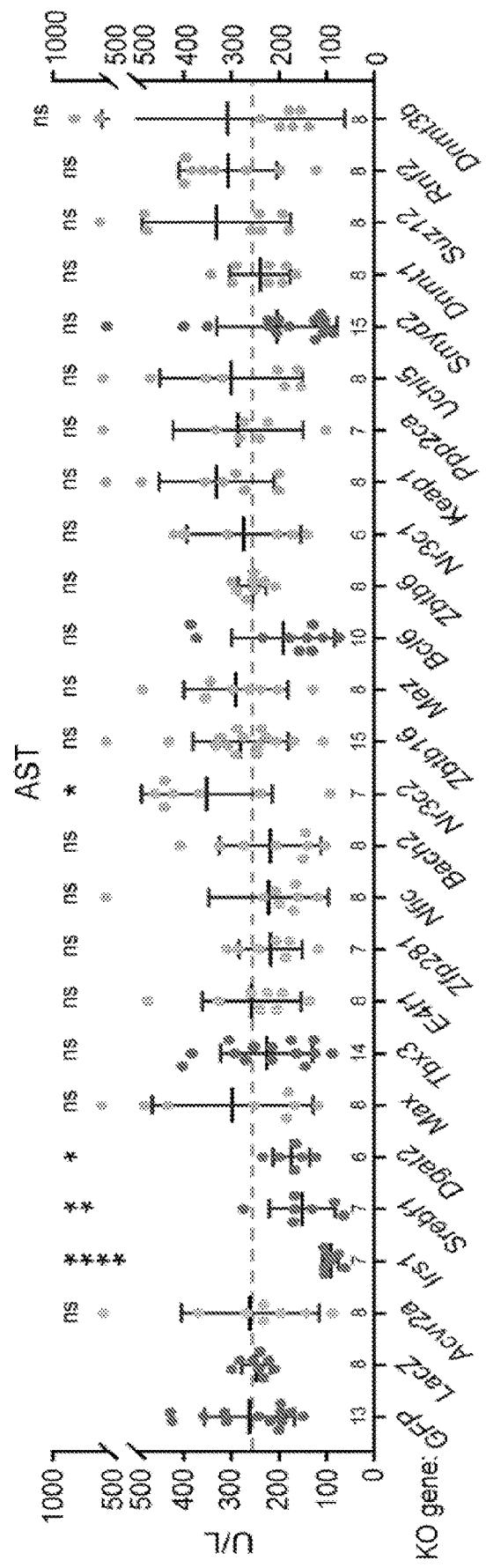
**FIG. 4A**



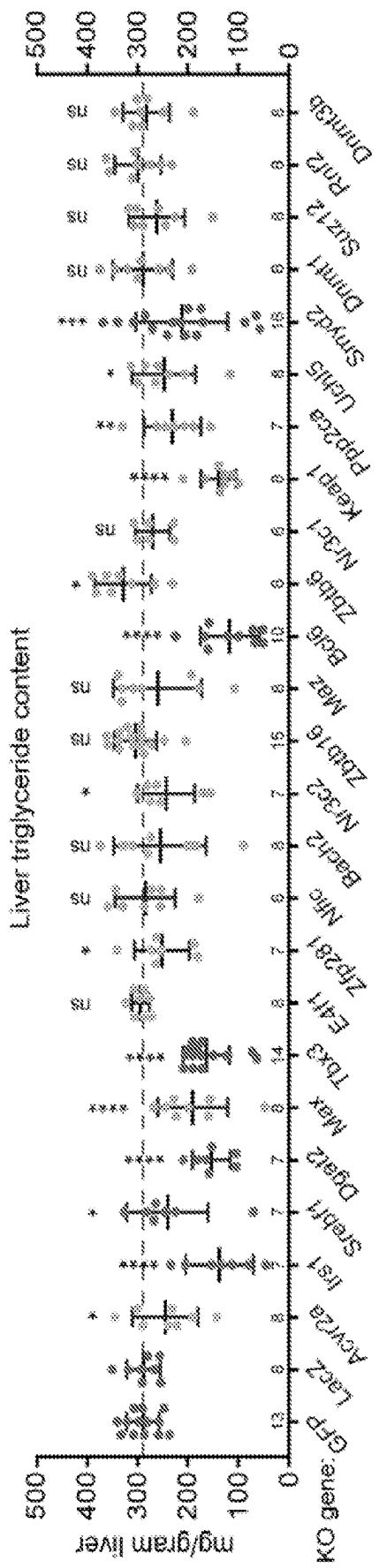
**FIG. 4B**



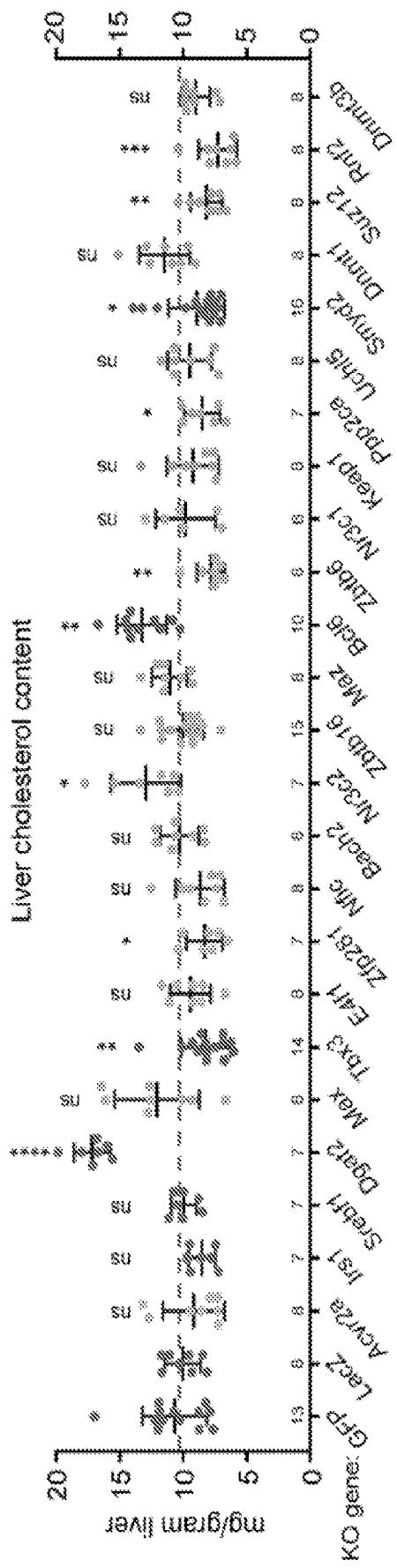
**FIG. 4C**



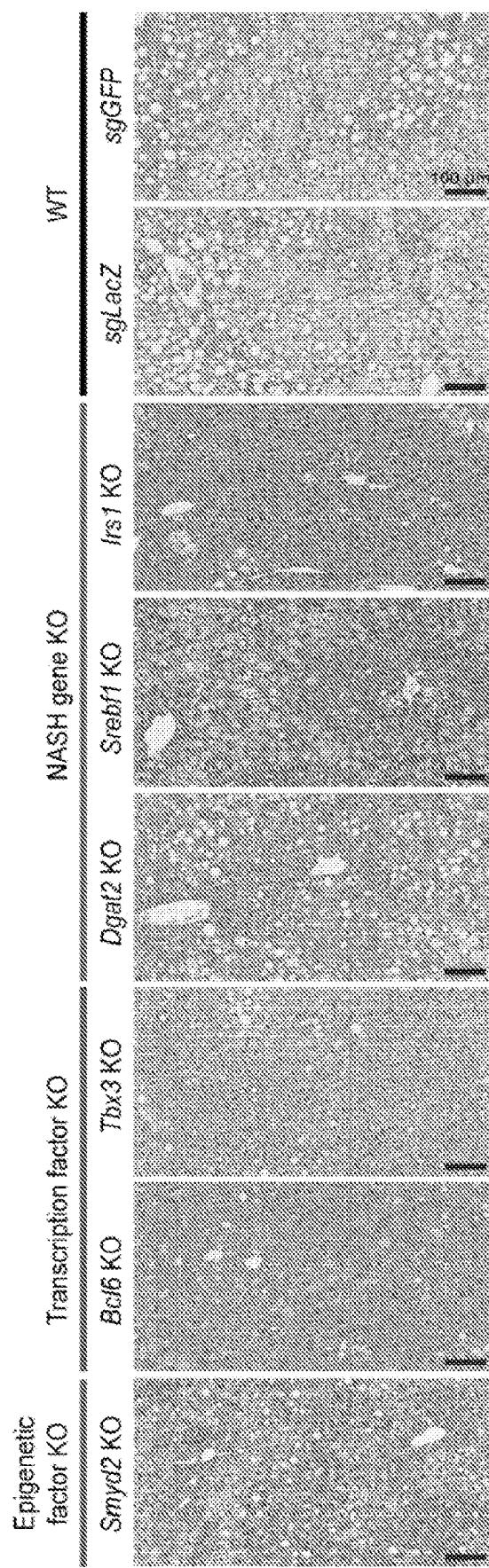
**FIG. 4D**



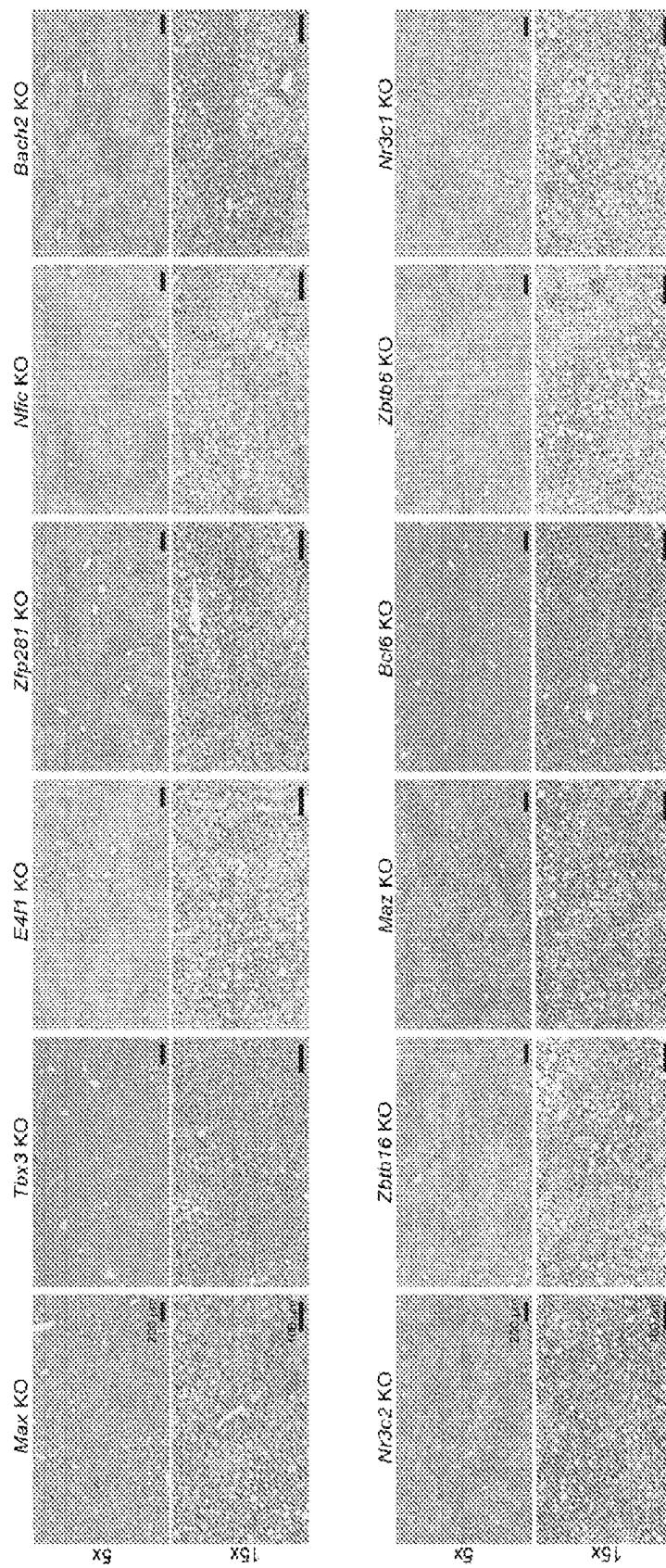
**FIG. 4E**



**FIG. 4F**



**FIG. 4G**



**FIG. 4H**

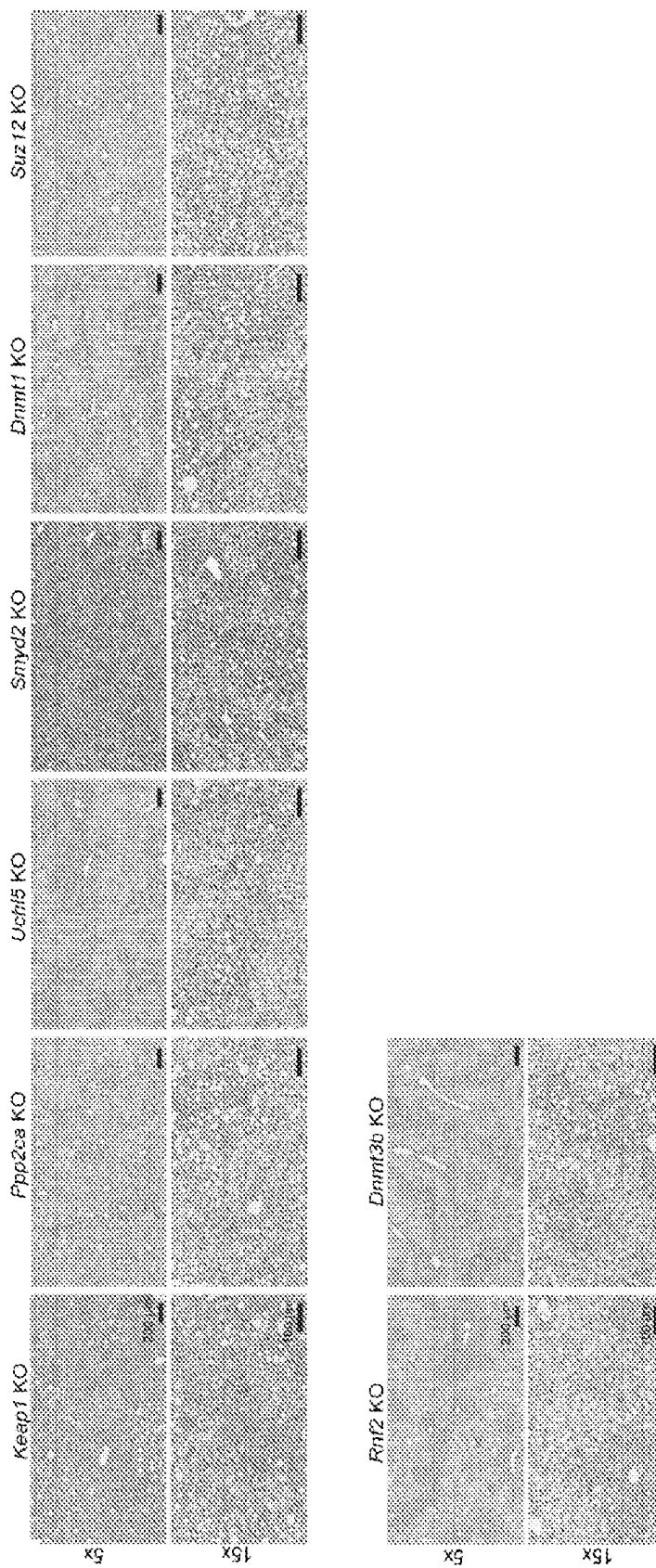


FIG. 4I

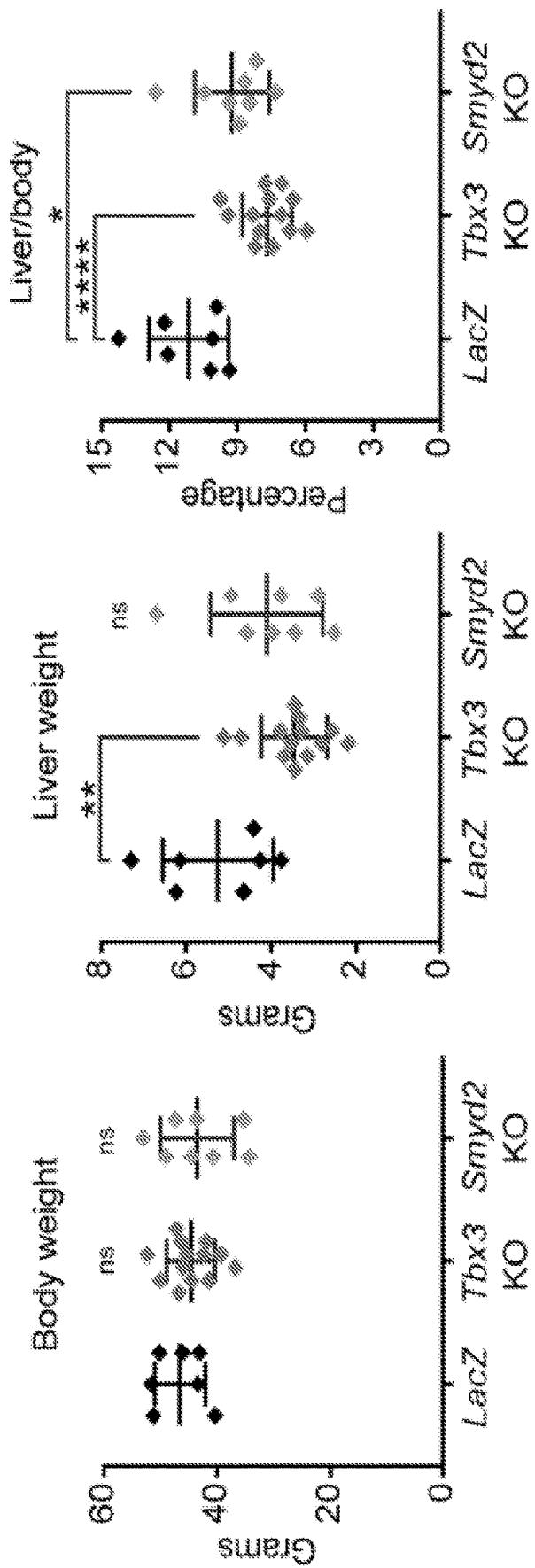
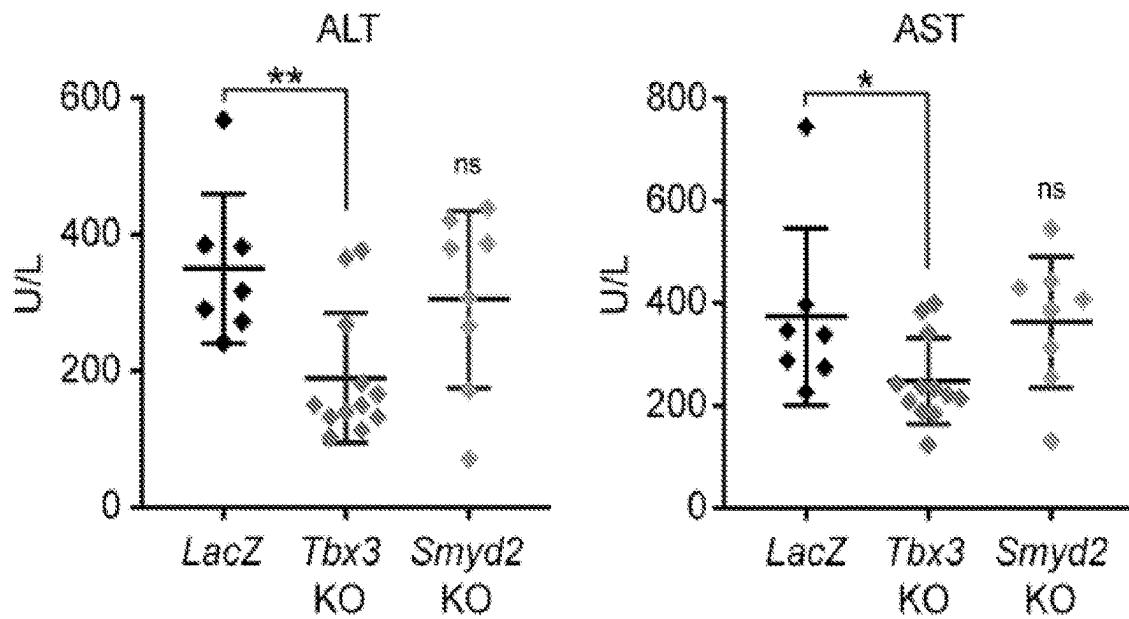
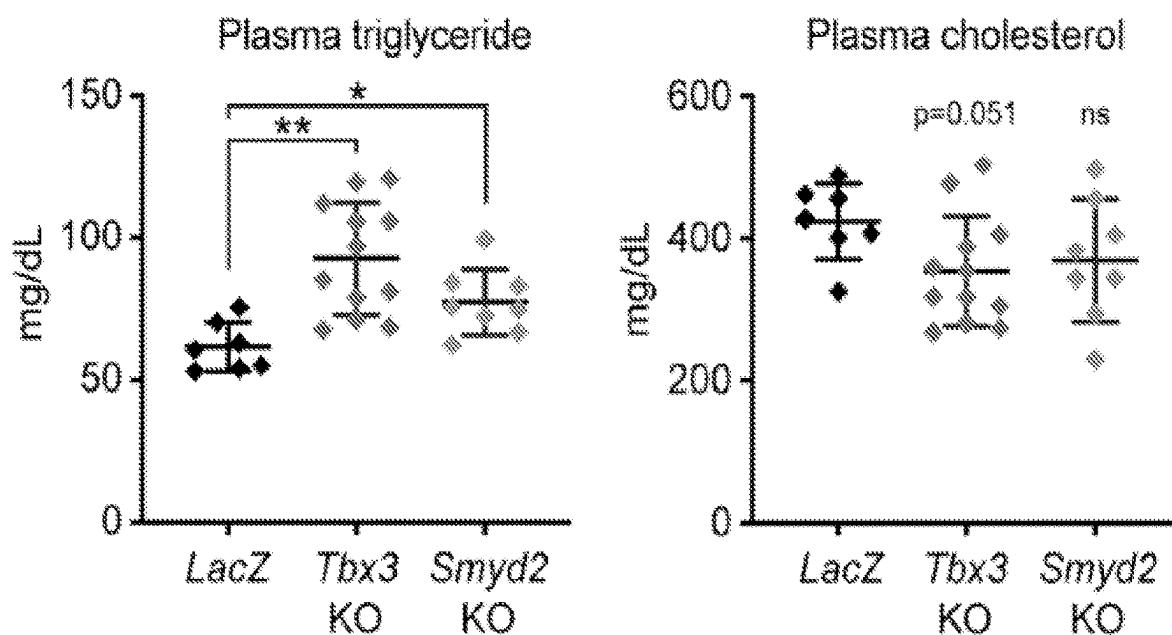


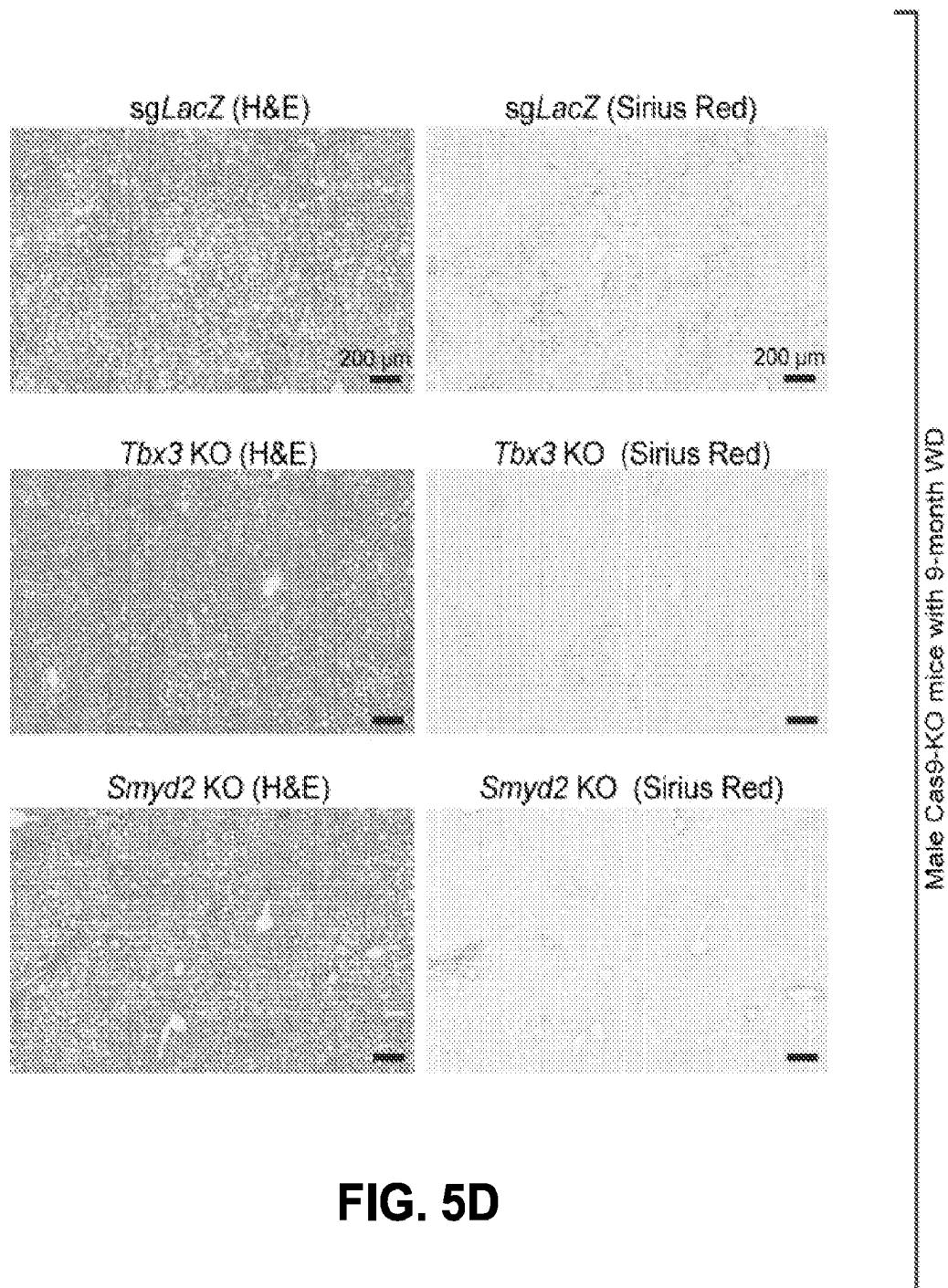
FIG. 5A



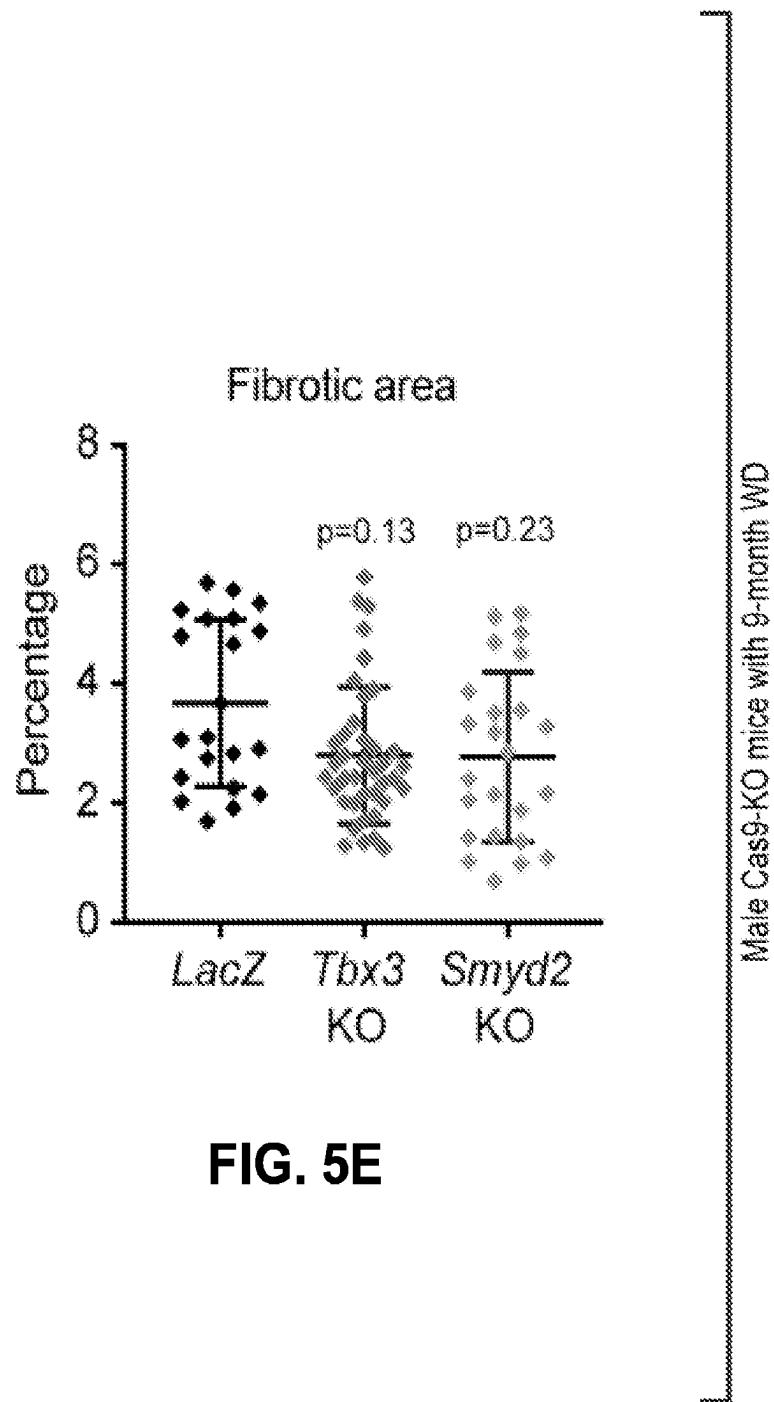
**FIG. 5B**

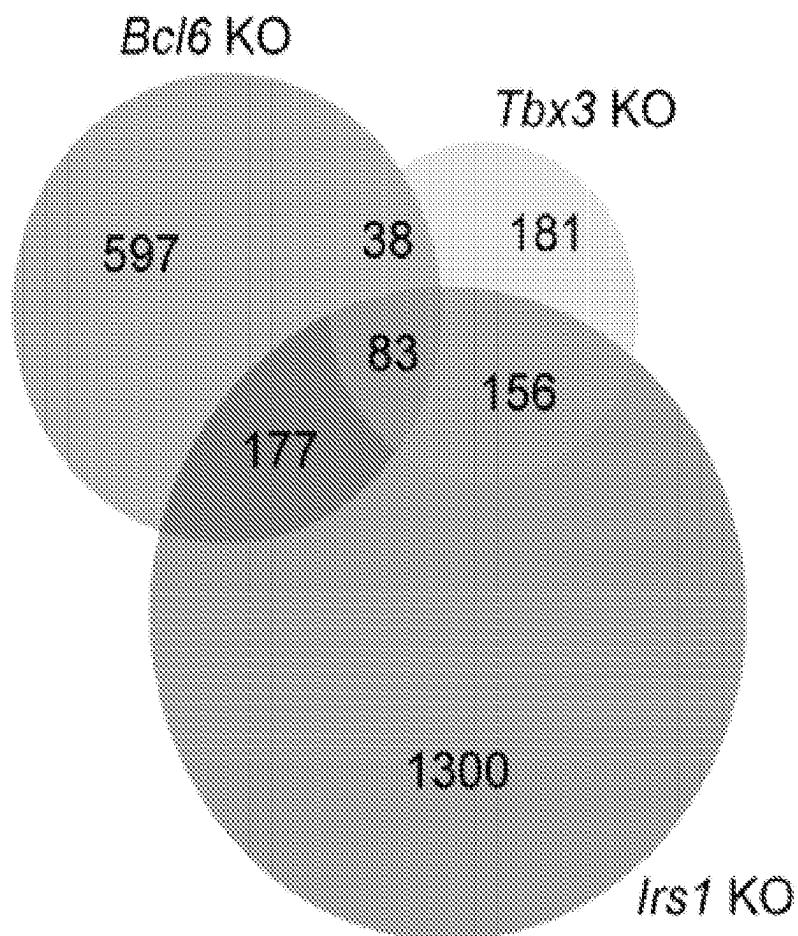


**FIG. 5C**

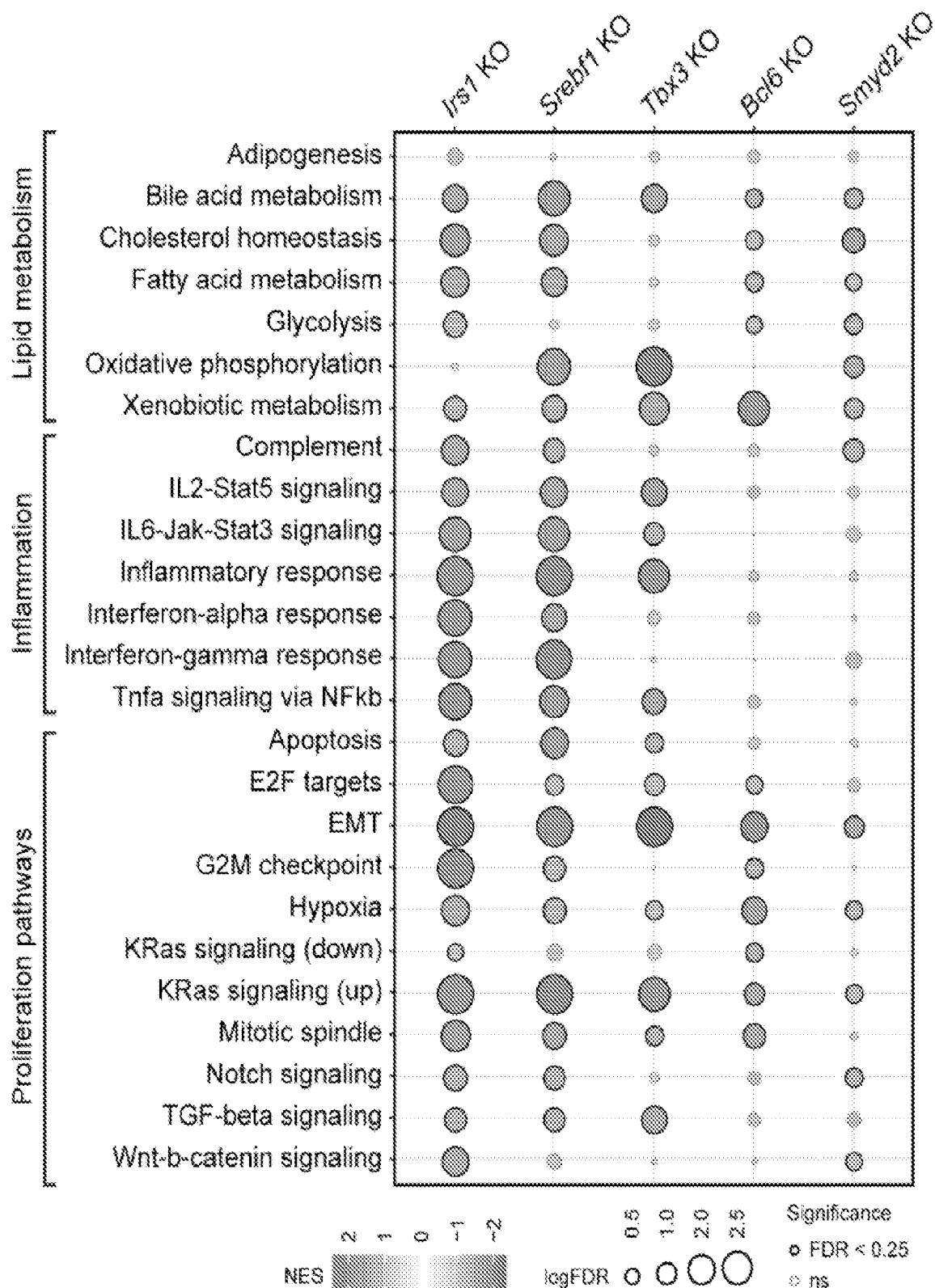


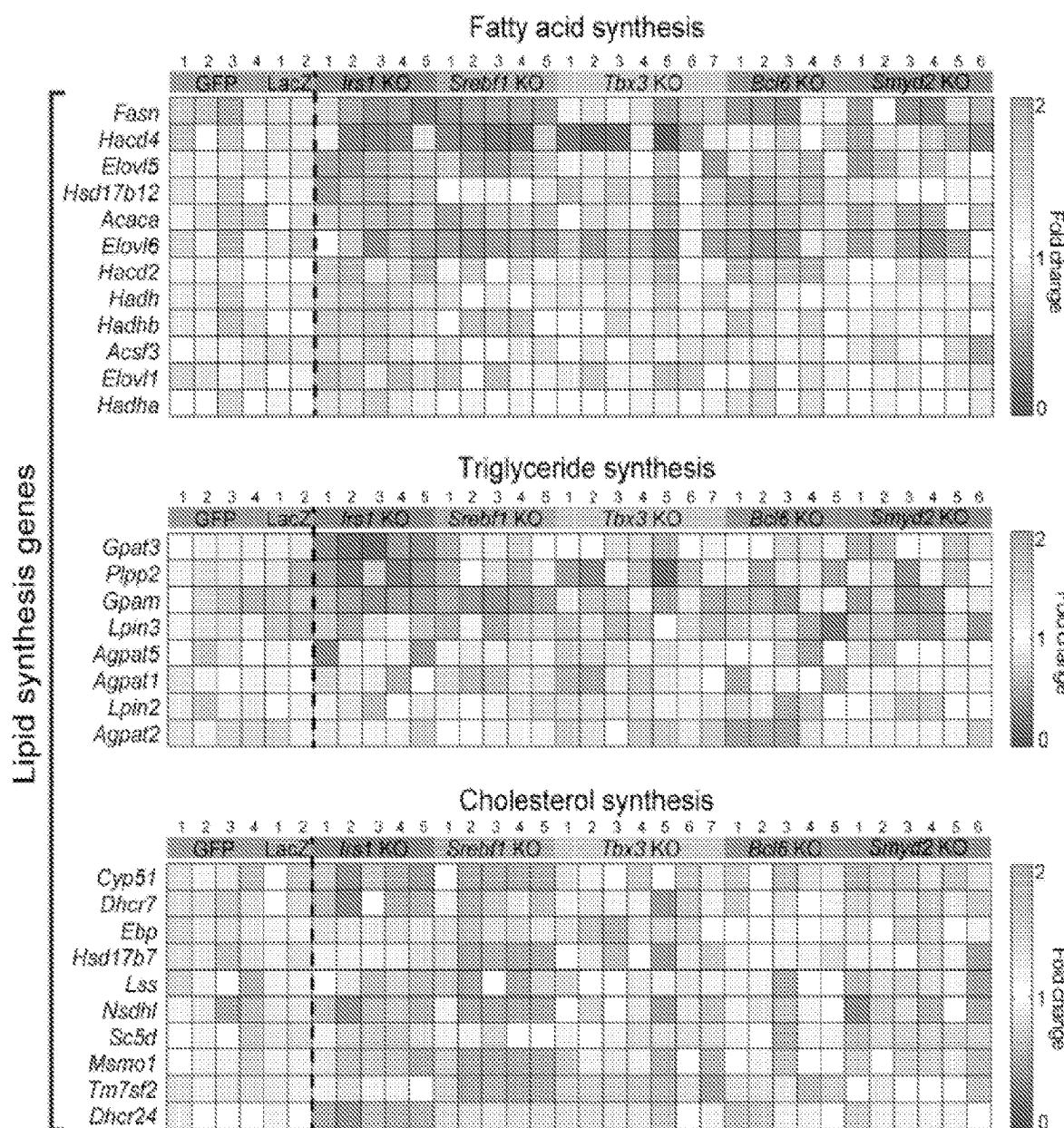
**FIG. 5D**

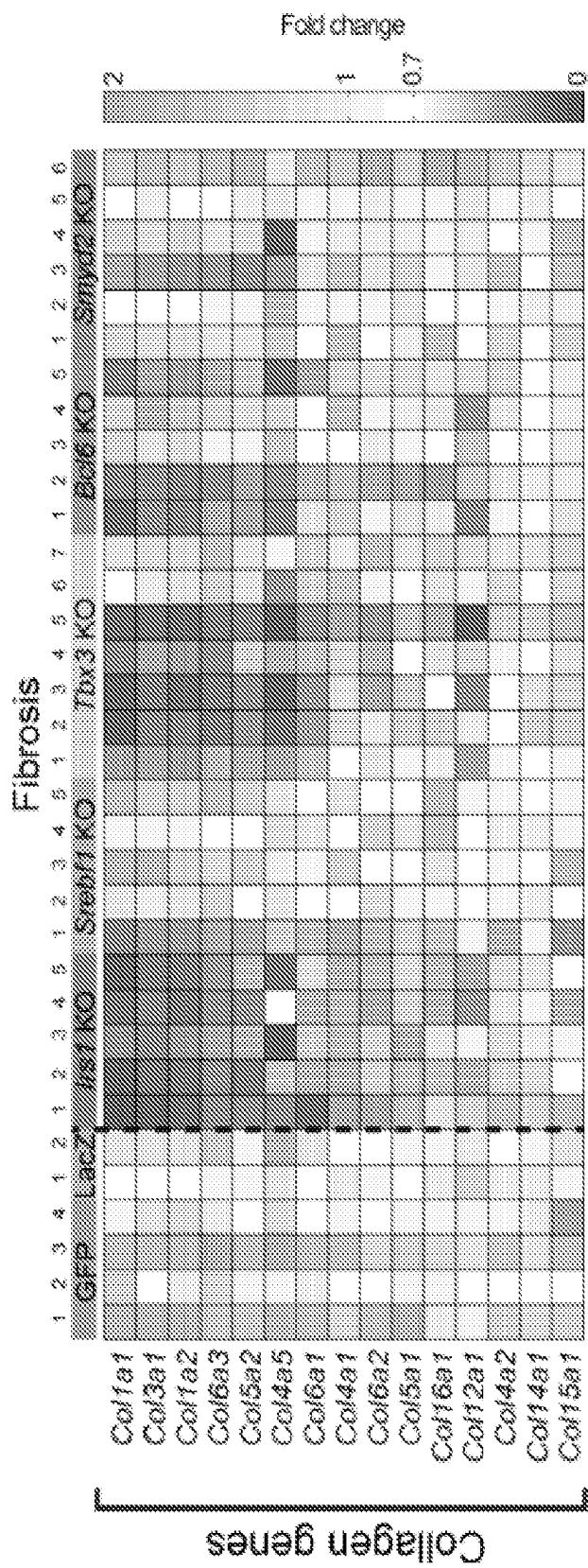




**FIG. 6A**

**FIG. 6B**


**FIG. 6C**



**FIG. 6D**

## COMPOSITIONS AND METHODS FOR TREATING LIVER DISEASES WITH SIRNAS TARGETING TBX3

### CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Application No. 63/328,549, filed Apr. 7, 2022, and U.S. Provisional Application No. 63/491,264, filed Mar. 20, 2023, the disclosures of which are incorporated herein by reference in their entireties for all purposes.

### SEQUENCE LISTING

[0002] This application contains a Sequence Listing that has been submitted in xml format via EFS-Web and is hereby incorporated by reference in its entirety. The xml copy, created on Apr. 7, 2023, is named "106546-755473.xml" and is about 281.4 KB in size.

### BACKGROUND

#### 1. Field

[0003] The present inventive concept is directed to compositions and methods of use thereof for liver disease treatment, specifically to metabolic liver diseases, including but not limited to non-alcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH).

#### 2. Discussion of Related Art

[0004] NAFLD is a spectrum of chronic liver disorders, which encompass, among others nonalcoholic fatty liver (NAFL) and nonalcoholic steatohepatitis (NASH). NASH, a common cause of chronic liver disease, is defined as having at least 5% hepatic steatosis and inflammation with or without fibrosis. Over time and without treatment, NASH may progress to cirrhosis and even hepatocellular carcinoma (HCC). Moreover, NAFLD/NASH is strongly associated with obesity and type II diabetes which together affect over 50% of the US population, leading to a heavy economic burden. Unfortunately, therapeutic options for NASH remain limited, with only slight benefits observed from vitamin E or obeticholic acid treatment. While NASH is the leading cause of chronic liver disease and cirrhosis, there are currently no clinically approved therapies. As such, new targets, therapeutics, and combinations thereof are needed to accelerate clinical progress in the treatment of liver diseases.

### SUMMARY OF THE INVENTION

[0005] The present disclosure is based, in part, on the novel finding that positive selection of somatic mutations in NASH patient livers can increase cell fitness and competitiveness through decreased lipid accumulation in NASH. In vivo genetic screening for transcriptional regulators that promote clonal fitness through the reversal of lipotoxicity identified TBX3 (T-box transcription factor 3). Accordingly, the present disclosure provides for novel compositions for modulating TBX3 expression and methods of preventing, attenuating and/or treating liver diseases.

[0006] In some aspects, the current disclosure encompasses a composition comprising a nucleic acid that downregulates expression of TBX3 (T-box transcription factor 3) or a variant thereof. In some aspects, the nucleic acid that

downregulates expression of TBX3 comprises a siRNA, a cluster regularly interspaced short palindromic repeats (CRISPR) related nucleic acid, a single guide RNA (sgRNA), a CRISPR-RNA (crRNA), or a trans-activating crRNA (tracrRNA). In some aspects, the nucleic acid is a small interfering RNA (siRNA) molecule. In some aspects, the composition comprises a plasmid or a viral vector, wherein the plasmid or the viral vector comprises a nucleic acid encoding the siRNA molecule as disclosed herein. In some aspects, the siRNA molecule comprises a nucleotide sequence that is 2 to 30 nucleotides in length and is at least 80% homologous to at least 2 to 30 contiguous nucleotides of a human TBX3 cDNA sequence. In some aspects, the human TBX3 cDNA sequence comprises SEQ ID NO: 1. In some aspects the siRNA molecule targets the open reading frame or the 5' or 3' UTRs of the TBX3 gene. In some aspects, the siRNA molecule comprises at least one sense sequence, at least one antisense sequence, or at least one sense sequence and at least one antisense sequence. In some aspects, the siRNA molecule comprises a nucleotide sequence SEQ ID NOs: 2-115 or any combination thereof. In some aspects, the at least one sense sequence comprises SEQ ID NOs: 2-58. In some aspects, the at least one antisense sequence comprises SEQ ID NOs: 59-115.

[0007] In some aspects, the current disclosure also encompasses a composition comprising a nucleic acid molecule that downregulates expression of TBX3, wherein the nucleic acid is a sgRNA or encodes an sgRNA. In some aspects, the composition comprises a plasmid or a viral vector, wherein the plasmid or the viral vector comprises a first nucleic acid encoding the sgRNA molecule as disclosed herein and optionally a second nucleic acid encoding an RNA guided nuclease. In some aspects, the sgRNA target sequence is at least about 80% identical to any one of SEQ ID NO: 118-217. In some aspects, the RNA guided nuclease is a Cas endonuclease.

[0008] In some aspects, the siRNA molecule as disclosed herein specifically downregulates gene expression of at least one variant of TBX3. In some aspects, the sgRNA molecule specifically downregulates gene expression of at least one variant of TBX3. In some aspects, the at least one variant of GPAM is associated with a liver disease. Non-limiting examples of liver disease comprises fatty liver disease (FLD), alcohol-related liver disease (ARLD), non-alcoholic fatty liver disease (NAFLD), non-alcoholic steatohepatitis (NASH), end stage liver disease (cirrhosis) from any etiology, liver cancer, or any combination thereof.

[0009] In some aspects, the nucleic acid molecule as disclosed herein may be conjugated to least one targeting ligand. In some aspects, at least one targeting ligand comprises a liver targeting ligand. In some aspects, the liver targeting ligand comprises at least one N-acetylgalactosamine (GalNAc) conjugate. In some aspects, the nucleic acid molecule is conjugated to about one to about three GalNAc conjugates. In some aspects, the nucleic acid molecule comprises at least one chemical modification. In some aspects, the nucleic acid molecule comprises a modification at least one ribosugar moiety of its nucleotide sequence. In some aspects, the at least one ribosugar moiety is modified with 2'-O-methyl (2'OMe), 2'-deoxy-2'-fluoro (2'F), 2'-deoxy, 5-C-methyl, 2'-O-(2-methoxyethyl) (MOE), 4'-thio, 2'-amino, 2'-C-allyl, or any combination thereof. In some aspects, less than about 10% to about 70% of ribosugar moieties of the total nucleotide sequence is modified.

[0010] In some aspects, the current disclosure also encompasses a pharmaceutical composition comprising any one of the compositions as disclosed herein and at least one pharmaceutically acceptable carrier. In some aspects, the pharmaceutical composition further comprises a nanoparticle. In some aspects, the pharmaceutical composition further comprises a lipid.

[0011] In some aspects, the current disclosure also encompasses a method of treating a subject in need thereof, the method comprising administrating a therapeutically effective amount of the composition as disclosed herein, or the pharmaceutical composition as disclosed herein. In some aspects, the subject in need thereof, is a human subject having or suspected of having a liver disease. In some aspects, the liver disease comprises fatty liver disease (FLD), alcohol-related liver disease (ARLD), non-alcoholic fatty liver disease (NAFLD), non-alcoholic steatohepatitis (NASH), end stage liver disease (cirrhosis) from any etiology, liver cancer, or any combination thereof. In some aspects, the method of administering comprises parenteral administration. In some aspects, the administration of a therapeutically effective amount of the composition as disclosed here or the pharmaceutical composition as disclosed herein increases life expectancy of the subject compared to an untreated subject with identical disease condition and predicted outcome. In some aspects of the method, administration of a therapeutically effective amount of the composition as disclosed herein or the pharmaceutical composition as disclosed herein increases liver function of the subject compared to an untreated subject with identical disease condition and predicted outcome. In some aspects of the method, administration of a therapeutically effective amount of the composition as disclosed herein or the pharmaceutical composition as disclosed herein attenuates liver fibrosis in the subject compared to an untreated subject with identical disease condition and predicted outcome. In some aspects, the administration of a therapeutically effective amount of the composition as disclosed herein or the pharmaceutical composition as disclosed herein prevents additional liver fibrosis in the subject compared to an untreated subject with identical disease condition and predicted outcome.

[0012] In some aspects, the current disclosure also encompasses a kit comprising: a container holding the composition as disclosed here or the pharmaceutical composition as disclosed herein, a pharmaceutical administrative means; and instructions for use.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0013] The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present disclosure, which can be better understood by reference to the drawing in combination with the detailed description of specific embodiments presented herein. Embodiments of the present inventive concept are illustrated by way of example in which like reference numerals indicate similar elements and in which:

[0014] FIG. 1 depicts a conditional liver gene deletion analysis using CRISPR, showing Tbx3 gene knockout resulting in reduced liver to body weight percentages. Single guide RNAs targeting GFP or LacZ delivered into mice with Cas9 expression served as controls and single guide RNAs targeting Tbx3 delivered into mice with Cas9 expression served as liver specific models of Tbx3 deletion. All mice

were fed with western diets and high sugar water for 3 months to induce a NASH phenotype, and each mouse is shown in FIG. 1 as a single red or blue dot. There was a statistically significant reduction in liver weight to body weight percentage after Tbx3 deletion, indicating protection from NASH.

[0015] FIG. 2A shows the liver histology after three (3) months of NASH diets in GFP knockout mice at 5x.

[0016] FIG. 2B shows the liver histology after three (3) months of NASH diets in GFP knockout mice at 15x.

[0017] FIG. 2C shows the liver histology after three (3) months of NASH diets in LacZ knockout mice at 5x.

[0018] FIG. 2D shows the liver histology after three (3) months of NASH diets in LacZ knockout mice at 15x.

[0019] FIG. 2E shows the liver histology after three (3) months of NASH diets in Tbx3 CRISPR knockout mice at 5x.

[0020] FIG. 2F shows the liver histology after three (3) months of NASH diets in Tbx3 CRISPR knockout mice at 15x. Representative images show significantly reduced hepatic steatosis in Tbx3 CRISPR knockout mice compared to the GFP and LacZ groups, which were representative controls.

[0021] FIG. 3A shows a luciferase-based screen used to test the relative knockdown efficacy for various TBX3 siRNAs.

[0022] FIG. 3B shows representative images of TBX3 protein levels examined using western blot analysis after siRNA knockdown.

[0023] FIG. 4A shows body weight ratios of control (sgGFP and sgLacZ) and liver-specific KO mice fed with 3 months of WD. Gray dots represent control mice, blue dots represent liver-specific KO mice for known NASH genes, red dots represent transcription factor KO mice, and green dots represent epigenetic factor KO mice. Darker dots represent mice that have the most significant differences in liver/body weight ratios. Each dot represents one mouse, and the n is denoted at the bottom of each plot.

[0024] FIG. 4B shows liver/body weight ratios of control (sgGFP and sgLacZ) and liver-specific KO mice fed with 3 months of WD. Gray dots represent control mice, blue dots represent liver-specific KO mice for known NASH genes, red dots represent transcription factor KO mice, and green dots represent epigenetic factor KO mice. Darker dots represent mice that have the most significant differences in liver/body weight ratios. Each dot represents one mouse, and the n is denoted at the bottom of each plot.

[0025] FIG. 4C shows liver function testing using plasma ALT. Gray dots represent control mice, blue dots represent liver-specific KO mice for known NASH genes, red dots represent transcription factor KO mice, and green dots represent epigenetic factor KO mice. Darker dots represent mice that have the most significant differences in liver/body weight ratios. Each dot represents one mouse, and the n is denoted at the bottom of each plot.

[0026] FIG. 4D shows liver function testing using plasma AST.

[0027] FIG. 4E shows liver function testing using liver triglycerides.

[0028] FIG. 4F shows liver function testing using liver cholesterol.

[0029] FIG. 4G shows representative H&E images of liver section are shown.

[0030] FIG. 4H shows H&E staining of liver sections from transcription factor KO mice.

[0031] FIG. 4I shows H&E staining of liver sections from epigenetic factor KO mice.

[0032] FIG. 5A shows body weight, liver weight, and liver/body weight ratios of male control, Tbx3 KO and Smyd2 KO mice fed with 9 months of WD (n=7, 13, 8 mice for each group).

[0033] FIG. 5B shows liver function testing with plasma ALT and AST for the mice (n=7, 12, 8 mice for each group).

[0034] FIG. 5C shows plasma cholesterol and triglycerides for the mice (n=7, 12, 8 mice for each group).

[0035] FIG. 5D shows H&E and Sirius Red staining of liver sections for the mice.

[0036] FIG. 5E shows quantification of Sirius Red staining images in 5D. Each dot represents one image field; three fields from each mouse liver were analyzed. Statistical analysis was performed on averaged image data from individual mice.

[0037] FIG. 6A shows a Venn diagram showing the shared and unique gene numbers with changed expression (fold change >=1.5) in Bcl6, Tbx3, and Irs1 KO livers.

[0038] FIG. 6B shows a hallmark pathway enrichment analysis of RNA-seq data.

[0039] FIG. 6C shows heatmaps showing the fold changes of differentially expressed genes in fatty acid, triglyceride, and cholesterol synthesis pathways. The average expression levels of control samples (four sgGFP and two sgLacZ) were normalized to 1 for each gene.

[0040] FIG. 6D shows heatmaps showing the fold changes of differentially expressed collagen genes. The average expression levels of control samples (four sgGFP and two sgLacZ) were normalized to 1 for each gene.

[0041] The drawing figures do not limit the present inventive concept to the specific embodiments disclosed and described herein. The drawings are not necessarily to scale, emphasis instead being placed on clearly illustrating principles of certain embodiments of the present inventive concept.

#### DETAILED DESCRIPTION

[0042] The following detailed description references the accompanying drawings that illustrate various embodiments of the present inventive concept. The drawings and description are intended to describe aspects and embodiments of the present inventive concept in sufficient detail to enable those skilled in the art to practice the present inventive concept. Other components can be utilized and changes can be made without departing from the scope of the present inventive concept. The following description is, therefore, not to be taken in a limiting sense. The scope of the present inventive concept is defined only by the appended claims, along with the full scope of equivalents to which such claims are entitled.

[0043] With increasing over-nutrition and obesity, non-alcoholic fatty liver disease (NAFLD) is rapidly becoming the leading cause of liver disease in the world. NAFLD is usually conceptualized at the organismal and tissue levels; however, little thought has been given to genetic heterogeneity within clones of the liver. Somatic mutations are common in most healthy individuals, and there is accumulating evidence that mutation burden increases with age and chronic tissue damage. The present disclosure is based, in part, on the novel finding that positive selection of somatic

mutations in NASH patient livers can increase cell fitness and competitiveness through decreased lipid accumulation in NASH. Based on this novel concept, in vivo genetic screening of somatically mutated transcriptional regulators dysregulated in human liver disease was performed. This genetic screen identified genes that when mutated, promoted clonal fitness through the reversal of lipotoxicity, including TBX3 (T-box transcription factor 3). Accordingly, provided herein are compositions and methods of use thereof for liver disease treatment, specifically to metabolic liver diseases, including but not limited to non-alcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH).

#### I. Terminology

[0044] The phraseology and terminology employed herein are for the purpose of description and should not be regarded as limiting. For example, the use of a singular term, such as, "a" is not intended as limiting of the number of items. Also, the use of relational terms such as, but not limited to, "top," "bottom," "left," "right," "upper," "lower," "down," "up," and "side," are used in the description for clarity in specific reference to the figures and are not intended to limit the scope of the present inventive concept or the appended claims.

[0045] Further, as the present inventive concept is susceptible to embodiments of many different forms, it is intended that the present disclosure be considered as an example of the principles of the present inventive concept and not intended to limit the present inventive concept to the specific embodiments shown and described. Any one of the features of the present inventive concept may be used separately or in combination with any other feature. References to the terms "embodiment," "embodiments," and/or the like in the description mean that the feature and/or features being referred to are included in, at least, one aspect of the description. Separate references to the terms "embodiment," "embodiments," and/or the like in the description do not necessarily refer to the same embodiment and are also not mutually exclusive unless so stated and/or except as will be readily apparent to those skilled in the art from the description. For example, a feature, structure, process, step, action, or the like described in one embodiment may also be included in other embodiments but is not necessarily included. Thus, the present inventive concept may include a variety of combinations and/or integrations of the embodiments described herein. Additionally, all aspects of the present disclosure, as described herein, are not essential for its practice. Likewise, other systems, methods, features, and advantages of the present inventive concept will be, or become, apparent to one with skill in the art upon examination of the figures and the description. It is intended that all such additional systems, methods, features, and advantages be included within this description, be within the scope of the present inventive concept, and be encompassed by the claims.

[0046] As used herein, the term "about," can mean relative to the recited value, e.g., amount, dose, temperature, time, percentage, etc.,  $\pm 10\%$ ,  $\pm 9\%$ ,  $\pm 8\%$ ,  $\pm 7\%$ ,  $\pm 6\%$ ,  $\pm 5\%$ ,  $\pm 4\%$ ,  $\pm 3\%$ ,  $\pm 2\%$ , or  $\pm 1\%$ .

[0047] The terms "comprising," "including," "encompassing" and "having" are used interchangeably in this disclosure. The terms "comprising," "including," "encompassing" and "having" mean to include, but not necessarily be limited to the things so described.

[0048] The terms “or” and “and/or,” as used herein, are to be interpreted as inclusive or meaning any one or any combination. Therefore, “A, B or C” or “A, B and/or C” mean any of the following: “A,” “B” or “C”; “A and B”; “A and C”; “B and C”; “A, B and C.” An exception to this definition will occur only when a combination of elements, functions, steps or acts are in some way inherently mutually exclusive.

[0049] “Biomarker” as used herein refers to any biological molecules (e.g., nucleic acids, genes, peptides, proteins, lipids, hormones, metabolites, and the like) that, singularly or collectively, reflect the current or predict future state of a biological system. Thus, as used herein, the presence or concentration of one or more biomarkers can be detected and correlated with a known condition, such as a disease state. In some aspects, detecting the presence and/or concentration of one or more biomarkers herein may be an indication of a liver disease risk in a subject. In some other aspects, detecting the presence and/or concentration of one or more biomarkers herein may be used in treating and/or preventing a chronic liver disease in a subject.

[0050] As used herein, the terms “treat”, “treating”, “treatment” and the like, unless otherwise indicated, can refer to reversing, alleviating, inhibiting the process of, or preventing the disease, disorder or condition to which such term applies, or one or more symptoms of such disease, disorder or condition and includes the administration of any of the compositions, pharmaceutical compositions, or dosage forms described herein, to prevent the onset of the symptoms or the complications, or alleviating the symptoms or the complications, or eliminating the condition, or disorder.

[0051] The term “biomolecule” as used herein refers to, but is not limited to, proteins, enzymes, antibodies, DNA, siRNA, and small molecules. “Small molecules” as used herein can refer to chemicals, compounds, drugs, and the like.

[0052] The term “nucleic acid” or “polynucleotide” refers to deoxyribonucleic acids (DNA) or ribonucleic acids (RNA) and polymers thereof in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides that have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions), alleles, orthologs, SNPs, and complementary sequences as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer et al., Nucleic Acid Res. 19:5081 (1991); Ohtsuka et al., J. Biol. Chem. 260:2605-2608 (1985); and Rossolini et al., Mol. Cell. Probes 8:91-98 (1994)).

[0053] The terms “peptide,” “polypeptide,” and “protein” are used interchangeably, and refer to a compound comprised of amino acid residues covalently linked by peptide bonds. A protein or peptide must contain at least two amino acids, and no limitation is placed on the maximum number of amino acids that can comprise a protein’s or peptide’s sequence. Polypeptides include any peptide or protein comprising two or more amino acids joined to each other by peptide bonds. As used herein, the term refers to both short

chains, which also commonly are referred to in the art as peptides, oligopeptides and oligomers, for example, and to longer chains, which generally are referred to in the art as proteins, of which there are many types. “Polypeptides” include, for example, biologically active fragments, substantially homologous polypeptides, oligopeptides, homodimers, heterodimers, variants of polypeptides, modified polypeptides, derivatives, analogs, fusion proteins, among others. A polypeptide includes a natural peptide, a recombinant peptide, or a combination thereof.

[0054] It should also be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one step or act, the order of the steps or acts of the method is not necessarily limited to the order in which the steps or acts of the method are recited.

## II. Compositions

### (a) TBX3

[0055] In certain embodiments, compositions for use in the methods disclosed herein can modulate TBX3 (T-box transcription factor 3) gene. The TBX3 gene is a member of a phylogenetically conserved family of genes that share a common DNA-binding domain, the T-box. These genes encode transcription factors involved in the regulation of developmental processes. Additionally, alternative splicing results in three transcript variants encoding different isoforms.

[0056] As used herein, compositions “modulating” TBX3 can include any biomolecule(s) capable of decreasing TBX3 gene expression, decreasing TBX3 protein expression, decreasing TBX3 activity, or a combination thereof. In some aspects, biomolecule(s) herein capable of modulating TBX3 can be an inhibitor of TBX3. As used herein, an inhibitor of TBX3 can inhibit TBX3 direct activity, inhibit TBX3 indirect activity, decrease expression of the TBX3 gene, decrease expression of the TBX3 protein, or a combination thereof.

[0057] In certain embodiments, compositions for use in the methods disclosed herein can include a nucleic acid molecule. The term “nucleic acid molecule” as used herein refers to a molecule having nucleotides. The nucleic acid can be single, double, or multiple stranded and may comprise modified or unmodified nucleotides or non-nucleotides or various mixtures and combinations thereof. In some embodiments, a nucleic acid molecule for use herein can be a double-stranded RNA. In some examples, a double stranded RNA suitable for use herein can be small temporal RNA, small nuclear RNA, small nucleolar RNA, short hairpin RNA, microRNA, or the like. In certain embodiments, a double stranded RNA suitable for use herein can be a small interfering RNA (siRNA).

[0058] The term “siRNA” as used herein refers to small inhibitory RNA duplexes that induce the RNA interference (RNAi) pathway. A siRNA molecule disclosed herein may be capable of silencing, reducing, and/or inhibiting expression of a target gene (e.g., TBX3). These molecules can vary in length (generally about 5-50 base pairs) and contain varying degrees of complementarity to their target mRNA in the antisense strand. Some, but not all, siRNA have unpaired overhanging bases on the 5' or 3' end of the sense strand and/or the antisense strand. The term “siRNA” as used

herein can include duplexes of two separate strands, as well as single strands that can form hairpin structures comprising a duplex region.

[0059] siRNA molecules disclosed herein may be any interfering RNA with a duplex length of about 2-60, about 5-50, or about 10-40 nucleotides in length, more typically about 2-30, about 5-25, or about 10-25 nucleotides in length. In some embodiments, siRNA molecules disclosed herein may have a nucleotide sequence that is about 2 to about 30 nucleotides in length, about 5 to about 25 nucleotides in length, or about 10 to about 20 nucleotides in length. In some embodiments, siRNA molecules disclosed herein may have a nucleotide sequence that is about 10 nucleotides in length, about 11 nucleotides in length, about 12 nucleotides in length, about 13 nucleotides in length, about 14 nucleotides in length, about 15 nucleotides in length, about 14 nucleotides in length, about 15 nucleotides in length, about 16 nucleotides in length, about 17 nucleotides in length, about 18 nucleotides in length, about 19 nucleotides in length, about 20 nucleotides in length, about 21 nucleotides

in length, about 22 nucleotides in length, about 23 nucleotides in length, about 24 nucleotides in length, or about 25 nucleotides in length. Each complementary sequence of a double-stranded siRNA disclosed herein may be about 2-60, about 5-50, about 10-40, about 2-30, about 5-25, or about 10-25 nucleotides in length, but other noncomplementary sequences may be present. For example, siRNA duplexes may comprise 3' overhangs of about 1 to about 4 or more nucleotides and/or 5' phosphate termini comprising about 1 to about 4 or more nucleotides.

[0060] In certain embodiments, siRNA molecules disclosed herein may have about 2-60, about 2-50, about 2-40, or about 2-30 contiguous nucleotides of homology with a target (e.g., TBX3) nucleotide sequence. In some aspects, a target nucleotide sequence herein may be a human TBX3 nucleotide sequence or a variant thereof. In some other aspects, a target nucleotide sequence herein may be human TBX3, RefSeq: NM\_005996.4, or a variant thereof. In still some other aspects, a target nucleotide sequence herein may be SEQ ID NO: 1 as follows, or a variant thereof:

Human TBX3 (SEQ ID NO: 1):  
AGCCTCTCCATGAGAGATCCGGTCATTCTGGACAAGCATGGCCTACCATCCGTTCT  
ACCTCACCGGGCGCCGGACTTCGCCATGAGCGCGTGCTGGGTACCAGCCGCCGTTCT  
TCCCCGGCGCTGAGCCTGCCCAACGGCGCGGCCGCTCTCGCTGCCGGCGCCCTG  
GCCAAGCCGATCATGGATCAATTGGTGGGGGGCGGCCGAGACGGCATCCCGTTCTCTC  
CCTGGGGCCCCAGGCGCATCTGAGGCCCTTGAAGACCATGGAGCCGAAGAAGAGGTGG  
AGGACGACCCCAGGTGCACCTGGAGGCTAAAGAACTTGGGATCAGTTCACAAAGCGG  
GGCACCGAGATGGTCATTACCAAGTCGGGAAGGCGAATGTTCTCCATTAAAGTGAG  
ATGTTCTGGGCTGGATAAAAAGCCAATACATTATTGATGGACATTAGCTGCTG  
ATGACTGTCGTTATAAATTTCACAATTCTCGTGGATGGTGGCTGGTAAGGCCGACCC  
GAAATGCCAAAGAGGATGTACATTACCCGGACAGCCCCGTAUTGGGAAACAGTGGAT  
GTCCAAGTCGTCACTTCCACAAACTGAAACTCACCAACAACATTCAAGACAAACATG  
GATTACTTTGGCTTCCAAGTGATCACGCTACGTCAGGGAAATTAGTTGGT  
ACTCAGACTATATTGAACTCCATGCACAAATCAGGCCGGTCCACATTGTAAGAGC  
CAATGACATCTGAAACTCCCTATAGTACATTGGACATACTGTTCCCGAAACTG  
AATTCACTGCTGTGACTGCATACCAAGATAAGATAACCCAGTTAAAAATAGACAAAC  
AACCTTTGCAAAGGTTCCGGACACTGGAAATGCCGAAGAGAAAAAGAAAAAC  
GCTCACCTGCAGTCCATGAGGGTGTGATGAAAGACACAAAAAGGAGAATGGACCT  
CTGATGAGTCCTCCAGTGAACAAGCAGCTTCAACTGCTCGCCAGGCTTCTC  
GCCGCCTCCACTGTAGGGACATCGAACCTCAAAGATTATGTCGGAGGGTGAG  
CGACGCCAGGCCAGAGCAAAGAGGAGCATGGCCCCAGGCCCTGCGACGCCAG  
TCTCCACCAACGTCGGAGGCCCTGCCGTGACAAGGGCAGCCCCGCCGTCAGG  
CACCTTTGCTGCTGAGCGGCCCGGACAGCGGGGGCTGGACAAGCGTCGCC  
CTCACGCCATAGCCCCGCCACCATCTCGTCCAGCACTCGCCGCTGGCGCGAG  
GCAGGAGCCGGTTCGGAGGGCACAGCGCCGCCAAGGTGGAAGAGGCCGCGC  
CCGGCAAGGAGGCCTTCGGCGCCGTCACGGTGCAGACGGACGCCGCCGCGCACCT

- continued

```

GCCCGAGGGCCCCCTGCTGGCTCGGCTCGCCCGGGCTGGCGGGCAAACAGTTCT
TCAACGGGCACCCGCTTCCTGCACCCCAGCCAGTTGCCATGGGGGCGCCTCTCC
AGCATGGCGGCCGCTGGCATGGTCCCCCTGGCACGGTTCTGGGCCTCCACCGG
TGCTCGGGCTGGATTCCACGGCATGGCTCTGCCGTGGGGCAGGGACTGTCCG
GGCGTCCGGCCACCCCTGCCCTCCACCTCAGCAGCACGTCCTGGCCTCAGGGC
CTGGCCATGTCCCCTTCGGAAGCCTGTTCCCTAACCTACACGTACATGGCCGCAGC
GGCGGCCGCCTCTCGGGCAGCCTCCAGCTGGTGCACGCCACCCCTCCTCAATC
TGAACACCATGCGCCCGGGCTGCGCTACAGCCCCTACTCCATCCCGTGCGGTCCC
GACGGCAGCAGTCTGCTCACCCGCCCTGCCCTCATGGCGGCCGCGGGGCCCT
GGACGGCAAAGTCGCCGCCCTGGCGCCAGCCGGCTCGTGGCAGTGGACTCGGGCT
CTGAACCTAACAGCCGCTCCCTCACGCTCTCCAGCTCATGTCCCTGTGCGCCAAA
CTCTGCGCGGAGAAAGAGGCCACCAGCGAAGTGCAGAGCATCCAGCGTTGGTTAG
CGGCTTGAAGCCAAGCCGGACAGGTCCCGCAGCGCGTCCCCG

```

**[0061]** In certain embodiments, siRNA molecules disclosed herein may have a nucleotide sequence that is about 2-30 contiguous in length and is at least about 80% homologous (e.g., about 80%, about 85%, about 90%, about 95%, or about 99%) to at least about 2 to about 30 contiguous nucleotides of a human TBX3 cDNA sequence. In some embodiments, siRNA molecules disclosed herein may have a nucleotide sequence that is about 2-30 contiguous in length and is at least about 80% homologous (e.g., about 80%, about 85%, about 90%, about 95%, or about 99%) to at least about 2 to about 30 contiguous nucleotides of SEQ ID NO: 1. In some other embodiments, siRNA molecules disclosed herein may have a nucleotide sequence that is about 2-30 contiguous in length and is homologous to at least about 2 to about 30 contiguous nucleotides of SEQ ID NO: 1.

**[0062]** A siRNA molecule disclosed herein may be synthesized in any number of conformations. One skilled in the art would recognize the type of siRNA conformation to be used for a particular purpose. Examples of siRNA confor-

mations include, but need not be limited to, a double-stranded polynucleotide molecule assembled from two separate stranded molecules, wherein one strand is the sense strand and the other is the complementary antisense strand; a double-stranded polynucleotide molecule assembled from a single-stranded molecule, where the sense and antisense regions are linked by a nucleic acid-based or non-nucleic acid-based linker; a double-stranded polynucleotide molecule with a hairpin secondary structure having complementary sense and antisense regions; or a circular single-stranded polynucleotide molecule with two or more loop structures and a stem having self-complementary sense and antisense regions. In the case of the circular polynucleotide, the polynucleotide may be processed either *in vivo* or *in vitro* to generate an active double-stranded siRNA molecule.

**[0063]** In certain embodiments, siRNA molecules disclosed herein may be double stranded siRNA molecules. In some aspects, double stranded siRNA molecules disclosed herein may have at least one sense sequence. In some other aspects, double stranded siRNA molecules disclosed herein may have at least one antisense sequence. In still some other aspects, double stranded siRNA molecules disclosed herein may have at least one antisense sequence and at least one sense sequence. In some embodiments, double stranded siRNA molecules disclosed herein may have at least one antisense sequence selected from Table 1, at least one sense sequence selected from Table 1, or both.

TABLE 1

siRNA ID	Sense Sequence (5'-3')	Over-hang	Antisense Sequence (5'-3')	Over-hang
TBX3_470	GCUGAUGACUGUCGUUAUA (SEQ ID NO: 2)	UU	UAAUACGACAGUCAUCAGC (SEQ ID NO: 59)	UU
TBX3_1080	GGACAUCAACCUAAAGA (SEQ ID NO: 3)	UU	UCUUUGAGGUUCGAUGUCC (SEQ ID NO: 60)	UU
TBX3_457	GGACAUUAUAGCUGCUGAU (SEQ ID NO: 4)	UU	AUCAGCAGCUAAUAGUCC (SEQ ID NO: 61)	UU
TBX3_707	GGUACUCAGACUAAUUGA (SEQ ID NO: 5)	UU	UCAAUAUAGCUGAGUACC (SEQ ID NO: 62)	UU

TABLE 1-continued

sirNA ID	Sense Sequence (5'-3')	Over- hang	Antisense Sequence (5'-3')	Over- hang
TBX3_754	CCACAUUGUAAGAGCCAAU (SEQ ID NO: 6)	UU	AUUGGCUCUUAACAAUGUGG (SEQ ID NO: 63)	UU
TBX3_966	GGGUGUUUGAUGAAAGACA (SEQ ID NO: 7)	UU	UGUCUUUCAUCAACACCC (SEQ ID NO: 64)	UU
TBX3_631	CAACAUUUCAGACAAACAU (SEQ ID NO: 8)	UU	AUGUUUGUCUGAAAUGUUG (SEQ ID NO: 65)	UU
TBX3_628	CAACAACAUUUCAGACAAA (SEQ ID NO: 9)	UU	UUUGUCUGAAAUGUUGUUG (SEQ ID NO: 66)	UU
TBX3_627	CCAACAACAUUUCAGACAA (SEQ ID NO: 10)	UU	UUGUCUGAAAUGUUGUUGG (SEQ ID NO: 67)	UU
TBX3_857	GAUAGAUAAACCCAGUUAA (SEQ ID NO: 11)	UU	UUAACUGGGUUUAUCUUUAUC (SEQ ID NO: 68)	UU
TBX3_389-411	CGAAUGUUUCUCUCCAUUUA (SEQ ID NO: 12)	AA	AAAUGGAGGAAACAUUCG (SEQ ID NO: 69)	CC
TBX3_390-412	GAAUGUUUCUCUCCAUUUA (SEQ ID NO: 13)	AG	UUAAAUGGAGGAAACAUUC (SEQ ID NO: 70)	UU
TBX3_438-460	CCAAAUAACAUUUUAUUGAU (SEQ ID NO: 14)	UU	AUCAAUAAAAUGUAUUGG (SEQ ID NO: 71)	CU
TBX3_619-641	GAAACUCACCAACAACAUU (SEQ ID NO: 15)	UC	AAUGUUGUJGGUGAGUUUC (SEQ ID NO: 72)	AG
TBX3_768-790	CCAAUGACAUUCUUGAAACU (SEQ ID NO: 16)	CC	AGUUUCAAGAUGUCAUUGG (SEQ ID NO: 73)	CU
TBX3_917-939	GGAAAUGGCCGAAGAGAAA (SEQ ID NO: 17)	AA	UUUCUCUUCCGGCAUUUCC (SEQ ID NO: 74)	AG
TBX3_428-450	GAUAAAAAAGCCAAAUAACA (SEQ ID NO: 18)	UU	UGUAUUUGGCCUUUUUAUC (SEQ ID NO: 75)	CA
TBX3_437-459	GCCAAAUACAUUUUAUUGA (SEQ ID NO: 19)	UG	UCAAUAAAAUGUAUUGGC (SEQ ID NO: 76)	UU
TBX3_778-800	CUUGAAACUCUCCUUUAUGU (SEQ ID NO: 20)	AC	ACUUAAGGGAGUUUCAAG (SEQ ID NO: 77)	AU
TBX3_860-882	AAGAUAAACCCAGUUAAAAAA (SEQ ID NO: 21)	UU	UUUUUAACUGGGUUUAUCUU (SEQ ID NO: 78)	AU
J-012197-05	CAUCGAACCUCAAAGAUUU (SEQ ID NO: 22)	UU	AAAUCUUUGAGGUUCGAUG (SEQ ID NO: 79)	UU
J-012197-06	CCAAAGAGGAUGUACAUUC (SEQ ID NO: 23)	UU	GAAUGUACAUCCUCUUUGG (SEQ ID NO: 80)	UU
J-012197-07	CAUUAUAGCUGCGUGAUGAC (SEQ ID NO: 24)	UU	GUCAUCAGCAGCUAAUAG (SEQ ID NO: 81)	UU
J-012197-08	UAUAGUACAUUUUCGGACAU (SEQ ID NO: 25)	UU	AUGUCCGAAAUGUACUUA (SEQ ID NO: 82)	UU
D-012197-01	CAUCGAACCUCAAAGAUUU (SEQ ID NO: 26)	UU	AAAUCUUUGAGGUUCGAUG (SEQ ID NO: 83)	UU
D-012197-02	GCUGAUGACUGUCGUUUA (SEQ ID NO: 27)	UU	UAUACGACAGUCAUCAGC (SEQ ID NO: 84)	UU
D-012197-03	CCAAUGACAUUCUUGAAACU (SEQ ID NO: 28)	UU	AGUUUCAAGAUGUCAUUGG (SEQ ID NO: 85)	UU
D-012197-04	CCAUUUAAAGUGAGAUGUU (SEQ ID NO: 29)	UU	AACAUUCACUUUUAAUGG (SEQ ID NO: 86)	UU

TABLE 1-continued

sirNA ID	Sense Sequence (5'-3')	Over-hang	Antisense Sequence (5'-3')	Over-hang
s13865	GAAACUCCUUUAUAGUACA (SEQ ID NO: 30)	TT	UGUACUAUAAGGGAGUUUC (SEQ ID NO: 87)	AA
s13866	CCAACAACAUUCAGACAA (SEQ ID NO: 31)	TT	UUGUCUGAAAUGUUGUUGG (SEQ ID NO: 88)	TG
s13867	GCUGAUGACUGUCGUUAUA (SEQ ID NO: 32)	TT	UAUAACGACAGUCAUCAGC (SEQ ID NO: 89)	AG
s529425	AAGUGAGACUAUUAGACAA (SEQ ID NO: 33)	TT	UUGUCUAUAAGUCUCACUU (SEQ ID NO: 90)	CT
s534311	CCAUUAAAAGUGAGAUGUU (SEQ ID NO: 34)	TT	AACAUCUCACUUAAAUGG (SEQ ID NO: 91)	AG
s534312	GAUAAACCCAGUUAAAAAUA (SEQ ID NO: 35)	TT	UAUUUUUAACUGGGUUAUC (SEQ ID NO: 92)	TT
107921	GGGUGUUUUGAUGAAAGACA (SEQ ID NO: 36)	TT	UGUCUUUCAUCAACACCC (SEQ ID NO: 93)	TC
107922	GGCUCACCUUUUCUGCUGCU (SEQ ID NO: 37)	TT	AGCAGCGAAAAGGUGAGCC (SEQ ID NO: 94)	TT
115750	GCAGCUUUCAACUGCUUCG (SEQ ID NO: 38)	TT	CGAACAGUUGAAAGCUGC (SEQ ID NO: 95)	TT
115751	CCUCAAAGAUUUUAUGUCCC (SEQ ID NO: 39)	TT	GGGACAUAAAUCUUUGAGG (SEQ ID NO: 96)	TT
139810	GGAGUGGGCAAAACACAU (SEQ ID NO: 40)	TT	UAUGUGUUUUGCCCACUCC (SEQ ID NO: 97)	TT
239194	UGGCCGAAGAGAAAAAAAAGA (SEQ ID NO: 41)	TT	UCUUUUUUUCUUCGGCCA (SEQ ID NO: 98)	TT
239195	CAAGCAGCUUCAACUGCU (SEQ ID NO: 42)	TT	AGCAGUUGAAAGCUGCUUG (SEQ ID NO: 99)	TT
239196	AACUCCCUUUAUAGUACAU (SEQ ID NO: 43)	TT	AAUGUACUAUAAGGGAGUU (SEQ ID NO: 100)	TC
3661	GGCUUUGAAGACCAUGGA (SEQ ID NO: 44)	TT	UCCAUGGUCUUCAAAGGCC (SEQ ID NO: 101)	TC
3754	GGUGCACCUUGAGGCCAAA (SEQ ID NO: 45)	TT	UUUAGCCUCCAGGUGCACC (SEQ ID NO: 102)	TT
3845	GGCUAAAGAACUUUGGGAU (SEQ ID NO: 46)	TT	AUCCCAAAGUUCUUUAGCC (SEQ ID NO: 103)	TC
TBX3 186	CGAUCAUGGAUCAAUUGGU (SEQ ID NO: 47)	UU	ACCAAUUGAUCCAUGAUCG (SEQ ID NO: 104)	UU
TBX3 454	GAUGGACAUUAUAGCUGCU (SEQ ID NO: 48)	UU	AGCAGCUUAAUGUCCAUC (SEQ ID NO: 105)	UU
TBX3 589	GAUGUCCAAGUGUCACU (SEQ ID NO: 49)	UU	AGUGACGACUUUGGACAUC (SEQ ID NO: 106)	UU
TBX3 540	CAAAGAGGAUGUACAUUCA (SEQ ID NO: 50)	UU	UGAAUGUACAUCCUCUUUG (SEQ ID NO: 107)	UU
TBX3 781	GAAACUCCUUUAUAGUACA (SEQ ID NO: 51)	UU	UGUACUAUAAGGGAGUUUC (SEQ ID NO: 108)	UU
TBX3 713	CAGACAUUAUUGAACUCCA (SEQ ID NO: 52)	UU	UGGAGUUCAAAUAGUCUG (SEQ ID NO: 109)	UU
TBX3 368	GUCAUUACCAAGUCGGAA (SEQ ID NO: 53)	UU	UUCCGACUUGGUAAUGAC (SEQ ID NO: 110)	UU

TABLE 1-continued

siRNA ID	Sense Sequence (5'-3')	Over- hang	Antisense Sequence (5'-3')	Over- hang
TBX3_324	CUAAAGAACUUUGGGAUCA (SEQ ID NO: 54)	UU	UGAUCCCCAAAGUUCUUUAG (SEQ ID NO: 111)	UU
TBX3_788	CCUUUAUAGUACAUUUCGGA (SEQ ID NO: 55)	UU	UCCGAAAUGUACUAUAAGG (SEQ ID NO: 112)	UU
TBX3_717	CUAUAUUGAACUCCAUGCA (SEQ ID NO: 56)	UU	UGCAUGGGAGGUUCAAUAUAG (SEQ ID NO: 113)	UU
TBX3_820	CGAAACUGAAUUCAUCGCU (SEQ ID NO: 57)	UU	AGCGAUGGAAUUCAGUUUCG (SEQ ID NO: 114)	UU
TBX3_1091	CUCAAAGAUUUUAUGUCCCA (SEQ ID NO: 58)	UU	UGGGACAUAAAUCUUUGAG (SEQ ID NO: 115)	UU

**[0064]** In some embodiments, double stranded siRNA molecules disclosed herein may have at least one sense sequence having at least about 80% homology (e.g., about 80%, about 85%, about 90%, about 95%, or about 99%) to any one of SEQ ID NOs: 2-58. In some embodiments, double stranded siRNA molecules disclosed herein may have at least one antisense sequence having at least about 80% homology (e.g., about 80%, about 85%, about 90%, about 95%, or about 99%) to any one of SEQ ID NOs: 59-115. In some embodiments, double stranded siRNA molecules disclosed herein may have at least one sense sequence having at least about 80% homology (e.g., about 80%, about 85%, about 90%, about 95%, or about 99%) to any one of SEQ ID NOs: 2-58 and at least one antisense sequence having at least about 80% homology (e.g., about 80%, about 85%, about 90%, about 95%, or about 99%) to any one of SEQ ID NOs: 59-115. In some embodiments, double stranded siRNA molecules disclosed herein may have at least one sense sequence of SEQ ID NOs: 2-58. In some embodiments, double stranded siRNA molecules disclosed herein may have at least one antisense sequence of SEQ ID NOs: 59-115. In some embodiments, double stranded siRNA molecules disclosed herein may have at least one sense sequence of SEQ ID NOs: 2-58 and at least one antisense sequence of SEQ ID NOs: 59-115. In some aspects, double stranded siRNA molecules disclosed herein may have at least one sense sequence of SEQ ID NOs: 2-21 and at least one antisense sequence of SEQ ID NOs: 59-78. In some other aspects, double stranded siRNA molecules disclosed herein may have at least one sense sequence of SEQ ID NOs: 47-58 and at least one antisense sequence of SEQ ID NOs: 104-155. In some aspects, double stranded siRNA molecules disclosed herein may have a sense sequence of SEQ ID NO: 3 and an antisense sequence of SEQ ID NO: 60. In some aspects, double stranded siRNA molecules disclosed herein may have a sense sequence of SEQ ID NO: 18 and an antisense sequence of SEQ ID NO: 75. In some aspects, double stranded siRNA molecules disclosed herein may have a sense sequence of SEQ ID NO: 19 and an antisense sequence of SEQ ID NO: 76.

**[0065]** In some embodiments, the present disclosure also encompasses nucleic acid sequences encoding a double stranded siRNA molecule as disclosed herein. In some embodiments, a nucleic acid sequence encoding the double stranded siRNA as disclosed herein may be a double stranded DNA, single stranded DNA, a plasmid vector, a

viral vector for example a retroviral vector, lentiviral vector, a pox viral vector, an adenoviral vector, or an adeno-associated viral vector. In some embodiments, the nucleic acid may encode at least one sense sequence having at least about 80% homology (e.g., about 80%, about 85%, about 90%, about 95%, or about 99%) to any one of SEQ ID NOs: 2-58. In some embodiments, the nucleic acids may encode at least one antisense sequence having at least about 80% homology (e.g., about 80%, about 85%, about 90%, about 95%, or about 99%) to any one of SEQ ID NOs: 59-115. In some embodiments, the nucleic acids may encode at least one sense sequence having at least about 80% homology (e.g., about 80%, about 85%, about 90%, about 95%, or about 99%) to any one of SEQ ID NOs: 2-58 and at least one antisense sequence having at least about 80% homology (e.g., about 80%, about 85%, about 90%, about 95%, or about 99%) to any one of SEQ ID NOs: 59-115. In some embodiments, the nucleic acids may encode at least one sense sequence of any one of SEQ ID NOs: 2-58. In some embodiments, the nucleic acids may encode at least one antisense sequence of any one of SEQ ID NOs: 59-115. In some embodiments, the nucleic acids may encode at least one sense sequence of SEQ ID NOs: 2-58 and at least one antisense sequence of SEQ ID NOs: 59-115. In some embodiments, the nucleic acid may encode a sense sequence of SEQ ID NO: 3 and an antisense sequence of SEQ ID NO: 60. In some aspects, the nucleic acid may encode a sense sequence of SEQ ID NO: 18 and an antisense sequence of SEQ ID NO: 75. In some aspects, nucleic acid molecule may encode a sense sequence of SEQ ID NO: 19 and an antisense sequence of SEQ ID NO: 76.

**[0066]** The present disclosure also provides for methods of making the siRNA molecules disclosed herein. Making of siRNA molecules disclosed herein may be through chemical synthesis or siRNA molecules disclosed herein may be encoded by a plasmid and transcribed or may be vectored by a virus engineered to express the siRNA. A siRNA molecule disclosed herein may be a single stranded molecule with complementary sequences that self-hybridize into duplexes with hairpin loops. An siRNA molecule disclosed herein can also be generated by cleavage of parent dsRNA through the use of an appropriate enzyme such as *E. coli* RNase III or Dicer. A parent dsRNA may be any double stranded RNA duplex from which a siRNA may be produced, such as a full or partial mRNA transcript. Use of cell lysates or in vitro processing may further involve the subsequent isolation of

the short, nucleotide siRNAs (e.g., about 2-25 nucleotides in length) from the lysate, etc., making the process somewhat cumbersome and expensive. Chemical synthesis proceeds by making two single stranded RNA-oligomers followed by the annealing of the two single stranded oligomers into a double stranded RNA. Methods of chemical synthesis are diverse and can be easily adapted to synthesis different siRNAs.

[0067] In certain embodiments, siRNA molecules disclosed herein may abolish gene expression of TBX3. Methods known in the art for the detection and quantification of RNA expression suitable for use herein can include, but are not limited to northern blotting and in situ hybridization, RNase protection assays, polymerase chain reaction (PCR), reverse transcription polymerase chain reaction (RT-PCR), real-time quantitative reverse transcription PCR (RT-qPCR or qPCR), sequencing-based gene expression analysis (e.g., Serial Analysis of Gene Expression (SAGE)), gene expression analysis by massively parallel signature sequencing (MPSS), and the like.

[0068] The siRNA molecules of the present disclosure specifically downregulate gene expression of TBX3 (T-Box Transcription Factor 3) or a variant thereof. The term downregulating may be interchangeably expressed as reducing, inhibiting, preventing, blocking or silencing. Here, the phrase "downregulating gene expression" refers to any reduced level of gene expression in comparing with an ordinary expression level. For example, the reduced level of gene expression can be from about 70% to about 0% of the ordinary expression level. In other words, about 30% to about 100% gene expression is downregulated, reduced, blocked, inhibited, prevented or silenced, comparing to the ordinary expression level. Specifically, the reduced level of gene expression is about 70%, 65%, 60%, 55%, 50%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, 8%, 5%, 4%, 3%, 2%, 1% or 0% of the ordinary expression level of TBX3. The reduced level of gene expression can also be any percentage or range as recited above. On the other hand, the term "specific" or "specifically" used in combination with downregulating refers to downregulation of a target gene's expression with minimal or no binding or downregulation of other nucleic acids or their expressions.

[0069] In some embodiments, siRNA molecules disclosed herein may reduce gene expression of TBX3 by at least about 50%. In some aspects, siRNA molecules disclosed herein may reduce gene expression of TBX3 by about 50% to about 99%, about 55% to about 98%, or about 60% to about 95%. In some aspects, siRNA molecules disclosed herein may reduce gene expression of TBX3 by about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 95%, or about 99%.

[0070] In certain embodiments, siRNA molecules disclosed herein may abolish protein expression of TBX3. Methods known in the art for the detection and quantification of protein expression suitable for use herein can include, but are not limited to ELISAs (enzyme-linked immunosorbent assays), immunoblot assays, flow cytometric assays, immunohistochemical assays, radioimmuno assays, Western blot assays, an immunofluorescent assays, chemiluminescent assays, mass spectrometry assays, matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass mapping, liquid chromatography/quadrupole time-of-flight electrospray ionization tandem mass spectrometry (LC/Q-

TOF-ESI-MS/MS), two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), and the like.

[0071] In some embodiments, siRNA molecules disclosed herein may reduce protein expression of TBX3 by at least about 50%. In some aspects, siRNA molecules disclosed herein may reduce protein of TBX3 by about 50% to about 99%, about 55% to about 98%, or about 60% to about 95%. In some other aspects, siRNA molecules disclosed herein may reduce protein expression of TBX3 by about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 95%, or about 99%.

[0072] In some embodiments, siRNA molecules disclosed herein may have one or more chemical modifications. Non-limiting examples of chemical modifications can include terminal cap moieties, phosphate backbone modifications, and the like. Examples of classes of terminal cap moieties include, without limitation, inverted deoxy abasic residues, glycerol modifications, 4',5'-methylene nucleotides, 1-( $\beta$ -D-erythrofuranosyl) nucleotides, 4'-thio nucleotides, carbocyclic nucleotides, 1,5-anhydrohexitol nucleotides, L-nucleotides,  $\alpha$ -nucleotides, modified base nucleotides, threo pentofuranosyl nucleotides, acyclic 3',4'-seco nucleotides, acyclic 3,4-dihydroxybutyl nucleotides, acyclic 3,5-dihydroxypentyl nucleotides, 3'-3'-inverted nucleotide moieties, 3'-3'-inverted abasic moieties, 3'-2'-inverted nucleotide moieties, 3'-2'-inverted abasic moieties, 5'-5'-inverted nucleotide moieties, 5'-5'-inverted abasic moieties, 3'-5'-inverted deoxy abasic moieties, 5'-amino-alkyl phosphate, 1,3-diamino-2-propyl phosphate, 3 aminopropyl phosphate, 6-aminohexyl phosphate, 1,2-aminododecyl phosphate, hydroxypropyl phosphate, 1,4-butanediol phosphate, 3'-phosphoramidate, 5'-phosphoramidate, hexylphosphate, aminohexyl phosphate, 3'-phosphate, 5'-amino, 3'-phosphorothioate, 5'-phosphorothioate, phosphorodithioate, and bridging or non-bridging methylphosphonate or 5'-mercapto moieties. Non-limiting examples of phosphate backbone modifications (i.e., resulting in modified internucleotide linkages) include phosphorothioate, phosphorodithioate, methylphosphonate, phosphotriester, morpholino, amide, carbamate, carboxymethyl, acetamide, polyamide, sulfonate, sulfonamide, sulfamate, formacetal, thioformacetal, and alkylsilyl substitutions. Such chemical modifications can occur at the 5'-end and/or 3'-end of the sense strand, antisense strand, or both strands of a siRNA molecule disclosed herein.

[0073] Chemical modification of a siRNA molecule disclosed herein may comprise modification of at least one ribosugar moiety of its nucleotide sequence. The ribosugar moiety may be modified with 2'-O-methyl (2'OMe), 2'-deoxy-2'-fluoro (2'F), 2'-deoxy, 5-C-methyl, 2'-O-(2-methoxyethyl) (MOE), 4'-thio, 2'-amino, 2'-C-allyl, or any combination thereof. In some aspects, a siRNA molecule disclosed herein may have less than about 10% to about 70% (e.g., about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%) of ribosugar moieties of the total nucleotide sequence modified.

[0074] Chemical modification of a siRNA molecule disclosed herein may comprise attaching a conjugate to the siRNA molecule. The type of conjugate used and the extent of conjugation to the siRNA can be evaluated for improved pharmacokinetic profiles, bioavailability, and/or stability of the siRNA while retaining activity. As such, one skilled in the art can screen siRNA molecules having various conjugates attached thereto to identify siRNA conjugates having

improved properties using any of a variety of well-known in vitro cell culture or in vivo animal models including the negative-controlled expression studies described above. The conjugate can be attached at the 5'- and/or the 3'-end of the sense and/or the antisense strand of the siRNA via a covalent attachment such as a nucleic acid or non-nucleic acid linker. The conjugate can be attached to the siRNA through a carbamate group or other linking group (see, e.g., U.S. Patent Publication Nos. 20050074771, 20050043219, and 20050158727, the content of each of which is incorporated by reference herein in its entirety). A conjugate may be added to siRNA for any of a number of purposes. For example, the conjugate may be a molecular entity that facilitates the delivery of siRNA into a cell or may be a molecule that comprises a drug or label. Examples of conjugate molecules suitable for attachment to siRNA of the present invention include, without limitation, steroids such as cholesterol, glycals such as polyethylene glycol (PEG), human serum albumin (HSA), fatty acids, carotenoids, terpenes, bile acids, folates (e.g., folic acid, folate analogs and derivatives thereof), sugars (e.g., galactose, galactosamine, N-acetyl galactosamine, glucose, mannose, fructose, fucose, etc.), phospholipids, peptides, ligands for cellular receptors capable of mediating cellular uptake, and combinations thereof. Other examples include the lipophilic moiety, vitamin, polymer, peptide, protein, nucleic acid, small molecule, oligosaccharide, carbohydrate cluster, intercalator, minor groove binder, cleaving agent, and cross-linking agent conjugate molecules described in e.g., U.S. Patent Publication Nos. 20050119470 and 20050107325, the content of each of which is incorporated by reference herein in its entirety. Other examples include the 2'-O-alkyl amine, 2'-O-alkoxy-alkyl amine, polyamine, C5-cationic modified pyrimidine, cationic peptide, guanidinium group, amidinium group, cationic amino acid conjugate molecules, and the like. Additional examples of conjugate molecules include a hydrophobic group, a membrane active compound, a cell penetrating compound, a cell targeting signal, an interaction modifier, or a steric stabilizer as described in U.S. Patent Publication No. 20040167090, incorporated by reference herein in its entirety.

**[0075]** In certain embodiments, siRNA molecules disclosed herein may be conjugated to at least one targeting ligand. Targeting ligands contemplated herein include ligands suitable for targeting siRNA molecules to a liver, a liver tissue, and/or a liver cell. Non-limiting examples of targeting ligands suitable for use herein may include galactose, galactosamine, N-formyl-galactosamine, N-acetylgalactosamine, N-propionyl-galactosamine, N-n-butanoyle-galactosamine, N-iso-butanoylegalactos-amine, galactose cluster, and N-acetylgalactosamine trimer and may optionally have a pharmacokinetic modulator selected from the group consisting of: hydrophobic group having 16 or more carbon atoms, hydrophobic group having 16-20 carbon atoms, palmitoyl, hexadec-8-enoyl, oleyl, (9E,12E)-octadeca-9,12dienoyl, dioctanoyl, and C16-C20 acyl, and cholesterol. In some embodiments, a liver targeting ligand suitable for use herein may be a N-Acetylgalactosamine (GalNAc) conjugate. In some embodiments, siRNA molecules disclosed herein may be conjugated to at least one GalNAc conjugate. In some embodiments, siRNA molecules disclosed herein may be conjugated to about 1 to about 10 GalNAc conjugates, about 2 to about 9 GalNAc conjugates, or about 3 to about 8 GalNAc conjugates. In

some embodiments, siRNA molecules disclosed herein may be conjugated to about 1, about 2, about 3, about 4, about 5, about 6, about 7, about 8, about 9, or about 10 GalNAc conjugates.

**[0076]** Any of the siRNA molecules disclosed herein may target the open reading frame or the 5' or 3' UTRs of the TBX3 gene or at least one variant thereof.

**[0077]** siRNA molecules disclosed herein can specifically downregulate gene expression of at least one variant of TBX3. The TBX3 gene and/or its variant may be associated with a liver disease, liver dysfunction, liver injury, and/or liver damage. The liver disease may comprise fatty liver disease (FLD), alcohol-related liver disease (ARLD), non-alcoholic fatty liver disease (NAFLD), non-alcoholic steatohepatitis (NASH), end stage liver disease (cirrhosis) from any etiology, liver cancer, or any combination thereof. Liver damage may include cirrhosis, chronic infection of hepatitis B virus (HBV), chronic infection of hepatitis C virus (HCV), non-alcoholic fatty liver disease (NAFLD), non-alcoholic steatohepatitis (NASH), primary biliary cirrhosis (PBC), hereditary hemochromatosis, type 2 diabetes, obesity, tobacco use, alcohol abuse, long-term anabolic steroid use, tyrosinemia, alpha1-antitrypsin deficiency, porphyria cutanea tarda, glycogen storage diseases, Wilson disease, or any combination thereof. In some embodiments, a suitable subject for the methods herein may have or be suspected of having one or more injuries to the liver that may predispose a subject to HCC.

**[0078]** In certain embodiments, methods and compositions provided herein can include a vector containing any one of the siRNA molecules disclosed herein. In some embodiments, a vector for use herein can be a viral vector. As used herein, the term "viral vector" can refer to a nucleic acid vector construct that includes at least one element of viral origin and has the capacity to be packaged into a viral vector particle and encodes at least an exogenous polynucleotide. In certain embodiments, the vector and/or particle can be utilized for the purpose of transferring any nucleic acids into cells either in vitro or in vivo. Numerous viral vectors are known in the art. The term virion can refer to a single infective viral particle. "Viral vector", "viral vector particle" and "viral particle" also refer to a complete virus particle with its DNA or RNA core and protein coat as it exists outside the cell. Non-limiting examples of viral vectors for use herein can include adenoviruses, adeno-associated viruses (AAV), herpesviruses, retroviruses, lentiviruses, integrase defective lentiviruses (IDLV), and the like. In some embodiments, a viral vector disclosed herein can be a lentiviral vector. Examples of lentiviruses include, but are not limited to, human lentiviruses such as HIV (in particular HIV-1 or HIV-2), simian immunodeficiency virus (SIV), equine infectious anemia virus (EIAV), feline immunodeficiency virus (FIV), Caprine Arthritis Encephalitis Virus (CAEV), visna and progressive pneumonia viruses of sheep, baboon pseudotype viruses, bovine immunodeficiency virus (BIV), and the like. In some embodiments, siRNA molecules and/or vectors described herein can be prepared by conventional recombinant technology known to one of skill in the art. In other embodiments, siRNA molecules and/or vectors described herein can be prepared by a gene editing method known in the art (e.g., by CRISPR). In certain embodiments, methods provided herein can include generating a cell to express any of the siRNA molecules and/or vectors described herein. In some embodiments, vectors,

viral particles, and the like as contemplated herein may be encapsulated into a liposome for delivery to a subject.

[0079] In some aspects, the current disclosure also encompasses use of gene editing systems for example CRISPR based systems for abolishing or downregulating gene expression of TBX3. As such, the current disclosure also encompasses compositions comprising a nucleic acid sequence related to a cluster regularly interspaced short palindromic repeats (CRISPR) system, for example a single guide RNA (sgRNA), a CRISPR-RNA (crRNA), a trans-activating crRNA (tracrRNA), and further a plasmid DNA (pDNA) or a viral vector encoding nucleic acid sequence related to a cluster regularly interspaced short palindromic repeats (CRISPR) system, that specifically target TBX3 or

variants thereof. In some aspects, the gene editing system comprises at least an sgRNA targeting TBX3 or a UTR thereof and an RNA guided endonuclease for example Cas9. In some aspects, the gene editing system comprises at least an sgRNA targeting TBX3 and an RNA guided endonuclease for example Cas9. In some aspects, the sgRNA target sequence is at least about 80% identical to any one of SEQ ID NO: 118-217 as provided in Table 2. In some aspects, the sgRNA targets a sequence at least about 80%, or about 85%, or about 90%, or about 95% or about 100% identical to SEQ ID NO: 118-217. In some aspects, the sgRNA targets 20 consecutive nucleotides in any one of SEQ ID NO: 218-317. None of the gRNA as provided in Table 2 are predicted to target the last exon or have 10 bp off target match.

TABLE 2

gRNA Sequence	SEQ ID NOS	Guide with flanking sequence (10 bp before and after the guide + PAM)	SEQ ID NOS	cut position in chromosome	On-target efficiency	Exon	Off-target score
GAAAAGGTGA GCCTTGACCG	118	GCTCAGCAGC GAAAAGGTGA GCCTTGACCG CGGCGGGGC TGCC	218	115112508	0.714618143	7	0
GCCCCGCTAC TGGGAAACA G	119	CACCCGGACA GCCCGCCTAC TGGGAAACA GTGGTGGATG TCCA	219	115118882	0.711616088	2	0.353571
CAAGATCTCC ACCACCACGT	120	GCGACGCGG CCAAGATCTC CACCAACCACG TCGGCGGAG GAGCC	220	115112068	0.708326443	7	0
CGAGGGGTGA GAGCGACGC CG	121	TATGTCCCAG CGAGGGTGA GAGCGACGC CGAGGAGGC CGAGAG	221	115112002	0.692170265	7	0.059091
TGTCATTGGC TCTTACAATG	122	AGTTTCAAGA TGTCAATTGGC TCTTACAATG TGGTGGAACC GGG	222	115117415	0.684737218	4	0
AAAGAGGATG TACATTCAACC	123	CCGAAATGCC AAAGAGGATG TACATTCAACC CGGGGGACA GCC	223	115118856	0.677873292	2	0
GTTCGATGTC CCTACAGTGG	124	TACCTTGAG GTTCGATGTC CCTACAGTGG AGGAGGCGG CTGG	224	115114144	0.674367102	6	0
GAACATCTCA CTTTAAATGG	125	ATCCAGCCCCA GAACATCTCA CTTTAAATGG AGGAGGAAAC ATT	225	115118938	0.672438822	2	0.109804

TABLE 2-continued

gRNA Sequence	SEQ ID NOS	Guide with flanking sequence (10 bp before and after the guide + PAM)	SEQ ID NOS	cut position in chromosome	On- target efficiency	Off- target Exon	score
CCGACCCCG AAATGCCAAA G	126	GCTGGTAAGG CCGACCCCG AAATGCCAAA GAGGAGGAT GTACA	226	115118840	0.671949373	2	0
CAAAGAGGA GCATGGCCC CG	127	AGGCGAGA GCAAAGAGGA GCATGGCCC CGAGGAGGC CTGCGA	227	115112032	0.669845158	7	0.043831
ATCCAGGCC GAGACACCG G	128	TGGCCGTGG AATCCAGGCC CGAGACACC GGTGGTGGA GGCCCC	228	115112088	0.658171921	7	0
TTCACTGGAG GACTCATCAG	129	AAGCTGCTTG TTCACTGGAG GACTCATCAG AGGAGGTCC CATT	229	115114210	0.653146782	6	0
GAAAGCTGCT TGTTCACTGG	130	CGAACGCAGTT GAAAGCTGCT TGTTCACTGG AGGAGGACTC ATC	230	115114198	0.6492486	6	0
GAGGTTCGAT GTCCCTACAG	131	GAGGTTCGAT GTCCCTACAG TGGTGGAGG CGGC	231	115114141	0.64192681	6	0
AGCTCACCT GCAGTCCATG	132	CACAGAAAAAC AGCTCACCT GCAGTCCATG AGGAGGTG TTTG	232	115114142	0.624268231	6	0.00293
GCGGAGAGC AAAGAGGAG CA	133	CGACGCCGA GGCCGAGAG CAAAGAGGAG CATGGTGGCC CCGAG	233	115112024	0.619070831	7	0
CATCCACTGT TCCCCAGTAG	134	CGACTTTGGA CATCCACTGT TCCCCAGTAG CGGCGGGGC TGTC	234	115118766	0.618411118	2	0
CCATGCACAA ATACCAGCCC	135	ATATTGAACT CCATGCACAA ATACCAGCCC CGGCGGTTC CACA	235	115117342	0.617360315	4	0.046524
ATAAATTTCA CAATTCTCGG	136	GACTGTCGTT ATAAATTTCA CAATTCTCGG TGGTGGATGG TGG	236	115118801	0.598188657	2	0

TABLE 2-continued

gRNA Sequence	SEQ ID NOS	Guide with flanking sequence (10 bp before and after the guide + PAM)	SEQ ID NOS	cut position in chromosome	On-target efficiency	Exon	Off-target score
GCGTCCGTCT GCACCGTGA G	137	CGCGGCCGGC CGCGTCGGTC TGCACCGTGA CGGGCGGCC CGAAG	237	115112296	0.5963183	7	0
TGTCTCGGGC CTGGATTCCA	138	CCTCCACCCG TGTCTCGGGC CTGGATTCCA CGGCGGCCA TGGC	238	115112545	0.594369303	7	0
GGTGCCCGTT GAAGAACTGT	139	AGGAAGAGC GGGTCCCCG TTGAAGAACT GTTGGTGGCC CGCCA	239	115112210	0.593081981	7	0
TCACAATTCT CGGTGGATG G	140	GTTATAAATT CACAAATTCTC GGTGGATGGT GGTGGCTGG TAA	240	115118808	0.591282693	2	0
CGATGTCCCT ACAGTGGAG G	141	CTTTGAGGTT CGATGTCCCT ACAGTGGAG GCGGGGGCT GGAGA	241	115114147	0.589956292	6	0
CCCCAGAAC CGTGGCCAG G	142	CCGGTGGAG GCCCCAGAAA CCGTGGCCA GGAGGGAGGG GACCCA	242	115112114	0.587176141	7	0
TTTGTGGAAA GTGACGACTT	143	TGAGTTTCAG TTTGTTGAA GTGACGACTT TGGTGGACAT CCA	243	115118742	0.576703566	2	0
CAGGCCGA GACACCGGT GG	144	CCGTGGAATC CAGGCCGA GACACCGGT GGAGGGAGGC CCCAGA	244	115112091	0.564917969	7	0.037427
GGACAGTCCC TGCAGCGCA G	145	CGGACGCC CGGACAGTCC CTGGCCCGC AGCGGGCGC AGAGGC	245	115112046	0.562625974	7	0.007792
TGGCCCCGA GGCCTGCGA CG	146	AAGAGGAGCA TGGCCCCGA GGCCTGCGA CGCGGGCGC CAAGAT	246	115112044	0.562084199	7	0
TCTTTCATCA AACACCCCTCA	147	CCTTTTTGTG TCTTTCATCA AACACCCCTCA TGGTGGACTG CAG	247	115114252	0.547366511	6	0

TABLE 2-continued

gRNA Sequence	SEQ ID NOS	Guide with flanking sequence (10 bp before and after the guide + PAM)	SEQ ID NOS	cut position in chromosome	On-target efficiency	Exon	Off-target score
GGAGGAGCC CTGCCGTGAC A	148	CCACCACGTC GGAGGAGCC CTGCCGTGAC AAGGAGGGC AGCCC	248	115112089	0.545951103	7	0
GGGGCCGCT CAGCAGCGAA A	149	CCGCTGTCCC GGGGCCGCT CAGCAGCGAA AAGGAGGTGA GCCT	249	115112492	0.542944147	7	0
GCGCAGGAG CCCCGTTCGC G	150	GCGCGGAGG AGCCAGGA GCCCGTTC GCGAGGAGG GCACAGC	250	115112248	0.542164198	7	0
GGTTTCCGGG ACACTGGAAA	151	TTTTGCAAAA GGTTTCCGGG ACACTGGAAA TGGTGGCCG AAGA	251	115115443	0.541096538	5	0.046753
AGGGAGTTTC AAGATGTCAT	152	ATGTACTATA AGGGAGTTTC AAGATGTCAT TGGTGGCTCT TAC	252	115117401	0.540968892	4	0
CCGGGGCTG GTATTGTGCA A	153	CAATGTGGAA CCGGGGCTG GTATTGTGCA ATGGTGGAGT TCAA	253	115117440	0.539444587	4	2.87266
ATTCACAATT CTCGGTGGA	154	GTCGTTATAA ATTTCACAGATT CTCGGTGGAT GGTGGTGGC TGG	254	115118805	0.530626362	2	0
CACTCGCGG CCTGGGCGC GG	155	TCTCGTCCAG CACTCGCGG CCTGGGCGC GGAGGAGGA GCGCAG	255	115112224	0.527300954	7	0
AGGCCCCAG AAACCGTGGC C	156	ACACCGGTG GAGGCCCA GAAACCGTGG CCAGGAGGA GGGGAC	256	115112111	0.526292158	7	0
CCGGGGGGC TGCCCTTGTC A	157	TGAGCCTTGA CCGGGGGGC TGCCCTTGTC ACGGCGGCA GGGCT	257	115112525	0.526023177	7	0.059524
GCGGGCCAA CAGTTCTTCA A	158	CCCGGGCCT GGCGGGCCA ACAGTTCTTC AACGGCGGG CACCCG	258	115112411	0.517318009	7	0

TABLE 2-continued

gRNA Sequence	SEQ ID NOS	Guide with flanking sequence (10 bp before and after the guide + PAM)	SEQ ID NOS	cut position in chromosome	On-target efficiency	Exon	Off-target score
TCTCCAGCCG CCTCCACTGT	159	CCAGGCTTCT TCTCCAGCCG CCTCCACTGT AGGAGGGAC ATCG	259	115114255	0.513501316	6	0.02795
TAAAGTGAGA TGTCTGGGC	160	TTCCTCCATT TAAAGTGAGA TGTCTGGGC TGGTGGATAA AAA	260	115118 721	0.51348377	2	0.023228
AGCTTTCAAC TGCTTCGCC	161	GTGAACAAAGC AGCTTTCAAC TGCTTCGCC AGGAGGCTTC TTC	261	115114227	0.512276609	6	0
TTCGGCCATT TCCAGTGTCC	162	CTTTTTCTCT TCGGCCATT CCAGTGTCCC GGCGGAAAC CTT	262	115115414	0.511671134	5	0
GTTGAAAGCT GCTTGTTCAC	163	GGGCGAAGC AGTTGAAAGC TGCTTGTCA CTGGTGGAG GACTC	263	115114195	0.511455039	6	0.012245
GCCATGCTCC TCTTGCTCT	164	AGGCCTCGG GGCCATGCTC CTCTTGCTC TCGGCGGCC TCGGC	264	115112601	0.51141366	7	0
AATTCTCGGT GGATGGTGG C	165	TAAATTCAC AATTCTCGGT GGATGGTGG CTGGTGGTAA GGCC	265	115118812	0.502328709	2	0.131868
CGCGCCGCT CACGGTGCA GA	166	AGGAGGCCTT CGCGCCGCT CACGGTGCA GACGGCGGA CGCGC	266	115112326	0.500475435	7	0
GGGACTGTCC GGGGCGTCC G	167	CTGGCGCGC AGGGACTGTG CGGGCGTC CGCGCGGC CACCTT	267	115112593	0.498017929	7	0.018803
GCAAAAGGTT TCCGGGACAC	168	CAACCTTTT GCAAAAGGTT TCCGGGACAC TGGTGGAAAT GGC	268	115115437	0.4925993	5	0.668662
ATGTTTGCT GAAATGTTGT	169	TTACAAATCC ATGTTTGCT GAAATGTTGT TGGTGGTGAG TTT	269	115118709	0.492461709	2	0

TABLE 2-continued

gRNA Sequence	SEQ ID NOS	Guide with flanking sequence (10 bp before and after the guide + PAM)	SEQ ID NOS	cut position in chromosome	On-target efficiency	Exon	Off-target score
TTACAATGTG GAACCGGGG C	170	TCATTGGCTC TTACAATGTG GAACCGGGG CTGGTGGTAT TTGT	270	115117427	0.491623031	4	0
CAAGGGCAG CCCCGCGGT CA	171	CCTGCCGTGA CAAGGGCAG CCCCGCGGT CAAGGAGGCT CACCT	271	115112107	0.491507388	7	0
CAAACACCCCT CATGGACTGC	172	TGTCTTTCAT CAAACACCCCT CATGGACTGC AGGAGGGTG AGCT	272	115114260	0.483843367	6	0
GTCCTTACAG TGGAGGCGG C	173	GAGGTTCGAT GTCCTTACAG TGGAGGCGG CTGGTGGAGA AGAA	273	115114151	0.479159771	6	0
CCTCTTTGGC ATTTCGGGGT	174	GAATGTACAT CCTCTTTGGC ATTTCGGGGT CGGCGGCCT TACC	274	115118811	0.476893705	2	0
CAGCACTCGC GCCCTGGGC G	175	CCATCTCGTC CAGCACTCGC GGCCTGGGC GCGGCGGAG GAGCG	275	115112221	0.469561023	7	0
AACAACATTT CAGACAAACA	176	GAAACTCACC AACAAACATT CAGACAAACA TGGTGGATT GTA	276	115118944	0.46931656	2	0.123077
GTGGATGAA AGACACAAAA	177	CCATGAGGGT GTGGATGAA AGACACAAAA AGGAGGAGA ATGG	277	115114167	0.466860073	6	0.141087
GGGCCTGGA TTCCACGGCC A	178	CCGGTGTCTC GGGCCTGGA TTCCACGGCC ATGGTGGCCT CTGC	278	115112551	0.46665298	7	0.01555
CACCCGGACA GCCCGCTAC	179	GATGTACATT CACCCGGACA GCCCGCTAC TGGTGGGA ACAG	279	115118872	0.462439623	2	0.095238
TGAAGAACTG TTGGCCCGCC	180	GGGTGCCCG TTGAAGAACT GTTGGCCCG CCAGGAGGC CCGGGG	280	115112219	0.453772064	7	0

TABLE 2-continued

gRNA Sequence	SEQ ID NOS	Guide with flanking sequence (10 bp before and after the guide + PAM)	SEQ ID NOS	cut position in chromosome	On-target efficiency	Exon	Off-target score
GCGGGGCT ATGGCGTGAG T	181	ACGAGATGGT GGCGGGGCT ATGGCGTGAG TCGGCGGGC GACGC	281	115112439	0.45003226	7	0
CCCGGGGCG AAGCCGAGG CC	182	GCCCCGCCAG GCCCGGGGC GAAGCCGAG GCCAGGAGG CAGGGGG	282	115112242	0.446022564	7	0
GCCCTGGGC CAGGTGCCG GG	183	CAGGCAGGG GCCCTGGG CCAGGTGCG CGGCAGGGCG CCGCGTC	283	115112271	0.440747673	7	0
TGGCTCTTAC AATGTGAAAC	184	AAGATGTCAT TGGCTCTTAC AATGTGAAAC CGGCAGGGC TGGT	284	115117421	0.440350741	4	0.237981
TGTTGGTGAG TTTCAGTTTG	185	TCTGAAATGT TGTTGGTGAG TTTCAGTTTG TGGTGGAAAG TGA	285	115118726	0.437592504	2	0.289593
TCGGTGGATG GTGGCTGGTA	186	TTCACAATTG TCGGTGGATG GTGGCTGGTA AGGAGGCCG ACCC	286	115118817	0.43734768	2	0.083612
AGGCCATGG CCGTGGAAATC C	187	GCAGCGGCA GAGGCCATG GCCGTGGAAAT CCAGGAGGC CCGAGA	287	115112072	0.436752641	7	0
GGGTGAATGT ACATCCTCTT	188	GGGGCTGTC CGGGTGAATG TACATCCTCT TTGGTGGCAT TTCG	288	115118797	0.434996943	2	0
TCCCTGCGCC GCAGCGGCA G	189	CCCCGGACA GTCCCCTGCGC CGCACGGC AGAGGAGGC CATGGC	289	115112052	0.433714562	7	0.041663
GGTTCGCGA GGGCACAGC GC	190	GCAGGAGCC CGGTTCGCGA GGGCACAGC GCCGGCGGC CAAGGT	290	115112260	0.432583783	7	0
GGAGATCTTG GCCGCGTCG C	191	ACGTGGTGGT GGAGATCTTG GCCGGTTCG CAGGAGGCC TCGGG	291	115112571	0.425267921	7	0

TABLE 2-continued

gRNA Sequence	SEQ ID NOS	Guide with flanking sequence (10 bp before and after the guide + PAM)	SEQ ID NOS	cut position in chromosome	On-target efficiency	Exon	Off-target score
TGGCAAATG GCTGGGGTGC	192	GCGCCCCC ATGGCAAATC GGCTGGGGT GCAGGAGGA AGAGCG	292	115112180	0.421342539	7	0.045918
ACGAGATGGT GGCGGGCTTA	193	CGAGTGCTG GACGAGATG GTGGCGGG CTATGTTGGC GTGAGT	293	115112429	0.418222381	7	0
TCCTGCGCTC CTCCGGGCC	194	CGAACCGGG CTCCCTGCGCT CCTCCGGCG CCAGGAGGC CGCGAG	294	115112393	0.415381631	7	0.047431
AACTGTTGGC CCGCCAGGC	195	CCCGTTGAAG AACTGTTGGC CCGCCAGGC CCGGCGGGG CGAAG	295	115112224	0.413331832	7	0
TATGTCCGAA ATGTACTATA	196	GGGGAACAA GTATGTCCGA AATGTACTAT AAGGAGGG GTTTC	296	115117381	0.411925573	4	0
TGTCCAGCCG CCCGCTGTCC	197	GGGCACGCTT TGTCCAGCCG CCCCTGTCC CGGGGGGG CGCT	297	115112471	0.404128254	7	0.049451
AGGC GGCTG GAGAAGAAG CC	198	CCTACAGTGG AGGC GGCTG GAGAAGAAGC CTGGTGGGC GAAGC	298	115114164	0.391417195	6	0.024913
CGCGCCTCTT CCACCTTGGC	199	CGGGAGCGC GCGCGCCTCT TCCACCTTGG CCGGGGCG CTGTG	299	115112347	0.384325854	7	0.051996
CTCCACCGGT GTCTCGGGC	200	TTTCTGGGGC CTCCACCGGT GTCTCGGGC CTGGTGGATT CCAC	300	115112536	0.379428798	7	0
TGGGGCCTC CACCGGTGTC	201	CCACGGTTTC TGGGGCCTC CACCGGTGTC TCGGCGGC CTGGA	301	115112530	0.375527286	7	0
GTTATAAATTT ACAATTCT	202	GATGACTGTC GTTATAAATTT CACAAATTCTC GGCGGTGGA TGG	302	115118798	0.368720329	2	0.015571

TABLE 2-continued

gRNA Sequence	SEQ ID NOS	Guide with flanking sequence (10 bp before and after the guide + PAM)	SEQ ID NOS	cut position in chromosome	On-target efficiency	Exon	Off-target score
GCGCTCGCT CTCACCCCTCG C	203	TCTCGGGCCTC GGCGTCGCT CTCACCCCTCG CTGGTGGGA CATAA	303	115112628	0.360741554	7	0
GGGGCTGCC CTTGTCA CGG C	204	CCTTGACCGC GGGGCTGCC CTTGTCA CGG CAGGAGGGC TCCTC	304	115112529	0.357111718	7	0
TCCCCAGTAG CGGGGCTGT C	205	CATCCACTGT TCCCCAGTAG CGGGGCTGT CCGGCGGGT GAATG	305	115118776	0.356138987	2	0.11655
AACTCCCTTA TAGTACATT T	206	GACATCTTGA AACTCCCTTA TAGTACATT T CGGCGGACA TACT	306	115117396	0.339348533	4	0
CGCCGCAGC GGCAGAGGC CA	207	ACAGTCCCTG CGCCGCAGC GGCAGAGGC CATGGTGGCC GTGGA	307	115112058	0.319953253	7	0
CACAGCGATG AATTCA GTT T	208	GGTATGCAGT CACAGCGATG AATTCA GTT T CGGCGGGGA ACAA	308	115117350	0.30653678	4	0
GCTGCGGCG CAGGGACTGT C	209	GGCCTCTGCC GCTGCGGCG CAGGGACTGT CCGGCGGGG CGTCC	309	115112582	0.304673453	7	0
CTCGTCCAGC ACTCGCGGC C	210	CCGCCACCAT CTCGTCCAGC ACTCGCGGC CTGGTGGGC GCGGA	310	115112215	0.299666415	7	0
CTTGGCCGC GTCGCAGGC CT	211	TGGTGGAGAT CTTGGCCGC GTCGCAGGC CTCGGCGGG GCCATG	311	115112577	0.297388466	7	0
GGCGCTGTG CCCTCGCGAA C	212	CACCTTGGCC GGCGCTGTG CCCTCGCGAA CCGGCGGGC TCCTG	312	115112368	0.28490812	7	0.039409
CCAGAACATC TCACTTTAAA	213	TTTATCCAGC CCAGAACATC TCACTTTAAA GGTGGAGGA AAC	313	115118935	0.276653116	2	0

TABLE 2 -continued

gRNA Sequence	SEQ ID NOS	Guide with flanking sequence (10 bp before and after the guide + PAM)	SEQ ID NOS	cut position in chromosome	On-target efficiency	Exon	Off-target score
GTCCCATCAAT AAAATGTATT	214	CAGCTATAAT GTCCCATCAAT AAAATGTATT GGTGGCTTT TT	314	115118898	0.268535242	2	0
GTACATCCTC TTTGGCATT	215	CCGGGTGAAT GTACATCCTC TTTGGCATT CGGCAGGGT CGGC	315	115118805	0.252451409	2	0
CCATTTAAAG TGAGATGTT	216	AATGTTCCCT CCATTTAAAG TGAGATGTT TGGTGGGCT GGAT	316	115118716	0.216921734	2	0
CCCCTCCTGG CCACGGTTTC	217	TGGCATGGGT CCCCTCCTGG CCACGGTTTC TGGTGGGCG CTCC	317	115112510	- 0.030230013	7	0

## (b) Pharmaceutical Compositions

[0080] The siRNA molecules targeting TBX3 disclosed herein for use according to the methods described herein may be provided per se and/or as part of a pharmaceutical composition, where modulators and/or inhibitors can be mixed with suitable carriers or excipients. As used herein a “pharmaceutical composition” refers to a preparation of one or more of the active ingredients described herein with other chemical components such as physiologically suitable carriers and excipients. The purpose of a pharmaceutical composition is to facilitate administration of a compound to an organism.

[0081] Herein the term “active ingredient” refers to any of the siRNA molecules. The term “active ingredient” as used herein can also include any vector, medium, microorganism, or cell culture wherein the siRNA molecule is synthesized, expressed and/or contained, such as a genetically modified cell, viral vector, plasmid, bacteria, yeast, fungus, and the culture or medium thereof.

## (i) Pharmaceutically Acceptable Carriers and Excipients

[0082] Hereinafter, the phrases “physiologically acceptable carrier” and “pharmaceutically acceptable carrier” are interchangeably used herein to refer to a carrier or a diluent that does not cause significant irritation to an organism and does not abrogate the biological activity and properties of the administered compound. An adjuvant is included under these phrases.

[0083] In certain embodiments, compositions disclosed herein may further comprise one or more pharmaceutically acceptable diluent(s), excipient(s), and/or carrier(s). As used herein, a pharmaceutically acceptable diluent, excipient, or carrier, refers to a material suitable for administration to a subject without causing undesirable biological effects or interacting in a deleterious manner with any of the compo-

nents of the composition in which it is contained. Pharmaceutically acceptable diluents, carriers, and excipients can include, but are not limited to, physiological saline, Ringer’s solution, phosphate solution or buffer, buffered saline, and other carriers known in the art.

[0084] In some embodiments, pharmaceutical compositions herein may also include stabilizers, anti-oxidants, colorants, other medicinal or pharmaceutical agents, carriers, adjuvants, preserving agents, stabilizing agents, wetting agents, emulsifying agents, solution promoters, salts, solubilizers, antifoaming agents, antioxidants, dispersing agents, surfactants, or any combination thereof. Herein, the term “excipient” refers to an inert substance added to a pharmaceutical composition to further facilitate administration of an active ingredient. Examples, without limitation, of excipients include calcium carbonate, calcium phosphate, various sugars and types of starch, cellulose derivatives, gelatin, vegetable oils and polyethylene glycols. Techniques for formulation and administration of drugs may be found in “Remington’s Pharmaceutical Sciences,” Mack Publishing Co., Easton, Pa., latest edition, which is incorporated herein by reference.

[0085] In certain embodiments, pharmaceutical compositions described herein may be formulated in conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries to facilitate processing of genetically modified endothelial progenitor cells into preparations which can be used pharmaceutically. In some embodiments, any of the well-known techniques, carriers, and excipients may be used as suitable and/or as understood in the art.

[0086] In certain embodiments, pharmaceutical compositions described herein may be an aqueous suspension comprising one or more polymers as suspending agents. In some embodiments, polymers that may comprise pharmaceutical compositions described herein include: water-soluble poly-

mers such as cellulosic polymers, e.g., hydroxypropyl methylcellulose; water-insoluble polymers such as cross-linked carboxyl-containing polymers; mucoadhesive polymers, selected from, for example, carboxymethylcellulose, carbomer (acrylic acid polymer), poly(methylmethacrylate), polyacrylamide, polycarbophil, acrylic acid/butyl acrylate copolymer, sodium alginate, and dextran; or a combination thereof. In some embodiments, pharmaceutical compositions disclosed herein may comprise at least about 5%, at least about 10%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, or at least about 50% total amount of polymers as suspending agent(s) by total weight of the composition. In some embodiments, pharmaceutical compositions disclosed herein may comprise about 5% to about 99%, about 10%, about 95%, or about 15% to about 90% total amount of polymers as suspending agent(s) by total weight of the composition.

**[0087]** In certain embodiments, pharmaceutical compositions disclosed herein may comprise a viscous formulation. In some embodiments, viscosity of composition herein may be increased by the addition of one or more gelling or thickening agents. In some embodiments, compositions disclosed herein may comprise one or more gelling or thickening agents in an amount to provide a sufficiently viscous formulation to remain on treated tissue. In some embodiments, pharmaceutical compositions disclosed herein may comprise at least about 5%, at least about 10%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, or at least about 50% total amount of gelling or thickening agent(s) by total weight of the composition. In some embodiments, pharmaceutical compositions disclosed herein may comprise about 5% to about 99%, about 10%, about 95%, or about 15% to about 90% total amount of gelling or thickening agent(s) by total weight of the composition. In some embodiments, suitable thickening agents for use herein can be hydroxypropyl methylcellulose, hydroxyethyl cellulose, polyvinylpyrrolidone, carboxymethyl cellulose, polyvinyl alcohol, sodium chondroitin sulfate, sodium hyaluronate. In other aspects, viscosity enhancing agents can be acacia (gum arabic), agar, aluminum magnesium silicate, sodium alginate, sodium stearate, bladderwrack, bentonite, carbomer, carrageenan, Carbopol, xanthan, cellulose, microcrystalline cellulose (MCC), ceratonia, chitin, carboxymethylated chitosan, chondrus, dextrose, furcellaran, gelatin, Ghatti gum, guar gum, hectorite, lactose, sucrose, maltodextrin, mannitol, sorbitol, honey, maize starch, wheat starch, rice starch, potato starch, gelatin, sternaria gum, xanthum gum, gum tragacanth, ethyl cellulose, ethylhydroxyethyl cellulose, ethylmethyl cellulose, methyl cellulose, hydroxyethyl cellulose, hydroxyethylmethyl cellulose, hydroxypropyl cellulose, poly(hydroxyethyl methacrylate), oxypropylgelatin, pectin, polygalacturonic acid, povidone, propylene carbonate, methyl vinyl ether/maleic anhydride copolymer (PVM/MA), poly(methoxyethyl methacrylate), poly(methoxyethoxyethyl methacrylate), hydroxypropyl cellulose, hydroxypropylmethyl-cellulose (HPMC), sodium carboxymethyl-cellulose (CMC), silicon dioxide, polyvinylpyrrolidone (PVP: povidone), Splenda (dextrose, maltodextrin and sucralose), or any combination thereof.

**[0088]** In certain embodiments, pharmaceutical compositions disclosed herein may comprise additional agents or additives selected from a group including surface-active

agents, detergents, solvents, acidifying agents, alkalizing agents, buffering agents, tonicity modifying agents, ionic additives effective to increase the ionic strength of the solution, antimicrobial agents, antibiotic agents, antifungal agents, antioxidants, preservatives, electrolytes, antifoaming agents, oils, stabilizers, enhancing agents, and the like. In some embodiments, pharmaceutical compositions disclosed herein may comprise at least about 5%, at least about 10%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, or at least about 50% total amount of one or more agents by total weight of the composition. In some embodiments, pharmaceutical compositions disclosed herein may comprise about 5% to about 99%, about 10%, about 95%, or about 15% to about 90% total amount of one or more agents by total weight of the composition. In some embodiments, one or more of these agents may be added to improve the performance, efficacy, safety, shelf-life and/or other property of the muscarinic antagonist composition of the present disclosure. In some embodiments, additives may be biocompatible, without being harsh, abrasive, and/or allergenic.

**[0089]** In certain embodiments, pharmaceutical compositions disclosed herein may comprise one or more acidifying agents. As used herein, "acidifying agents" refers to compounds used to provide an acidic medium. Such compounds include, by way of example and without limitation, acetic acid, amino acid, citric acid, fumaric acid and other alpha hydroxy acids, such as hydrochloric acid, ascorbic acid, and nitric acid and others known to those of ordinary skill in the art. In some embodiments, any pharmaceutically acceptable organic or inorganic acid may be used. In some embodiments, pharmaceutical compositions disclosed herein may comprise at least about 5%, at least about 10%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50% total amount of one or more acidifying agents by total weight of the composition. In some embodiments, pharmaceutical compositions disclosed herein may comprise about 5% to about 99%, about 10%, about 95%, or about 15% to about 90% total amount of one or more acidifying agents by total weight of the composition.

**[0090]** In certain embodiments, pharmaceutical compositions disclosed herein may comprise one or more alkalizing agents. As used herein, "alkalizing agents" are compounds used to provide alkaline medium. Such compounds include, by way of example and without limitation, ammonia solution, ammonium carbonate, diethanolamine, monoethanolamine, potassium hydroxide, sodium borate, sodium carbonate, sodium bicarbonate, sodium hydroxide, triethanolamine, and trolamine and others known to those of ordinary skill in the art. In some embodiments, any pharmaceutically acceptable organic or inorganic base can be used. In some embodiments, pharmaceutical compositions disclosed herein may comprise at least about 5%, at least about 10%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50% total amount of one or more alkalizing agents by total weight of the composition. In some embodiments, pharmaceutical compositions disclosed herein may comprise about 5% to about 99%, about 10%, about 95%, or about 15% to about 90% total amount of one or more alkalizing agents by total weight of the composition.

**[0091]** In certain embodiments, pharmaceutical compositions disclosed herein may comprise one or more antioxi-

dants. As used herein, "antioxidants" are agents that inhibit oxidation and thus can be used to prevent the deterioration of preparations by the oxidative process. Such compounds include, by way of example and without limitation, ascorbic acid, ascorbyl palmitate, butylated hydroxyanisole, butylated hydroxytoluene, hypophosphorous acid, monothioglycerol, propyl gallate, sodium ascorbate, sodium bisulfite, sodium formaldehyde sulfoxylate, sodium metabisulfite and other materials known to one of ordinary skill in the art. In some embodiments, pharmaceutical compositions disclosed herein may comprise at least about 5%, at least about 10%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50% total amount of one or more antioxidants by total weight of the composition. In some embodiments, pharmaceutical compositions disclosed herein may comprise about 5% to about 99%, about 10%, about 95%, or about 15% to about 90% total amount of one or more antioxidants by total weight of the composition.

[0092] In certain embodiments, pharmaceutical compositions disclosed herein may comprise a buffer system. As used herein, a "buffer system" is a composition comprised of one or more buffering agents wherein "buffering agents" are compounds used to resist change in pH upon dilution or addition of acid or alkali. Buffering agents include, by way of example and without limitation, potassium metaphosphate, potassium phosphate, monobasic sodium acetate and sodium citrate anhydrous and dihydrate and other materials known to one of ordinary skill in the art. In some embodiments, any pharmaceutically acceptable organic or inorganic buffer can be used. In some embodiments, pharmaceutical compositions disclosed herein may comprise at least about 5%, at least about 10%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50% total amount of one or more buffering agents by total weight of the composition. In some embodiments, pharmaceutical compositions disclosed herein may comprise about 5% to about 99%, about 10%, about 95%, or about 15% to about 90% total amount of one or more buffering agents by total weight of the composition.

[0093] In some embodiments, the amount of one or more buffering agents may depend on the desired pH level of a composition. In some embodiments, pharmaceutical compositions disclosed herein may have a pH of about 6 to about 9. In some embodiments, pharmaceutical compositions disclosed herein may have a pH greater than about 8, greater than about 7.5, greater than about 7, greater than about 6.5, or greater than about 6.

[0094] In certain embodiments, pharmaceutical compositions disclosed herein may comprise one or more preservatives. As used herein, "preservatives" refers to agents or combination of agents that inhibits, reduces or eliminates bacterial growth in a pharmaceutical dosage form. Non-limiting examples of preservatives include Nipagin, Nipasol, isopropyl alcohol and a combination thereof. In some embodiments, any pharmaceutically acceptable preservative can be used. In some embodiments, pharmaceutical compositions disclosed herein may comprise at least about 5%, at least about 10%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50% total amount of one or more preservatives by total weight of the composition. In some embodiments, pharmaceutical com-

positions disclosed herein may comprise about 5% to about 99%, about 10%, about 95%, or about 15% to about 90% total amount of one or more preservatives by total weight of the composition.

[0095] In certain embodiments, pharmaceutical compositions disclosed herein may comprise one or more surface-acting reagents or detergents. In some embodiments, surface-acting reagents or detergents may be synthetic, natural, or semi-synthetic. In some embodiments, compositions disclosed herein may comprise anionic detergents, cationic detergents, zwitterionic detergents, ampholytic detergents, amphoteric detergents, nonionic detergents having a steroid skeleton, or a combination thereof. In some embodiments, pharmaceutical compositions disclosed herein may comprise at least about 5%, at least about 10%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50% total amount of one or more surface-acting reagents or detergents by total weight of the composition. In some embodiments, pharmaceutical compositions disclosed herein may comprise about 5% to about 99%, about 10%, about 95%, or about 15% to about 90% total amount of one or more surface-acting reagents or detergents by total weight of the composition.

[0096] In certain embodiments, pharmaceutical compositions disclosed herein may comprise one or more stabilizers. As used herein, a "stabilizer" refers to a compound used to stabilize an active agent against physical, chemical, or biochemical process that would otherwise reduce the therapeutic activity of the agent. Suitable stabilizers include, by way of example and without limitation, succinic anhydride, albumin, sialic acid, creatinine, glycine and other amino acids, niacinamide, sodium acetyltryptophanate, zinc oxide, sucrose, glucose, lactose, sorbitol, mannitol, glycerol, polyethylene glycols, sodium caprylate and sodium saccharin and others known to those of ordinary skill in the art. In some embodiments, pharmaceutical compositions disclosed herein may comprise at least about 5%, at least about 10%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50% total amount of one or more stabilizers by total weight of the composition. In some embodiments, pharmaceutical compositions disclosed herein may comprise about 5% to about 99%, about 10%, about 95%, or about 15% to about 90% total amount of one or more stabilizers by total weight of the composition.

[0097] In some embodiments, pharmaceutical compositions disclosed herein may comprise one or more tonicity agents. As used herein, a "tonicity agents" refers to a compound that can be used to adjust the tonicity of the liquid formulation. Suitable tonicity agents include, but are not limited to, glycerin, lactose, mannitol, dextrose, sodium chloride, sodium sulfate, sorbitol, trehalose and others known to those of ordinary skill in the art. Osmolarity in a composition may be expressed in milliosmoles per liter (mOsm/L). Osmolarity may be measured using methods commonly known in the art. In some embodiments, a vapor pressure depression method is used to calculate the osmolarity of the compositions disclosed herein. In some embodiments, the amount of one or more tonicity agents comprising a pharmaceutical composition disclosed herein may result in a composition osmolarity of about 150 mOsm/L to about 500 mOsm/L, about 250 mOsm/L to about 500 mOsm/L, about 250 mOsm/L to about 350 mOsm/L, about 280

mOsm/L to about 370 mOsm/L or about 250 mOsm/L to about 320 mOsm/L. In some embodiments, a composition herein may have an osmolality ranging from about 100 mOsm/kg to about 1000 mOsm/kg, from about 200 mOsm/kg to about 800 mOsm/kg, from about 250 mOsm/kg to about 500 mOsm/kg, or from about 250 mOsm/kg to about 320 mOsm/kg, or from about 250 mOsm/kg to about 350 mOsm/kg or from about 280 mOsm/kg to about 320 mOsm/kg. In some embodiments, a pharmaceutical composition described herein may have an osmolarity of about 100 mOsm/L to about 1000 mOsm/L, about 200 mOsm/L to about 800 mOsm/L, about 250 mOsm/L to about 500 mOsm/L, about 250 mOsm/L to about 350 mOsm/L, about 250 mOsm/L to about 320 mOsm/L, or about 280 mOsm/L to about 320 mOsm/L. In some embodiments, pharmaceutical compositions disclosed herein may comprise at least about 5%, at least about 10%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50% total amount of one or more tonicity modifiers by total weight of the composition. In some embodiments, pharmaceutical compositions disclosed herein may comprise about 5% to about 99%, about 10%, about 95%, or about 15% to about 90% total amount of one or more tonicity modifiers by total weight of the composition.

#### (ii) Dosage Formulations

**[0098]** In certain embodiments, the present disclosure provides compositions formulated for one or more routes of administration. Suitable routes of administration may, for example, include oral, rectal, transmucosal, transnasal, intestinal, and/or parenteral delivery. In some embodiments, compositions herein formulated can be formulated for parenteral delivery. In some embodiments, compositions herein formulated can be formulated intramuscular, subcutaneous, intramedullary, intravenous, intraperitoneal, and/or intranasal injections.

**[0099]** In certain embodiments, one may administer a composition herein in a local or systemic manner, for example, via local injection of the pharmaceutical composition directly into a tissue region of a patient. In some embodiments, a pharmaceutical composition disclosed herein can be administered parenterally, e.g., by intravenous injection, intracerebroventricular injection, intra-cisterna magna injection, intra-parenchymal injection, or a combination thereof. In some embodiments, a pharmaceutical composition disclosed herein can be administered to subject as disclosed herein. In some embodiments, a pharmaceutical composition disclosed herein can be administered to human patient. In some embodiments, a pharmaceutical composition disclosed herein can be administered to a human patient via at least two administration routes. In some embodiments, the combination of administration routes by be intracerebroventricular injection and intravenous injection; intrathecal injection and intravenous injection; intra-cisterna magna injection and intravenous injection; and/or intra-parenchymal injection and intravenous injection.

**[0100]** In certain embodiments, pharmaceutical compositions of the present disclosure may be manufactured by processes well known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

**[0101]** In certain embodiments, pharmaceutical compositions for use in accordance with the present disclosure thus may be formulated in conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries, which facilitate processing of the active ingredients into preparations which, can be used pharmaceutically. Proper formulation is dependent upon the route of administration chosen. For injection, the active ingredients of a pharmaceutical composition herein may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hank's solution, Ringer's solution, physiological salt buffer, or any combination thereof.

**[0102]** In certain embodiments, pharmaceutical compositions described herein may be formulated in the form of a nanoparticle. The nanoparticle may have a monolayer enclosing the nanoparticle core, wherein the siRNA molecule is disposed within the nanoparticle core. In an embodiment, the nanoparticle core includes a solid lipid (i.e., lipid that remains solid at room temperature and body temperature) or a liquid lipid (i.e., oil, which remains liquid at room temperature and body temperature, for example, vegetable oil or a lipid extracted from human adipose tissue). In particular, embodiments of the present disclosure include nanoparticles and compositions for the controlled and/or sustained release (e.g., release at a predetermined rate to maintain a certain concentration for a certain period of time) of an agent, such as a small interfering RNA (siRNA) from the nanoparticle.

**[0103]** In certain embodiments, pharmaceutical compositions described herein may be formulated for parenteral administration, e.g., by bolus injection or continuous infusion. Formulations for injection herein may be presented in unit dosage form, e.g., in ampoules or in multidose containers with optionally, an added preservative. In some embodiments, compositions herein may be suspensions, solutions or emulsions in oily or aqueous vehicles, and/or may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

**[0104]** In certain embodiments, pharmaceutical compositions herein formulated for parenteral administration may include aqueous solutions of the active preparation (e.g., a siRNA molecule) in water-soluble form. In some embodiments, compositions herein comprising suspensions of the active preparation may be prepared as oily or water-based injection suspensions. Suitable lipophilic solvents and/or vehicles for use herein may include, but are not limited to, fatty oils such as sesame oil, or synthetic fatty acids esters such as ethyl oleate, triglycerides or liposomes. In some embodiments, compositions herein comprising aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, and/or dextran. In some embodiments, compositions herein comprising a suspension may also contain one or more suitable stabilizers and/or agents which increase the solubility of the active ingredients (e.g., a siRNA molecule) to allow for the preparation of highly concentrated solutions.

**[0105]** In some embodiments, compositions herein may comprise the active ingredient in a powder form for constitution with a suitable vehicle, e.g., sterile, pyrogen-free water-based solution, before use.

**[0106]** Pharmaceutical compositions suitable for use in context of the present disclosure may include compositions wherein the active ingredients can be contained in an

amount effective to achieve the intended purpose. In some embodiments, a therapeutically effective amount means an amount of active ingredients (e.g., a siRNA molecule) effective to prevent, slow, alleviate or ameliorate symptoms of a disorder (e.g., lymphoproliferative disorders, lymphoid malignancy) or prolong the survival of the subject being treated. Determination of a therapeutically effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any preparation used in the methods of the present disclosure, the therapeutically effective amount or dose can be estimated initially from in vitro and cell culture assays and or screening platforms disclosed herein. For example, a dose can be formulated in animal models to achieve a desired concentration or titer. Such information can be used to more accurately determine useful doses in humans.

[0107] In some embodiments, toxicity and therapeutic efficacy of the active ingredients disclosed herein (e.g., a siRNA molecule) can be determined by standard pharmaceutical procedures in vitro, in cell cultures or experimental animals. In some embodiments, data obtained from these in vitro and cell culture assays and animal studies can be used in formulating a range of dosage for use in a human subject. In some embodiments, a dosage for use herein may vary depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. (See e.g., Fingl, et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1).

[0108] In certain embodiments, dosage amounts and/or dosing intervals may be adjusted individually to brain or blood levels of the active ingredient that are sufficient to induce or suppress the biological effect (minimal effective concentration, MEC). In some embodiments, the MEC for an active ingredient (e.g., a siRNA molecule or composition disclosed herein) may vary for each preparation but can be estimated from in vitro data. In some embodiments, dosages necessary to achieve the MEC herein may depend on individual characteristics and route of administration. Detection assays can be used to determine plasma concentrations.

[0109] In certain embodiments, depending on the severity and responsiveness of the condition to be treated, dosing with compositions herein can be of a single or a plurality of administrations, with course of treatment lasting from several days to several weeks or until cure is effected or diminution of the disease state is achieved.

[0110] In certain embodiments, amounts of a composition herein to be administered will be dependent on the subject being treated, the severity of the affliction, the manner of administration, the judgment of the prescribing physician, and the like. In some embodiments, effective doses may be extrapolated from dose-responsive curves derived from in vitro or in vivo test systems.

### III. Methods

[0111] The present disclosure provides for methods of treating, attenuating, and preventing liver disease in a subject in need thereof. In several embodiments, a method for treating, attenuating, or preventing liver disease in a subject can include administering to a subject, including a human subject, an effective amount of one or more siRNA molecules targeting TBX3 disclosed herein or a nucleic acid encoding an siRNA molecule targeting TBX3 as disclosed

herein. In several embodiments, a method for treating, attenuating, or preventing liver disease in a subject can include administering to a subject, including a human subject, an effective amount of a nucleic acid encoding a suitable sgRNA, or a suitable sgRNA targeting TBX3 and an RNA guided endonuclease.

[0112] Methods disclosed herein may include treating a subject in need thereof by administrating a therapeutically effective amount of one or more siRNA molecules or a pharmaceutical composition disclosed herein. The subject may be a human subject having or suspected of having, or at risk of having liver disease, liver damage, liver dysfunction, liver injury. The term "liver disease", "liver injury" or "liver dysfunction" may be used interchangeably and refer to any injury of the liver, including but not limited to hardening of the liver, scarring of the liver, decreased or abnormal biliary tract function, abnormal liver enzyme activity, cirrhosis of the liver, abnormal physiology as determined by common diagnostic methods include but not limited to ultrasound, or biopsy/histopathology, necrosis of the liver and the like. Non-limiting examples of liver disease to be treated using the methods disclosed herein may include fatty liver disease (FLD), alcohol-related liver disease (ARLD), non-alcoholic fatty liver disease (NAFLD), non-alcoholic steatohepatitis (NASH), end stage liver disease (cirrhosis) from any etiology, liver cancer HCC, or any combination thereof. In some embodiments, the subject benefits of an increase in life expectancy compared to an untreated subject with identical disease condition and predicted outcome. In some other embodiments, the treatment improves the subject's liver function as compared to an untreated subject with identical disease condition and predicted outcome. In yet some other embodiments, the treatment attenuates the subject's liver fibrosis as compared to an untreated subject with identical disease condition and predicted outcome. In some embodiments, the treatment prevents additional liver fibrosis in the subject compared to an untreated subject with identical disease condition and predicted outcome.

[0113] A subject suitable for the liver disease treatment as disclosed herein may be selected based on the subject's diagnosis. In some embodiments, a method of diagnosis may detect one or more serum markers indicative of liver disease. Non-limiting examples of serum markers indicative of a liver disease (e.g., NAFLD, NASH, or HCC) may include alpha-fetoprotein (AFP) (e.g., an AFP level of 20 ng/mL or higher), des-gamma-carboxy prothrombin, lens culinaris agglutinin-reactive AFP (AFP-L3), and the like. The diagnosis method may also include the evaluation of at least one clinical symptom associated with a liver disease. Non-limiting examples of clinical symptoms associated with a liver disease may include mild to moderate upper abdominal pain, weight loss, early satiety, or a palpable mass in the upper abdomen, paraneoplastic syndrome, hypoglycemia, erythrocytosis, hypercalcemia, intractable diarrhea and associated electrolyte disturbances (e.g., hyponatremia, hypokalemia, metabolic alkalosis), cutaneous manifestations (e.g., dermatomyositis, pemphigus foliaceus, seborrheic keratosis, pityriasis rotunda), intraperitoneal bleeding, jaundice, fever, pyogenic liver abscess, and the like. Other aspects of diagnosis may include at diagnosis and/or a determination of severity of cirrhosis, chronic infection of hepatitis B virus (HBV), chronic infection of hepatitis C virus (HCV), non-alcoholic fatty liver disease (NAFLD), non-alcoholic ste-

atohepatitis (NASH), primary biliary cirrhosis (PBC), hereditary hemochromatosis, type 2 diabetes, obesity, tobacco use, alcohol abuse, long-term anabolic steroid use, tyrosinemia, alpha1-antitrypsin deficiency, porphyria cutanea tarda, glycogen storage diseases, Wilson disease, or any combination thereof.

[0114] In some embodiments, a subject can be diagnosed and/or predicted to have high or low risk for a liver disease (e.g., NASH or NAFLD) by histological or imaging-based examinations, such as contrast-enhanced multiphase CT, ultrasound, and/or MRI. Imaging features used to diagnose may include liver size, kinetics, and pattern of contrast enhancement, and growth on serial imaging wherein size may be measured as the maximum cross-section diameter on the image where the lesion is most clearly seen. The histologic appearance of NASH or NAFLD biopsies can include steatosis, inflammation, and fibrosis.

[0115] The methods and compositions of the present disclosure are useful for the treatment of subjects having fatty liver related disorders, such as NAFLD and/or NASH. The subject may have normal or substantially normal biliary tract function. Normal or substantially normal biliary tract function may be determined in a subject using any suitable methods known in the art. Generally, preferred tests for biliary tract function in NASH patients may be characterized in two groups: physiological based tests and biochemical based tests. Physiological based tests may include but are not limited to abdominal ultrasound, abdominal CT scan, endoscopic retrograde cholangiopancreatography (ERCP), Percutaneous transhepatic cholangiogram (PTCA) or Magnetic resonance cholangiopancreatography (MRCP). Biochemical based tests may include but are not limited to GGT tests, liver function tests, bilirubin tests, alkaline phosphatase (ALP) tests, liver enzyme tests, amylase blood test, lipase blood test, prothrombin time, and measurement of urine bilirubin. In some cases, one or more tests may be used to characterize biliary function. In some cases, a combination of tests may be used to assess biliary function in NASH subjects.

[0116] In some embodiments, any of the methods disclosed herein can further include monitoring for an occurrence of one or more adverse effects in the subject. Adverse effects may include, but are not limited to, hepatic impairment, hematologic toxicity, neurologic toxicity, cutaneous toxicity, gastrointestinal toxicity, or a combination thereof. When one or more adverse effects are observed, the methods disclosed herein can further include reducing or increasing the dose of one or more of the treatment regimens depending on the adverse effect or effects in the subject. For example, when a moderate to severe hepatic impairment is observed in a subject after treatment, compositions of use to treat the subject can be reduced in concentration or frequency.

[0117] In certain embodiments, treatments administered according to the methods disclosed herein can improve patient life expectancy compared to the life expectancy of an untreated subject with identical disease condition (e.g., NAFLD or NASH) and predicted outcome. As used herein, "patient life expectancy" is defined as the time at which 50 percent of subjects are alive and 50 percent have passed away. In some embodiments, patient life expectancy can be indefinite following treatment according to the methods disclosed herein. In other aspects, patient life expectancy can be increased at least about 5% or greater to at least about 100%, at least about 10% or greater to at least about 95% or

greater, at least about 20% or greater to at least about 80% or greater, at least about 40% or greater to at least about 60% or greater compared to an untreated subject with identical disease condition and predicted outcome. In some embodiments, patient life expectancy can be increased at least about 5% or greater, at least about 10% or greater, at least about 15% or greater, at least about 20% or greater, at least about 25% or greater, at least about 30% or greater, at least about 35% or greater, at least about 40% or greater, at least about 45% or greater, at least about 50% or greater, at least about 55% or greater, at least about 60% or greater, at least about 65% or greater, at least about 70% or greater, at least about 75% or greater, at least about 80% or greater, at least about 85% or greater, at least about 90% or greater, at least about 95% or greater, at least about 100% compared to an untreated subject with identical disease condition and predicted outcome. In some embodiments, patient life expectancy can be increased at least about 5% or greater to at least about 10% or greater, at least about 10% or greater to at least about 15% or greater, at least about 15% or greater to at least about 20% or greater, at least about 20% or greater to at least about 25% or greater, at least about 25% or greater to at least about 30% or greater, at least about 30% or greater to at least about 35% or greater, at least about 35% or greater to at least about 40% or greater, at least about 40% or greater to at least about 45% or greater, at least about 45% or greater to at least about 50% or greater, at least about 50% or greater to at least about 55% or greater, at least about 55% or greater to at least about 60% or greater, at least about 60% or greater to at least about 65% or greater, at least about 65% or greater to at least about 70% or greater, at least about 70% or greater to at least about 75% or greater, at least about 75% or greater to at least about 80% or greater, at least about 80% or greater to at least about 85% or greater, at least about 85% or greater to at least about 90% or greater, at least about 90% or greater to at least about 95% or greater, at least about 95% or greater to at least about 100% compared to an untreated patient with identical disease condition and predicted outcome.

[0118] In some embodiments, treatment of a liver disease, such as NAFLD or NASH, according to the methods disclosed herein can result in an attenuating, a shrinking, a reducing or a preventing of a liver fibrosis in comparison to the starting size of the liver fibrosis. In some embodiments, liver fibrosis attenuating or shrinking may be at least about 5% or greater to at least about 10% or greater, at least about 10% or greater to at least about 15% or greater, at least about 15% or greater to at least about 20% or greater, at least about 20% or greater to at least about 25% or greater, at least about 25% or greater to at least about 30% or greater, at least about 30% or greater to at least about 35% or greater, at least about 35% or greater to at least about 40% or greater, at least about 40% or greater to at least about 45% or greater, at least about 45% or greater to at least about 50% or greater, at least about 50% or greater to at least about 55% or greater, at least about 55% or greater to at least about 60% or greater, at least about 60% or greater to at least about 65% or greater, at least about 65% or greater to at least about 70% or greater, at least about 70% or greater to at least about 75% or greater, at least about 75% or greater to at least about 80% or greater, at least about 80% or greater to at least about 85% or greater, at least about 85% or greater to at least about 90% or greater, at least about 90% or greater to at least about 95% or greater, at least about 95% or greater to at least about 100% (meaning that the liver

fibrosis is completely gone after treatment) compared to the starting size of the liver fibrosis.

[0119] In some embodiments, treatment of a liver disease, such as NAFLD or NASH, according to the methods disclosed herein can result in an improved liver function. Liver function may be tested through routine biochemical methods. Biochemical based tests may include but are not limited to GGT tests, liver function tests, bilirubin tests, alkaline phosphatase (ALP) tests, liver enzyme tests, amylase blood test, lipase blood test, prothrombin time, and measurement of urine bilirubin. In some cases, one or more tests may be used to characterize biliary function. In some cases, a combination of tests may be used to assess biliary functions. The liver function improvement can also be assessed by subject's symptom relief, such as relieving in fatigue, weight loss, and weakness of the subject. Other relieved symptoms may include reduced fluid retention, muscle wasting, bleeding from the intestines, and any combination thereof.

#### IV. Kits

[0120] The present disclosure provides kits for use in the methods disclosed herein. In some aspects, the present disclosure provides a kit for treating a liver disease (e.g., NASH or NAFLD) as disclosed herein and/or for diagnosing the liver disease. Such a kit may comprise a means for holding and/or administering such a siRNA composition or a pharmaceutical composition.

[0121] In some embodiments, kits disclosed herein can have a medical container, which holds the composition in a safe, stable and durable way. In some examples, kits disclosed herein may also comprise a means to administer the composition, such as a needle or a spatula.

[0122] Any of the kits may further comprise an instruction manual providing guidance for using the kit for treatment. The manual may be written with the physician or the liver specialist as the intended reader.

[0123] Having described several embodiments, it will be recognized by those skilled in the art that various modifications, alternative constructions, and equivalents may be used without departing from the spirit of the present inventive concept. Additionally, a number of well-known processes and elements have not been described in order to avoid unnecessarily obscuring the present inventive concept. Accordingly, this description should not be taken as limiting the scope of the present inventive concept.

[0124] Those skilled in the art will appreciate that the presently disclosed embodiments teach by way of example and not by limitation. Therefore, the matter contained in this description or shown in the accompanying drawings should be interpreted as illustrative and not in a limiting sense. The following claims are intended to cover all generic and specific features described herein, as well as all statements of the scope of the method and assemblies, which, as a matter of language, might be said to fall there between.

#### EXAMPLES

[0125] The following examples are included to demonstrate preferred embodiments of the disclosure. It should be appreciated by those of skill in the art that the techniques disclosed in the examples that follow represent techniques discovered by the inventor to function well in the practice of the present disclosure, and thus can be considered to constitute preferred modes for its practice. However, those of

skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the present disclosure.

#### Example 1

[0126] Somatic mutation screening of transcription and epigenetic factors identified therapeutic targets in NASH. It was found that somatic mutations in hepatocytes could be promoting cellular fitness through the suppression of lipotoxicity. Using transcriptomic data from HCV cirrhosis and NASH patients, transcriptional or epigenetic regulators with altered gene expression in these disease states were identified. AAV-sgRNA libraries containing pools of sgRNAs targeting either transcription or epigenetic factors were generated, and these libraries were used to perform two independent somatic mosaic screens in the mouse liver. Each library contained 200-300 genes. After the establishment of somatic mutations in these genes using in vivo expression of Cas9, mice are put on normal chow or western diet/sugar water for up to 6 months. The sgRNAs associated with the most clonal expansion in both of these screens targeted a set of 23 genes (13 from the TF library and 10 from the epifactor library). Among the top hits from this in vivo genetic screen analysis was the T-box transcription factor (TBX3, see FIG. 1). TBX3 has not been studied in fatty liver disease before.

[0127] To further ascertain which genes from the genetic screens were increasing clonal expansion through an influence on metabolic fitness, conditional knockout mouse models were needed. The CRISPR plus AAV-sgRNA approach described herein was first used to generate liver-wide conditional deletion models. AAV-sgRNAs against GFP or LacZ were implemented as independent control models. High titer AAVs-sgRNAs against Tbx3 were injected into Cas9 expressing mice at 8 weeks of age, then after 2 weeks, mice are given western diet (WD) for three months. For each conditional CRISPR KO model, body weight, liver weight, histology, steatosis, fibrosis, serum tests (liver function tests, cholesterol, TGs, non-esterified free fatty acids (NEFAs)) are measured. Tbx3 KO mice showed reduced liver weight and liver to body weight ratios compared to controls (FIG. 1). In addition, TBX3 KO mice showed significantly reduced liver triglycerides. Moreover, there was reduced liver damage and inflammation as measured by AST/ALT and reduced hepatic steatosis on histology (See FIGS. 2A-2F). These genetic data showed that TBX3 was a promising therapeutic target in human NASH.

#### Example 2

[0128] TBX3 floxed mice were used to more rigorously examine loss of TBX3 in the liver. This knockout experiment allows nearly 100% hepatocyte deletion, which cannot be achieved using CRISPR AAV KO approaches, which are less efficient. Spatial and temporal conditional deletion are achieved with AAV-TBG-Cre, which expresses Cre recombinase only in hepatocytes. Mice are given AAV at 6 weeks and initiated on NASH diets at 8 weeks of age. Tbx3<sup>f/f</sup> and Tbx3<sup>+/+</sup> mice are given AAV-TBG-Cre (n=15 for each genotype, males and females tested separately). Four groups are compared: Tbx3<sup>+/+</sup>+AAV-TBG-Cre on normal chow (NC), Tbx3<sup>f/f</sup>+AAV-Cre on NC, Tbx3<sup>+/+</sup>+AAV-TBG-Cre

on WD, *Tbx3<sup>fl/fl</sup>*+AAV-TBG-Cre on WD. The NC groups are used to rule out phenotypic effects of *Tbx3* loss that may arise independently of diet. These rigorous reference standard KO mice are used to determine the extent to which NASH can be prevented. Effects of *Tbx3* deletion on liver steatosis, inflammation, and fibrosis after 12 and 24 weeks of NASH diets are characterized. In WT control mice fed WD, hepatic steatosis and inflammation are prominent at 12 weeks, while hepatocyte ballooning and fibrosis (features that are more NASH specific), are observed by 24 weeks. [0129] NASH progression is characterized. Mice are examined twice weekly when food and sugar water are replenished. The mice are visually inspected for signs of distress/illness. Body weights are recorded once weekly. Survival times are recorded in case of death (unexpected). Liver toxicity is monitored prior to liver harvesting (every 12 weeks) by blood AST, ALT, total bilirubin, albumin, and complete blood counts using the UTSW Molecular Phenotyping Core. Examined biomarkers include plasma and liver triglycerides (TG), non-esterified free fatty acids (NEFA), and cholesterol, all of which are the critical lipid markers of NASH and metabolic syndrome. Livers are photographed and/or sectioned for macro/microscopic evidence of tissue injury, inflammation, and fibrosis. Histology and whole slide imaging are performed using an Axioscan microscope. Oil Red-O and Sirius Red stains are used to examine steatosis and fibrosis, respectively. Livers are examined in a blinded fashion for components of the NASH Activity Score (NAS): steatosis, lobular inflammation, and hepatocyte ballooning. Separately, the degree of fibrosis is quantified. If cancers appear, it is quantified. TBX3 suppression resulted in reduced steatosis (FIGS. 1 and 2A-2F) and inflammation. Since it was possible for steatosis to be reduced without concomitant changes in ballooning or fibrosis, long-term WD feeding is critical to ascertain the distinct pathological components of NASH.

[0130] TBX3 hepatocyte specific KO mice have a strong protective effect against steatosis and inflammation than the prior CRISPR KO mice because they have a more complete deletion of TBX3 in more hepatocytes. Over 12 and 24 weeks, it is reasonable to expect a delay, but not complete elimination of steatosis, inflammation, ballooning, and fibrosis.

[0131] To summarize, one of the most promising genetic hits was *Tbx3*, a gene not known to be involved in NASH prior to this disclosure. Liver specific conditional knockout mouse models were generated to further study *Tbx3* in NASH. A CRISPR plus AAV-sgRNA approach was used to generate a liver-wide conditional deletion model for *Tbx3* and AAV-sgRNAs against GFP or LacZ were used as independent controls (FIG. 1). In the context of 3 months of NASH-inducing diets, deletion of *Tbx3* showed significantly reduced liver to body weight ratios (FIG. 1), a highly specific biomarker of liver fat accumulation and NASH. Moreover, there was reduced liver damage and inflammation as measured by AST/ALT, and significantly reduced hepatic steatosis on histology (FIGS. 2A-2F).

### Example 3

[0132] In order to address if a mouse siRNA against TBX3 prevents and/or reverses NASH in a mouse model, an siRNA tool compound used in mice is used to study translational aspects of TBX3 biology. Potent siRNA sequences were identified. Next, dosing is optimized to maximize

effectiveness in vivo. Eight siRNAs have been tested against mouse TBX3 in Cos7 cells to identify the siRNA sequence with the highest knockdown efficiency based on mouse *Tbx3* reporter assays. Then this siRNA is modified and conjugated with GalNAc in the standard fashion (Foster et al., 2018). Briefly, in order to minimize nucleolytic degradation and immune responses in vivo, every 2'-position is modified with 2'-O-methyl. To determine if these stabilizing modifications influence silencing activities, siRNA activity in vitro prior to GalNAc conjugation is determined. In accord with the industry standard approach, GalNAc conjugation allows for high efficiency delivery to hepatocytes without lipid nanoparticle packaging. The best in vivo dosing regimen is then determined for GalNAc-siTBX3. Three doses of siRNAs in PBS (1, 5, 10, 15 mg/kg) are given subcutaneously (SC), once every two weeks starting at 4 weeks of age to WT B6 mice. After siRNA treatments, mice are sacrificed for TBX3 qPCR, western blotting, and histology analysis (n=5 mice/group). Maximum tolerated dose is determined based on lack of severe toxicities such as weight loss greater than 10%, organ (heart/kidney/liver) failure, and death. In all of the following siRNA experiments, GalNAc-siLuc serves as the control.

[0133] To determine if GalNAc-siTbx3 prevents NASH, the optimized dose defined above is used. Then it is determined if GalNAc-siRNA can mimic the TBX3 KO model. In a parallel fashion as the genetic experiments, GalNAc-siRNA SC injections are started at 6 weeks of age and NASH diets are started at 8 weeks and continued for 12 weeks (6 total siRNA doses). Mice are euthanized at 20 weeks of age, a time point at which steatosis, inflammation, and fibrosis are assessed. The siRNA approach closely mimics the genetic NASH prevention model.

[0134] To determine if NASH can be reversed by GalNAc-siTBX3, siRNA dosing is initiated in mice that have already received 24 weeks of WD diets (30 weeks of age) and the WD and siRNAs are continued for 12 weeks total. Then the livers are assessed for pathological features of NASH. Assessment includes hepatocyte ballooning, inflammation, and fibrosis, features that characterize NASH as described above. These siRNA experiments provide immense translational value and pave the path for clinical therapeutics.

### Example 4

[0135] Specific siRNA sequences were designed to target human TBX3 for the treatment of NASH. Effective siRNAs that can knockdown human TBX3 were identified (FIGS. 3A-3B). Tests identified optimized siRNA sequences corresponding to the sequences of human TBX3. This involved screening candidate siRNAs per gene target using in vitro luciferase reporter-based assays. In brief, full length TBX3 cDNA was cloned into a psicheck2 plasmid backbone containing a luciferase gene. This created a *Renilla* luciferase-TBX3 fusion gene. The luciferase assay was a dual reporter system with Firefly luciferase as a control and *Renilla* luciferase as the read out for transcription and translation. Thus, the ratio of *Renilla* to Firefly signal accounted for variations in transfection efficiency and cell viability. The assay was carried out in a 96 well plate format with technical replicates using a 48-hour time point for the assay readout. The reporter along with siRNAs were co-transfected into Cos7 monkey kidney cells using lipofectamine. After 48 hours, the cells were lysed and the signal was captured by a

luminometer using the substrates for Firefly and *Renilla*. Effective siRNAs against a target gene showed a reduced ratio for *Renilla* vs. Firefly signal compared to non-targeting controls. A total of 45 siRNAs targeting the entire TBX3 cDNA were designed and tested. siRNAs designed by Dharmacon and ThermoFisher were also tested. At least 15 effective siRNAs were identified against human TBX3 (Table 3). In other words, 15 siRNAs shown in Table 2 had less than 40% of TBX3 mRNA remaining. The siRNAs designed and tested according to the methods disclosed herein shown in FIGS. 3A-3B. Optimized siRNAs are modified by conjugating to acetylgalactosamine (GalNAC) in the standard fashion and tested in vivo.

TABLE 3

siRNA ID	% of TBX3 mRNA remaining (compared to neg. ctrl siRNA, at 3.3 nM or 0.5 pmol)	Experimental Design or Manufacturer Design
TBX3 470	81.3	Experimental
TBX3 1080	22.7	Experimental
TBX3 457	83.0	Experimental
TBX3 707	85.0	Experimental
TBX3 754	89.1	Experimental
TBX3 966	79.9	Experimental
TBX3 631	81.0	Experimental
TBX3 628	77.9	Experimental
TBX3 627	76.5	Experimental
TBX3 857	77.2	Experimental
TBX3 389-411	88.2	Experimental
TBX3 390-412	86.7	Experimental
TBX3 438-460	82.8	Experimental
TBX3 619-641	80.3	Experimental
TBX3 768-790	81.2	Experimental
TBX3 917-939	127.8	Experimental
TBX3 428-450	19.3	Experimental
TBX3 437-459	16.5	Experimental
TBX3 778-800	73.9	Experimental
TBX3 860-882	76.7	Experimental
J-012197-05	59.0	Dharmacon
J-012197-06	35.3	Dharmacon
J-012197-07	40.2	Dharmacon
J-012197-08	70.0	Dharmacon
D-012197-01	62.7	Dharmacon
D-012197-02	20.2	Dharmacon
D-012197-03	23.3	Dharmacon
D-012197-04	29.3	Dharmacon
s13865	26.1	ThermoFisher
s13866	21.5	ThermoFisher
s13867	14.1	ThermoFisher
s529425	82.5	ThermoFisher
s534311	18.9	ThermoFisher
s534312	27.3	ThermoFisher
107921	50.3	ThermoFisher
107922	54.4	ThermoFisher
115750	59.4	ThermoFisher
115751	66.6	ThermoFisher
139810	126.8	ThermoFisher
239194	30.6	ThermoFisher
239195	64.2	ThermoFisher
239196	55.7	ThermoFisher
3661	36.0	ThermoFisher
3754	43.0	ThermoFisher
3845	31.4	ThermoFisher

[0136] Some of the human siRNAs against TBX3 in Table 2 are tested in human cell line models of NASH. The experimental design has three steps. First, use the identified siRNAs in human liver cancer cells, such as Huh7 or HepG2, to knockdown a target genes of interest, i.e. TBX3. Specifically, Huh7 cells are grown to 60-80% confluence and then transfected with siRNAs. Opti-MEM (Thermo Fisher 31985062) containing the siRNAs and Lipofectamine

are combined, incubated for 10 minutes, and then added to cells (400 µl/well to a 6-well plate or 2.4 ml to a 100 mm plate). Transfection medium is replaced after 6 hours with medium containing 100 µM FA. Then, cells are fed lipids and labeled with lipid dyes. Lastly, fluorescence-activated cell sorting (FACS) is performed to purify specific cell populations based on phenotypes detected by flow cytometry. This method enables characterization of a single cell population without the influence of other cells. Cells are stained with 1 ml PBS containing 10 µg of BODIPY® 493/503 (BD) (Thermo Fisher D3922) for 15 minutes at room temperature (23° C.±5° C.), then subjected to FACS measurements. Studies show that in lipophilic fluorophore stained cells, fluorescence intensity measured by FC reflects lipid levels.

[0137] An aliquot of the cells grown under each condition is assayed for triacylglycerol (TG). In brief, cells are trypsinized, washed once with PBS, resuspended in PBS with 10 mM EDTA, and then counted. Cells are divided in triplicate into 13×100 mm glass tubes in a final volume of 200 µl. Triolein standards (Sigma T7140) are also prepared in a final volume of 200 µl PBS/10 mM EDTA in 13×100 mm glass tubes. TGs are extracted and quantified. Briefly, 2 ml of isopropanol:hexane:water (40:10:1) is added to cells or standards and the samples are vortexed, covered, and incubated at room temperature for 30 min. Then, 500 µl of a 1:1 mixture of hexane:diethyl ether is added to the samples followed by vortexing and incubating for an additional 10 min at room temperature. Next, 1 ml of water is added to samples, tubes are vortexed, and layers are allowed to separate at room temperature while covered for 30-45 min. Using Pasteur pipettes, the top layer is transferred to 12×75 mm glass tubes and dried under N<sub>2</sub> to completion. Following the drying step, 400 µl of Infinity triglyceride reagent (Thermo Scientific TR22421) is added to each tube and vortexed. Tubes are covered and incubated for 90 minutes at 37° C. with shaking at 250 rpm. Finally, 300 µl of each sample is transferred to 96-well plates and absorbance is measured at 540 nm using a microplate reader. To determine whether fixing cells affected the TG measurement, TG is quantified in fixed and unfixed cells from the same batch of Huh7 cells incubated in 2 mM FA.

[0138] The tests described above were performed using optimized siRNAs and/or their conjugates with acetylgalactosamine (GalNAc) to improve liver targeting.

#### Example 5

[0139] Deletion of the transcription factor Tbx3 showed significant reductions in body weights (FIG. 4A) and this deletion resulted in one of the most substantial reductions in liver/body weight ratios compared to controls after 3 months of WD (dark dots in FIG. 4B). TBX3, a T-box transcription factor, has not been studied in fatty liver disease prior to the current disclosure. The KO model showed the most significant reductions in liver injury as measured by ALT (FIG. 4C), a trend toward reduced AST (FIG. 4D), and the most significant reductions in liver triglyceride levels (FIG. 4E). While Tbx3KO mice had reduced liver cholesterol, Bcl6 KO mice had increased liver cholesterol (FIG. 4F). Compared to controls and other KO models, there were clear improvements in hepatic steatosis in Tbx3, Bcl6, and Smyd2 KO livers (FIGS. 4G-4I). The phenotypes of Tbx3, Bcl6, and

*Smyd2* KO mice were comparable to *Srebf1* and *Dgat2* KO mice, suggesting potent regulation of lipid metabolism by these three genes.

[0140] To examine late-stage NASH features, some CRISPR KO models for 9 months with WD were fed (FIG. 5 and Table 4). In male *Tbx3* KO mice, liver weight and liver/body weight ratios decreased, while body weights were unchanged (FIG. 5A). Significant decreases of ALT and AST were also observed (FIG. 5B), suggesting reduced liver

damage. In male *Smyd2* KO mice, liver/body weight ratios decreased, but liver and body weights were unchanged (Figure S8A), and ALT/AST levels were unchanged (FIG. 5B). In the plasma, both *Tbx3* and *Smyd2* KOs showed moderate increases in triglyceride, but cholesterol levels remained similar to controls (FIG. 5C). These data showed that both KO models protect the liver from steatosis, but male *Tbx3* KO had more pronounced effects under long-term WD (FIG. 5D). Trends toward decreased fibrosis was observed in both KO models (FIGS. 5D and 5E).

TABLE 4

KO #	gene	sex	body (gram)	liver (gram)	liver/body (%)	AST (U/L)	ALT (U/L)	plasma-Cholesterol (mg/dL)	plasma-Triglyceride (mg/dL)
3671	lacZ	M	50.20	6.14	12.23	746	568	488	55.0
3676	lacZ	M	40.37	3.77	9.34	347	316	324	70.2
3679	lacZ	M	43.12	4.27	9.90	275	290	401	75.4
3680	lacZ	M	46.15	4.65	10.08	288	272	427	60.6
3681	lacZ	M	51.22	7.30	14.25	339	385	455	54.0
3682	lacZ	M	51.62	6.23	12.07	397	382	461	62.9
3683	lacZ	M	43.42	4.42	10.18	226	241	407	53.1
3732	smyd2	M	53.12	6.70	12.61	430	439	497	75.7
3737	smyd2	M	49.03	4.58	9.34	386	387	405	74.6
3739	smyd2	M	43.62	3.78	8.67	255	265	343	99.4
3740	smyd2	M	35.32	2.89	8.18	314	171	293	62.3
3741	smyd2	M	44.54	3.97	8.91	406	306	343	72.3
3755	smyd2	M	34.42	2.52	7.32	131	71	229	84.1
3757	smyd2	M	40.72	3.45	8.47	546	380	382	66.7
3758	smyd2	M	47.37	4.95	10.45	442	421	455	82.9
3673	tbx3	M	52.35	5.11	9.76	385	366	503	78.7
3674	tbx3	M	47.19	3.69	7.82	342	268	405	70.9
3687	tbx3	M	41.43	3.33	8.04	123	102	304	105.5
3688	tbx3	M	42.38	3.53	8.33	209	182	358	80.9
3691	tbx3	M	36.77	2.19	5.96	237	113	280	112.1
3692	tbx3	M	39.31	2.57	6.54	234	140	267	68.4
3703	tbx3	M	46.76	3.29	7.04	184	149	316	67.5
3709	tbx3	M	44.52	3.15	7.08	189	150	317	85.1
3730	tbx3	M	50.10	4.71	9.40	399	377	477	105.7
3731	tbx3	M	45.40	3.45	7.60	214	132	273	119.5
3735	tbx3	M	41.92	2.81	6.70	243	166	387	120.7
3738	tbx3	M	45.99	3.79	8.24			Failed to collect	
3742	tbx3	M	46.32	3.45	7.45	221	132	353	96.9
3901	gfp	F	41.48	3.25	7.84	319	190	275	64.4
3902	gfp	F	33.2	2.28	6.87	248	72	172	85.1
3909	gfp	F	41.26	2.55	6.18	150	99	252	73.8
3912	gfp	F	46.03	4.38	9.52	175	122	159	77.3
3916	gfp	F	47.73	5.29	11.08	266	246	376	69.1
3917	gfp	F	48.26	4.46	9.24	284	243	306	65.4
3920	gfp	F	36.58	2.45	6.70	149	75	245	73.8
3921	gfp	F	49.79	4.73	9.50	330	287	404	62.6
3933	gfp	F	35.28	2.26	6.41	108	63	126	68.3
3923	smyd2	F	43.72	3.33	7.62	227	154	280	47.5
3924	smyd2	F	43.98	3.95	8.98	302	225	385	72.6
3927	smyd2	F	37.02	2.33	6.29	204	108	307	77.4
3928	smyd2	F	40.72	2.19	5.38	176	114	228	74.8
3929	smyd2	F	36.14	2.46	6.81	250	114	186	55.0
3930	smyd2	F	34.96	1.75	5.01	109	52	152	55.8
3931	smyd2	F	25.57	1.41	5.51	82	36	103	63.9
3932	smyd2	F	42.68	2.72	6.37	164	98	270	74.7
3903	tbx3	F	37.94	2.18	5.75	213	114	220	86.0
3904	tbx3	F	34.08	1.87	5.49	152	34	104	75.0
3908	tbx3	F	41.64	3.12	7.49	173	139	221	92.6
3911	tbx3	F	30.04	1.74	5.79	188	122	121	83.6
3913	tbx3	F	43.43	3.11	7.16	294	183	300	72.5
3914	tbx3	F	46.27	3.56	7.69	331	285	365	82.6
3918	tbx3	F	54.82	4.91	8.96	404	271	365	64.8
3926	tbx3	F	45.53	2.9	6.37	245	213	287	106.4

### Example 6

**[0141]** To investigate the expression changes in WD fed livers carrying fitness promoting mutations, RNA-seq on control (sgGFP, sgLacZ), Irs1, Srebf1, Tbx3, Bcl6, and Smyd2 KO livers generated with MOSAICS AAVs was performed. Comparing differentially regulated genes in control vs. KO livers showed a relatively large intersection between Tbx3 and Irs1 KO livers (FIG. 6A), suggesting shared regulatory circuits. For lipid metabolism pathways in GSEA analysis, some shared but mostly unique expression patterns for each KO group were observed (FIG. 6B). The expression of genes involved in lipogenesis and fatty acid oxidation were further examined. Deletion of Irs1, Srebf1, Tbx3, Bcl6, and Smyd2 each led to decreased expression of fatty acid and triglyceride synthesis genes, but to different extents (FIG. 6C). The mRNA expression of collagen, a major component of fibrosis were also analyzed, and a downregulation of multiple collagen mRNAs in KO livers was observed (FIG. 6D). These data indicate shared and unique mechanisms by which mutant hepatocytes converge on decreased steatosis and fibrosis.

#### Experimental Methods and Subjects Used in the Examples

##### Mouse Strains and Breeding

**[0142]** All mice were handled in accordance with the guidelines of the Institutional Animal Care and Use Committee at UT Southwestern. All experiments were done in an age and sex controlled fashion unless otherwise noted. All mice used in this study were male. Genotyping and Sanger sequencing was used to confirm homologous recombination in the genome-edited pups. LSL-tdTomato (strain #007914) and Rosa-rtTA; TetO-Cas9 mice (#029415) were obtained from The Jackson Laboratory. Mice homozygous for both Rosa-rtTA and TetO-Cas9 were used to ensure a high Cas9 expression level in the liver. Western Diet (WD) used for NAFLD/NASH modeling is described in. It is composed of high fat solid food (ENVIGO #TD.120528) and high sugar water containing 23.1 g/L d-fructose (Sigma-Aldrich #F0127) and 18.9 g/L d-glucose (Sigma-Aldrich #G8270).

##### Fluorescent Imaging and Image Processing

**[0143]** For fluorescent imaging, liver pieces were fixed in buffered formalin (Fisherbrand #245-685) for 24 h with gentle shaking at 4° C. and then transferred into 30% sucrose (w/v) solution for another 24 h with shaking at 4° C. The livers were then embedded and frozen in Cryo-Gel (Leica #39475237), and sectioned at a thickness of 16 µm. Images were taken using a Zeiss Axionscan Z1 system in the UTSW Whole Brain Microscopy Facility to visualize Tomato clones. To statistically analyze the percentage of Tomato+ cells, black and white fluorescent images were taken from the same slide using an Olympus IX83 microscope at 4x magnification. Two different fields were taken for each liver. The percentage of Tomato+ cells (bright areas) was analyzed using ImageJ.

H&E, Immunohistochemistry (IHC), Immunofluorescence (IF), TUNEL, and Sirius Red Staining

**[0144]** Liver pieces were fixed in buffered formalin (Fisherbrand #245-685) for 24 h with gentle shaking at 4° C. and then transferred to 70% EtOH for another 24 h with shaking

at 4° C. Paraffin embedding, liver sectioning (4 µm thickness), and H&E staining were performed at the UT Southwestern Tissue Management Shared Resource Core. IHC was performed as previously described. Briefly, paraffin-embedded sections were dewaxed in xylene and hydrated using ethanol gradients. The slides were then boiled in antigen retrieval buffer (10 mM sodium citrate, 0.05% Tween 20, pH 6.0) for 20 min and soaked in 3% hydrogen peroxide (in methanol) for 10 min. After blocking with 5% goat serum for 1 h at room temperature, the slides were incubated with primary antibody overnight at 4° C. After washing, the slides were incubated with secondary antibody at room temperature for 0.5 h. The secondary antibody was coupled with HRP using ABC-HRP Kit (Vector laboratories, #PK-6101). The slides were developed using the DAB Kit (Vector laboratories, #SK-4100). For IF staining, the following primary antibodies were used: RFP (Rockland #600-401-379, IF 1:500); Ki67 (Invitrogen #14-5698-82, IF 1:500); HNF4a (Abcam #ab41898, IF 1:500), and the following secondary antibodies were used: Goat anti-rat IgG (H&L) Alexa Fluor Plus 488 (Invitrogen #A-48262, IF 1:500); Donkey anti-rabbit IgG (H&L) Alexa Fluor 594 (Invitrogen #A-21207, IF 1:500); Goat anti-mouse IgG2a Alexa Fluor 647 (Invitrogen #A-21241, IF 1:500). IF was performed on paraffin embedded mouse liver sections using the same protocol as IHC except that secondary antibodies were substituted by Alexa Fluor conjugated antibodies. TUNEL staining was performed on paraffin embedded liver sections using In Situ Cell Death Detection Kit, Fluorescein (Roche #C755B40) according to the manufacturer's protocol. Sirius Red staining was performed on paraffin embedded liver sections using the Picro Sirius Red Staining Kit (Abcam #ab150681) according to the manufacturer's protocol. QuPath software was used to quantify TUNEL staining and IHC staining of Ki67. ImageJ was used to quantify Sirius Red staining.

##### Plasma and Liver Metabolic Assays

**[0145]** Blood was taken using heparinized tubes from the inferior vena cava immediately after sacrificing the mouse, and then transferred into 1.5 ml tubes and centrifuged at 2000 g for 15 min at 4° C. The supernatant after centrifugation (plasma) was analyzed for AST, ALT, cholesterol, and triglyceride (Manufacturer's Reference Numbers 8433815, 1655281, 1669829, and 1336544, respectively) using a fully automated OCD Vitros 350 dry chemistry analyzer following the protocols provided by the reagent kit manufacturer (Ortho Clinical Diagnostics, Raritan, NJ) at the UT Southwestern Metabolic Phenotyping Core. 100-150 mg of liver per mouse was weighed and used for lipid extraction and quantification at the UT Southwestern Metabolic Phenotyping Core. Briefly, flash frozen tissue samples were homogenized with 2:1 chloroform:methanol mixture (v/v) using a multiplexed automatic tissue disruptor (TissueLYser II, Qiagen, Germantown, MD). The organic extract was transferred to a 5 ml graduated flask and the total volume was brought up to 5 ml. Total cholesterol and triacylglycerol analyses were performed in triplicate using 100 µl and 25 µl of lipid extracts, respectively. Total cholesterol and triacylglycerol concentrations were determined by commercial enzymatic colorimetric assays following the protocols described by the manufacturer (INFINITY™ Cholesterol Liquid Stable

Reagent #TR13421; INFINITY™ Triglycerides Liquid Stable Reagent #TR22421; Matrix PLUS™ Chemistry Reference Kit #NC9592194).

#### MOSAICS Reagent Construction

**[0146]** The MOSAICS plasmid uses the pX602 plasmid as a backbone. The sequence between the two AAV ITRs were removed using the NsiI and NotI restriction enzymes. The following fragments were cloned between the two AAV ITRs: the first SB100 binding IR, a U6 driven sgRNA scaffold, a CAG promoter driven SB100-P2A-Cre fusion cDNA with a beta-globin poly(A) signal, and the second SB100 binding IR. For library construction, mouse candidate genes for all of the *in vivo* screens were generated by using the mouse homologs of the human genes. A few genes were not included in the mouse gene lists due to the lack of a homolog or because they were known tumor suppressor genes. The individual sgRNA sequences corresponding to mouse candidate genes were extracted from the Brie library or obtained from the GUIDES server, and synthesized by CustomArray. Most genes had 5 distinct sgRNAs, 4 from Brie and 1 from GUIDES. A few genes had 4 targeting sgRNAs due to the overlap of sgRNA sequences from Brie and GUIDES. The library construction protocol was as well established in the field. Briefly, synthesized oligonucleotide libraries were amplified by PCR, purified using a PCR Purification Kit (Qiagen, #28104), and assembled in BsaI digested MOSAICS vector using DNA Assembly Kit (NEB, #E5520A). 1 µl of the assembled vector was then electroporated into 25 µl competent cells (Lucigen, #60052-2). After recovery in SOC medium for 1 h, bacteria were spread on a 245\*245 mm LB agar plate and incubated at 37° C. overnight. The bacteria were then harvested for plasmid preparation using the HiSpeed Plasmid Maxi Kit (Qiagen, #12663). Each sgRNA maintained a >1000-fold representation during construction. For individual sgRNA cloning, forward and reverse primers were annealed and fused to BsaI digested MOSAICS plasmid using T4 ligase. See Table 5 for the primers associated with the sgRNAs and the sgRNA.

TABLE 5

sgRNA primers	Sequence	SEQ ID NO
Tbx3 - sgRNA-F	CACCGGAGCACCTCACTTTAACCGG	116
Tbx3 - sgRNA-R	AAACCCGTTAAAGTGAGGTGCTCC	117
Mouse TBX3 - sgRNA1	GAGCACCTCACTTTAACCGG	318
Mouse TBX3 - sgRNA1	CATCATGGATCAGTTAGTGG	319
Mouse TBX3 - sgRNA1	AAAGAGAATGTATATACACC	320
Mouse TBX3 - sgRNA1	TCCCAAGCGATCACGCAACG	321

#### AAV Production and Purification

**[0147]** AAV8 was produced using AAV-Pro 293T cells (Takara #632273) cultured in one or more 15 cm dishes. Cells were plated one day before transfection at 50% confluence, which would allow the cells to reach 80-90% confluence the next day. For transfection of one 15 cm dish, 10 µg MOSAICS vector, 10 µg pAAV2/8 (Addgene #112864) and 20 µg pAdDeltaF6 (Addgene #112867) plas-

mids were mixed with 1 ml Opti-MEM medium in one tube. In another tube, 160 µl PEI solution (1 mg/ml in water, pH7.0, powder from ChemCruz #sc-360988) was mixed with 1 ml Opti-MEM medium. The solutions from both tubes were then mixed and incubated for 10 min before adding to cell culture. 48 h after transfection, the cells were scraped off the dish and collected by centrifugation at 500 g for 10 min. The supernatant was disinfected and discarded, and the cell pellets were lysed in 1.5 ml/15 cm dish lysis buffer (PBS supplemented with NaCl powder to final concentration of 200 mM, and with CHAPS powder to final concentration of 0.5% (w/v)). The cell suspension was put on ice for 10 min with intermittent vortexing, and then centrifuged at 20,000 g for 10 min at 4° C. The supernatant containing the AAV was collected. To set up the gravity column for AAV purification, 0.5 ml of AAV8-binding slurry beads (ThermoFisher #A30789), enough to purify AAV from three 15 cm dishes, was loaded into an empty column (Bio-Rad #731-1550). After the beads were tightly packed at the bottom, they were washed with 5 ml of wash buffer (PBS supplemented with NaCl powder to a final concentration of 500 mM). The supernatant containing AAV was then loaded onto the column. After all of the supernatant flowed through, the beads were washed with 10 ml wash buffer twice. The AAV was then eluted with 3 ml elution buffer (100 mM glycine, 500 mM NaCl in water, pH 2.5) and the eluate was immediately neutralized with 0.12 ml 1M Tris-HCl (pH 7.5-8.0). The AAV was concentrated by centrifugation at 2000 g for 3-5 min at 4° C. using an 100 k Amicon Ultra Centrifugal Filter Unit (Millipore #UFC810024). After centrifugation, the volume of AAV should be equal to or less than 0.5 ml. The concentrated AAV was diluted with 4-5 ml AAV dialysis buffer (PBS supplemented with powders to final concentrations of 212 mM NaCl and 5% sorbitol (w/v)) and centrifuged at 2000 g for 3-5 min at 4° C. The dilution and centrifugation processes were repeated 3 times. The final concentrated AAV was transferred into a 1.5 ml tube and centrifuged at 20,000 g for 5 min to remove debris. The supernatant was aliquoted, flash frozen using liquid nitrogen, and stored at -80° C.

#### Genomic DNA Extraction, sgRNA Amplification, and Amplicon Library Construction

**[0148]** To extract genomic DNA containing the integrated sgRNA, the entire liver (except a small piece used for sectioning and H&E staining) was minced into about 1 mm<sup>3</sup> pieces using a blade and weighed. Small nodules observed in some epigenetic factor screening livers given WD (<=3 nodules per liver in 5 out of 8 livers) were excluded from samples being processed for genomic DNA extraction.

Minced liver in two volumes (w/v) of homogenizing buffer (100 mM NaCl, 25 mM EDTA, 0.5% SDS, 10 mM Tris-HCl, pH 8) was transferred into a glass Wheaton Dounce Tissue Grinder and stroked 50 times or until no bulk tissues were seen. After homogenizing, 200 µl chow fed liver lysate or 300 µl WD fed liver lysate was transferred to a 15 ml tube for genomic DNA extraction using the Blood & Cell Culture DNA Midi Kit (Qiagen #13343) according to the manufacturer's protocol. The remaining lysates were frozen in -80° C. as backup samples. Briefly, 10 ml Buffer G2 from the kit, 100 µl Proteinase K (Roche #03115828001, or Proteinase K from the Qiagen kit) and 100 µl RNase A (Invitrogen #12091-021) were added to the 15 ml tube containing the lysate and digested in a 50° C. water bath overnight. The next day, the tubes were centrifuged at 4000 g for 10 min and the lipid layer on the top was discarded. The remaining supernatant was loaded on the column, washed, and genomic DNA elution/precipitation were performed according to the manufacturer's protocol. The precipitated DNA was resuspended in 100 µl 10 mM Tris (pH 8.0) and shaken on a 55° C. shaker for 2 h to help it dissolve. For amplicon library preparation, 5 pg genomic DNA, 5 µl general forward primer mix (5 pM), 5 µl barcode specific reverse primer (5 pM), 1 µl Q5 DNA polymerase, 10 µl Q5 buffer, 10 µl HighGC buffer, 1 µl dNTP, and water was mixed for a 50 µl PCR reaction, and two reactions were made for each genomic DNA sample. The PCR cycle was 95° C. 3 min-(95° C. 30 s-60° C. 30 s-72° C. 20 s)\*n-72° C. 2 min. The PCR cycle number was pre-optimized using the same PCR reactions with a smaller volume. The cycle numbers that gave a weak but sharp band on the DNA gel were used. In the final PCR reaction, 23 cycles were used for preparing the NASH gene, transcription factor, and epigenetic factor screens, and 30 cycles were used for preparing the guide mini-pool validation screen. After PCR, the two tubes of reactions with the same genomic DNA template were combined (total 100 µl) and 70 µl was resolved on a DNA gel. The 250 bp band corresponding to the amplicon was cut and purified using the QIAquick Gel Extraction Kit (Qiagen #28704). The DNA concentration was determined using Qubit kit (Invitrogen #Q32853) and high-throughput sequencing was performed using an Illumina NextSeq500 system at the CRI at UT Southwestern Sequencing Facility.

#### Bioinformatic Analysis of MOSAICS Screening Results

**[0149]** The reads from the sequencing of amplicon libraries described above were trimmed with cutadapt (version 1.9.1) to remove the excessive adaptor sequences so that only the sgRNA sequences were retained. The 5' sequences were trimmed with the options —O 32--discard-untrimmed-g CTTTATATCTTGTGGAAAGGACGAAACACCG. The 3' sequences were trimmed with the options —O 12-a GTTTTAGAGCTAGAAATAGCA. The abundance of each sgRNA was calculated with the count function in MAGeCK (version 0.5.6) with the default option. The trimmed fastq files were assigned to chow-fed and WD-fed groups and uploaded together with library files containing sgRNA sequences and targeted gene names to a server preloaded with MAGeCK. The enrichment of each sgRNA was calculated with the test function in MAGeCK.

#### RNA-Seq Library Preparation and Transcriptome Analysis of Mouse Fatty Livers

**[0150]** Total liver RNA was extracted from 4 sgGFP, 2 sgLacZ, 7 Tbx3, KO livers using TRIzol reagent (Invitrogen

#15596026) followed by purification using the RNeasy Mini kit (Qiagen #74014). Briefly, a liver fragment with a volume of about 3\*3\*3 mm<sup>3</sup> from each sample was homogenized in 1 ml Trizol, followed by adding 200 µl chloroform and vortexed. After centrifugation at 20,000 g for 10 min at 4° C., 350 µl supernatant from each sample was transferred to a new tube and mixed with equal volume of 75% EtOH, and then loaded on an RNeasy column. The following wash steps using RW1 and RPE buffers and RNA precipitation step were performed according to the manufacturer's protocol. RNA-seq libraries were prepared with the SMARTer Stranded Total RNA Sample Prep Kit—HI Mammalian (Takara #634875). 75 bp single-end sequencing was performed using an Illumina NextSeq500 system at the CRI at UT Southwestern Sequencing Facility. Alignment, quantification, and differential expression analysis were performed using the QBRC\_BulkRnaSeqDE pipeline ([https://github.com/QBRC/QBRC\\_BulkRnaSeqDE](https://github.com/QBRC/QBRC_BulkRnaSeqDE)). Briefly, the alignment of reads to the mouse reference genome (mm10) was done using (v2.7.2b) FeatureCounts (v1.6.4) was then used for gene count quantification. Differential expression analysis was performed using the R package DEseq2 (v1.26). Cutoff values of absolute fold change greater than 2 and FDR<0.05 were used to select for differentially expressed genes between sample group comparisons. Finally, GSEA was carried out with the R package fgsea (v1.14.0) using the 'KEGG' and 'Hallmark' libraries from MsigDB.

#### Transcriptomic Analysis of Human NASH Livers

**[0151]** Dataset 1 (NAFLD/NASH cohort): The RNA-Seq transcriptome profiles of biopsied liver tissues (GEO: GSE130970) were downloaded and analyzed 72 NAFLD/NASH patients with a range of disease severities (NASH activity scores of 1 to 6) and relevant histological features, i.e., steatosis, inflammation, fibrosis, and hepatocyte ballooning. The raw sequence reads were aligned to the GENCODE human reference genome (GRCh37, p13) using the STAR aligner (ver 2.6.1b), and gene-level count data were generated by the feature Counts function in the Subread package (ver 1.6.1), and the GENCODE genome annotation (GRCh37, v19). The count data were normalized using "Relative Log Expression" normalization (RLE) implemented in the DESeq2 package Dataset 2 (HCV cirrhosis cohort): The microarray gene expression profile of formalin-fixed needle biopsy specimens from the livers of 216 patients with hepatitis C-related early-stage (Child-Pugh class A) cirrhosis (GEO: GSE156540) was analyzed. This cohort was prospectively followed for a median of 10 years at an Italian center with relevant time event outcomes collected, including child, death, HCC and decomposition. Male and female patients were included in these studies.

#### Curating Transcription and Epigenetic Factors with Putative Pathogenic Activity in NAFLD/NASH

**[0152]** Gene expression of transcription factors do not directly reflect their functional activities, due to the low correlation between gene expression and protein abundance as well as the co-regulation with co-factors. Transcription factor activities can be estimated using enrichment of their downstream targets. However, traditional methods such as Gene Set Enrichment Analysis (GSEA) overestimate the number of significant transcription factors, largely confounded by the large number of overlaps among putative target genes of different TFs. Therefore, global modeling of putative transcription factors was performed with gene

expression data from NAFLD/NASH (GEO: GSE130970) and HCV cirrhosis patients (GEO: GSE15654) to directly infer the transcription factor activities from their downstream targets with adjustment for overlapping targets. A similar linear regression-based model was previously proposed to predict transcription factor regulatory activities and motifs from yeast gene expression data. The method regresses the fold-change of a gene on its putative regulatory transcription factor(s). The coefficient (Z score) of a transcription factor, estimated using genome-wide fold changes and predicted targets of all transcription factors, represents the regulatory activity change of the transcription factor across all the liver patients.

[0153] The regression model is defined as following:

$$f_q = \sum_i a_i T_{ig} + \sum_j \beta_i M_{jg} + c$$

where  $f_g$  is the fold change of g-th gene between two conditions;  $T_{ig}$  is the number of binding sites of i-th TF on the promoter of the g-th gene;  $M_{jg}$  is the number of binding sites of the j-th microRNA on the 3' UTR of the g-th gene; and  $\alpha_i$ ,  $\beta_j$  and c (a constant) can be inferred based on the values of  $f_g$ ,  $T_{ig}$  and  $M_{jg}$  for all the genes in the RNA-seq data. The Z scores of coefficients  $\alpha_i$  and  $\beta_j$  represent the activity changes of the i-th TF and j-th microRNA. Global transcription factor binding sites represented by 190 position-weighted matrices (PWMs) covering 500 mammalian TFs were based on the union of JASPAR, TRANSFAC and additional motifs from chromatin immunoprecipitation with DNA microarray and ChIP-seq data. The initial regression analysis was done using ISMARA before further integrative analysis with patient clinical histological features and time event outcome; sample-specific transcription factor activity

was estimated by the same regression model, where the fold changes were calculated between a single sample and all the samples combined together.

**[0154]** In the NAFLD/NASH cohort, the sample-level activities for each TF were associated with the four histological features, including fibrosis, inflammation, ballooning, and steatosis based on Pearson correlation. In the HCV cirrhosis cohort, the activities for each TF were used to perform outcome analysis on four-time events, including child, death, HCC and decomposition using cox proportional regression model. The p-values were calculated for both analyses respectively, followed by the calculation of False Discovery Rate (FDR) for multiple testing correction.

**[0155]** Similarly, to screen the putative pathogenic epigenetic regulators, the Pearson correlation was calculated between the gene expression of epigenetic regulators and the four histological features (fibrosis, inflammation, ballooning, and steatosis) in the NAFLD/NASH cohort, and used gene expression of known epigenetic regulators as independent variables to perform the cox proportional regression on four time events (Child-Pugh, death, HCC, decompensation) in the HCV cirrhosis cohort. The p-values were calculated for both analyses respectively, followed by the calculation of FDR for multiple testing correction.

## Quantification and Statistical Analysis

**[0156]** The data in most panels reflect multiple experiments performed on different days using mice derived from different litters. Variation in all panels is indicated using standard deviation presented as mean $\pm$ SD. Two-tailed unpaired Student's t-tests were used to test the significance of differences between two groups. Statistical significance is displayed as ns (not significant, or  $p>=0.05$ ), \*( $p<0.05$ ), \*\*( $p<0.01$ ), \*\*\*( $p<0.001$ ), \*\*\*\*( $p<0.0001$ ) unless specified otherwise. Image analysis for quantification was blinded.

SEQUENCE LISTING

---

-continued

---

acagcgccgg ccaagggtgga	agaggcgccgcgcgtccggg	gcaaggaggc cttcgcggccg	1440
ctcaccggtgc agaccgacgc	ggccgcggccgcg cacttggccc	aggccccct gcctggctc	1500
ggettcgcggccggccggcc	ggggcaacag ttcttcaacg	ggcaccggctt cttectgcac	1560
cccagcggagt ttggcatggg	ggggcccttc tccagcatgg	eggccgctgg catgggtccc	1620
ctcttgccca cggtttctgg	ggcctccacc ggtgtctcg	gcctggattc cacggccatg	1680
gctctcgccg ctgcggcgca	gggactgtccg ggggcgtccg	eggccacact gcccctccac	1740
ctccagcage acgttgcgc	ctctcagggc ctggccatgg	cccccttcgg aagctgttc	1800
ccttaccctt acacgtacat	ggccgcggccg cctctgggc	ctctctggc agcctccagc	1860
tccgtgeacc gccaccctt	ccteaatctg aacaccatgc	gccegggt gcgetacagc	1920
cctacttcca tcccggtgcc	ggccctggat ggccggatgc	tgtcaaccac cgcctgccc	1980
tccatgggggggggggggggg	ggccctggaa ggccaaatgg	ccggccctggc cgccagccg	2040
gctctcggtgg cagtggactc	gggtctgttcaacagc	gctctccac gctctccac	2100
agtcatgtt ctttgtcgcc	caaactctgc gcgagaaag	aggcgccac cagcgaactg	2160
caagacatcc agcggttgtt	tagggcttg gaagccaacg	eggacaggc ccgcagcg	2220
cccccg			2226
 SEQ ID NO: 2	moltype = RNA length = 19		
FEATURE	Location/Qualifiers		
source	1..19		
	mol_type = other RNA		
	organism = synthetic construct		
SEQUENCE: 2			
gctgatgact gtcgttata		19	
 SEQ ID NO: 3	moltype = RNA length = 19		
FEATURE	Location/Qualifiers		
source	1..19		
	mol_type = other RNA		
	organism = synthetic construct		
SEQUENCE: 3			
ggacatcgaa cctcaaaga		19	
 SEQ ID NO: 4	moltype = RNA length = 19		
FEATURE	Location/Qualifiers		
source	1..19		
	mol_type = other RNA		
	organism = synthetic construct		
SEQUENCE: 4			
ggacattata gctgctgat		19	
 SEQ ID NO: 5	moltype = RNA length = 19		
FEATURE	Location/Qualifiers		
source	1..19		
	mol_type = other RNA		
	organism = synthetic construct		
SEQUENCE: 5			
ggtaactcaga ctatattga		19	
 SEQ ID NO: 6	moltype = RNA length = 19		
FEATURE	Location/Qualifiers		
source	1..19		
	mol_type = other RNA		
	organism = synthetic construct		
SEQUENCE: 6			
ccacatttga agagccaat		19	
 SEQ ID NO: 7	moltype = RNA length = 19		
FEATURE	Location/Qualifiers		
source	1..19		
	mol_type = other RNA		
	organism = synthetic construct		
SEQUENCE: 7			
gggtgtttga tgaaaagaca		19	
 SEQ ID NO: 8	moltype = RNA length = 19		
FEATURE	Location/Qualifiers		
source	1..19		
	mol_type = other RNA		
	organism = synthetic construct		
SEQUENCE: 8			
caacatttca gacaaacat		19	
 SEQ ID NO: 9	moltype = RNA length = 19		
FEATURE	Location/Qualifiers		
source	1..19		
	mol_type = other RNA		

---

-continued

---

SEQUENCE: 9 caacaacatt tcagacaaa	organism = synthetic construct  19
SEQ ID NO: 10 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 10 ccaacaacat ttcagacaa	 19
SEQ ID NO: 11 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 11 gataagataa cccagttaa	 19
SEQ ID NO: 12 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 12 cgaatgttc ctccatcaa	 19
SEQ ID NO: 13 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 13 gaatgttcc tccatcaa	 19
SEQ ID NO: 14 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 14 ccaaatacat ttatttgat	 19
SEQ ID NO: 15 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 15 gaaactcacc aacaacatt	 19
SEQ ID NO: 16 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 16 ccaatgacat cttgaaact	 19
SEQ ID NO: 17 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 17 ggaaatggcc gaagagaaaa	 19
SEQ ID NO: 18 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 18 gataaaaaag ccaaataca	 19

---

-continued

---

SEQ ID NO: 19	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 19	
gccaaataca ttttattga	19
SEQ ID NO: 20	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 20	
cttgaaaactc ccttatagt	19
SEQ ID NO: 21	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 21	
aagataaccc agttaaaaaa	19
SEQ ID NO: 22	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 22	
catcgAACCT caaagattt	19
SEQ ID NO: 23	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 23	
ccaaAGAGGA tgtacatcc	19
SEQ ID NO: 24	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 24	
cattatagct gctgatgac	19
SEQ ID NO: 25	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 25	
tatagtacat ttccggacat	19
SEQ ID NO: 26	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 26	
catcgAACCT caaagattt	19
SEQ ID NO: 27	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 27	
gctgatgact gtcgttata	19
SEQ ID NO: 28	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA

---

-continued

---

SEQUENCE: 28 ccaatgacat cttgaaaact	organism = synthetic construct  19
SEQ ID NO: 29 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 29 ccattnaaag tgagatgtt	 19
SEQ ID NO: 30 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 30 gaaaactccct tatagtaca	 19
SEQ ID NO: 31 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 31 ccaacaacat ttcaagacaa	 19
SEQ ID NO: 32 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 32 gctgtatgact gtcgttata	 19
SEQ ID NO: 33 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 33 aagtggact attagacaa	 19
SEQ ID NO: 34 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 34 ccattnaaag tgagatgtt	 19
SEQ ID NO: 35 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 35 gataaccagg taaaaata	 19
SEQ ID NO: 36 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 36 gggtgtttga tgaaagaca	 19
SEQ ID NO: 37 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 37 ggctcacctt ttccgtgtct	 19

---

-continued

---

SEQ ID NO: 38	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 38	
gcagcttca actgcttcg	19
SEQ ID NO: 39	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 39	
cctcaaagat ttatgtccc	19
SEQ ID NO: 40	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 40	
ggagtgggca aaacacata	19
SEQ ID NO: 41	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 41	
tggccgaaga gaaaaaaga	19
SEQ ID NO: 42	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 42	
caaggcagtt tcaactgct	19
SEQ ID NO: 43	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 43	
aactccctta tagtacatt	19
SEQ ID NO: 44	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 44	
ggcctttgaa gaccatgga	19
SEQ ID NO: 45	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 45	
ggtgcacctg gaggctaaa	19
SEQ ID NO: 46	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 46	
ggctaaagaa cttagggat	19
SEQ ID NO: 47	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA

---

-continued

---

SEQUENCE: 47	organism = synthetic construct	
cgatcatgga tcaattgggt		19
SEQ ID NO: 48	moltype = RNA length = 19	
FEATURE	Location/Qualifiers	
source	1..19	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 48		
gatggacatt atagctgct		19
SEQ ID NO: 49	moltype = RNA length = 19	
FEATURE	Location/Qualifiers	
source	1..19	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 49		
gatgtccaaa gtcgtcact		19
SEQ ID NO: 50	moltype = RNA length = 19	
FEATURE	Location/Qualifiers	
source	1..19	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 50		
caaagaggat gtacattca		19
SEQ ID NO: 51	moltype = RNA length = 19	
FEATURE	Location/Qualifiers	
source	1..19	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 51		
gaaactccct tatagtaca		19
SEQ ID NO: 52	moltype = RNA length = 19	
FEATURE	Location/Qualifiers	
source	1..19	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 52		
cagactatat tgaactcca		19
SEQ ID NO: 53	moltype = RNA length = 19	
FEATURE	Location/Qualifiers	
source	1..19	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 53		
gtcattacca agtcgggaa		19
SEQ ID NO: 54	moltype = RNA length = 19	
FEATURE	Location/Qualifiers	
source	1..19	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 54		
ctaaagaact ttgggatca		19
SEQ ID NO: 55	moltype = RNA length = 19	
FEATURE	Location/Qualifiers	
source	1..19	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 55		
ccttatagta catttcgga		19
SEQ ID NO: 56	moltype = RNA length = 19	
FEATURE	Location/Qualifiers	
source	1..19	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 56		
ctatattgaa ctccatgca		19

---

-continued

---

SEQ ID NO: 57	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 57	
cgaaactgaa ttcatcgct	19
SEQ ID NO: 58	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 58	
ctcaaagatt tatgtccca	19
SEQ ID NO: 59	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 59	
tataacgaca gtcatcagc	19
SEQ ID NO: 60	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 60	
tctttgaggt tcgatgtcc	19
SEQ ID NO: 61	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 61	
atcagcagct ataatgtcc	19
SEQ ID NO: 62	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 62	
tcaatatagt ctgagtacc	19
SEQ ID NO: 63	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 63	
atggctctt acaatgtgg	19
SEQ ID NO: 64	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 64	
tgtctttcat caaacacccc	19
SEQ ID NO: 65	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 65	
atgtttgtct gaaatgttg	19
SEQ ID NO: 66	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA

---

-continued

---

SEQUENCE: 66 tttgtctgaa atgttgg	organism = synthetic construct	
		19
SEQ ID NO: 67 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 67 ttgtctgaaa tggttgg		19
SEQ ID NO: 68 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 68 ttaactgggt tatcttac		19
SEQ ID NO: 69 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 69 ttaatggagg aaacattcg		19
SEQ ID NO: 70 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 70 ttaaatggag gaaacattc		19
SEQ ID NO: 71 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 71 atcaataaaa tgtatgg		19
SEQ ID NO: 72 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 72 aatgttggt gtgagttc		19
SEQ ID NO: 73 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 73 agttcaaga tgtcatgg		19
SEQ ID NO: 74 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 74 tttctcttcg gccattcc		19
SEQ ID NO: 75 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 75 tgtatggc tttttatc		19

---

-continued

---

SEQ ID NO: 76	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 76	
tcaataaaaat gtatttggc	19
SEQ ID NO: 77	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 77	
actataaggg agtttcaag	19
SEQ ID NO: 78	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 78	
tttttaactg ggtttatctt	19
SEQ ID NO: 79	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 79	
aaatcttgaa ggttcgatg	19
SEQ ID NO: 80	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 80	
gaatgtacat cctctttgg	19
SEQ ID NO: 81	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 81	
gtcatcgaca gctataatg	19
SEQ ID NO: 82	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 82	
atgtccgaaa tgtactata	19
SEQ ID NO: 83	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 83	
aaatcttgaa ggttcgatg	19
SEQ ID NO: 84	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 84	
tataacgaca gtcatcagc	19
SEQ ID NO: 85	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA

---

-continued

---

SEQUENCE: 85 agtttcaaga tgtcattgg	organism = synthetic construct  19
SEQ ID NO: 86 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 86 aacatctac tttaaatgg	 19
SEQ ID NO: 87 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 87 tgtactataa gggagtttc	 19
SEQ ID NO: 88 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 88 ttgtctgaaa tggtgttgg	 19
SEQ ID NO: 89 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 89 tataacgaca gtcatcagc	 19
SEQ ID NO: 90 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 90 ttgtctaata gtctcaactt	 19
SEQ ID NO: 91 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 91 aacatctac tttaaatgg	 19
SEQ ID NO: 92 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 92 tatttttaac tgggttatac	 19
SEQ ID NO: 93 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 93 tgtctttcat caaacaccc	 19
SEQ ID NO: 94 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct
SEQUENCE: 94 agcagcgaaa aggtgagcc	 19

---

-continued

---

SEQ ID NO: 95	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 95	
cgaaggcagtt gaaagctgc	19
SEQ ID NO: 96	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 96	
gggacataaa tctttgagg	19
SEQ ID NO: 97	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 97	
tatgtgtttt gcccaactcc	19
SEQ ID NO: 98	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 98	
tcttttttctt cttcggcca	19
SEQ ID NO: 99	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 99	
agcagttgaa agctgttgt	19
SEQ ID NO: 100	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 100	
aatgtactat aaggaggatt	19
SEQ ID NO: 101	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 101	
tccatgttct tcaaaggcc	19
SEQ ID NO: 102	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 102	
tttagcctcc aggtgcacc	19
SEQ ID NO: 103	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 103	
atcccaaagt tcttttagcc	19
SEQ ID NO: 104	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA

---

-continued

---

SEQUENCE: 104 accaattgat ccatgatcg	organism = synthetic construct	
		19
SEQ ID NO: 105 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 105 agcagctata atgtccatc		19
SEQ ID NO: 106 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 106 agtgacgact ttggacatc		19
SEQ ID NO: 107 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 107 tgaatgtaca tcctcttg		19
SEQ ID NO: 108 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 108 tgtactataa gggagtttc		19
SEQ ID NO: 109 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 109 tggagttcaa tatagtcg		19
SEQ ID NO: 110 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 110 ttcccgactt ggtaatgac		19
SEQ ID NO: 111 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 111 tgatccaaa gttcttag		19
SEQ ID NO: 112 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 112 tccgaaatgt actataagg		19
SEQ ID NO: 113 FEATURE source	moltype = RNA length = 19 Location/Qualifiers 1..19 mol_type = other RNA organism = synthetic construct	
SEQUENCE: 113 tgcatggagt tcaatatacg		19

---

-continued

---

SEQ ID NO: 114	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 114	
agcgatgaat tcagttcg	19
SEQ ID NO: 115	moltype = RNA length = 19
FEATURE	Location/Qualifiers
source	1..19
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 115	
tgggacataa atcttttag	19
SEQ ID NO: 116	moltype = DNA length = 25
FEATURE	Location/Qualifiers
source	1..25
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 116	
caccggagca cctcacttta aacgg	25
SEQ ID NO: 117	moltype = DNA length = 25
FEATURE	Location/Qualifiers
source	1..25
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 117	
aaacccgttt aaagttaggt gctcc	25
SEQ ID NO: 118	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 118	
aaaaagggtga gccttgaccg	20
SEQ ID NO: 119	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 119	
cccccgctac tgggaaacag	20
SEQ ID NO: 120	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 120	
caagatctcc accaccacgt	20
SEQ ID NO: 121	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 121	
cgagggttag agcgacgccc	20
SEQ ID NO: 122	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 122	
tgtcattggc tcttacaatg	20
SEQ ID NO: 123	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA

---

-continued

---

SEQUENCE: 123 aaagaggatg tacattcacc	organism = synthetic construct  20
SEQ ID NO: 124 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 124 gttcgatgtc cctacagtgg	20
SEQ ID NO: 125 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 125 gaacatctca cttaaatgg	20
SEQ ID NO: 126 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 126 ccgaccccgaa atgccaaag	20
SEQ ID NO: 127 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 127 caaagaggag catggcccg	20
SEQ ID NO: 128 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 128 atccaggccc gagacaccgg	20
SEQ ID NO: 129 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 129 ttcaactggag gactcatcag	20
SEQ ID NO: 130 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 130 gaaaagctgct tgttcactgg	20
SEQ ID NO: 131 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 131 gagggttcgat gtccctacag	20
SEQ ID NO: 132 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 132 agtcacccct gcagtccatg	20

---

-continued

---

SEQ ID NO: 133	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 133	
gcccgagagca aagaggagca	20
SEQ ID NO: 134	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 134	
catccactgt tccccagtag	20
SEQ ID NO: 135	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 135	
ccatgcacaa ataccagccc	20
SEQ ID NO: 136	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 136	
ataaatttca caattctcg	20
SEQ ID NO: 137	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 137	
gcgtccgtct gcaccgtgag	20
SEQ ID NO: 138	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 138	
tgtctcgggc ctggattcca	20
SEQ ID NO: 139	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 139	
ggtgcccggt gaagaactgt	20
SEQ ID NO: 140	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 140	
tcacaattct cggtgatgg	20
SEQ ID NO: 141	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 141	
cgtatgtccct acagtggagg	20
SEQ ID NO: 142	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA

---

-continued

---

SEQUENCE: 142 ccccagaaac cgtggccagg	organism = synthetic construct  20
SEQ ID NO: 143 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 143 tttgtggaaa gtgacgactt	20
SEQ ID NO: 144 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 144 caggcccggag acaccgggtgg	20
SEQ ID NO: 145 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 145 ggacagtcggac tgcgcggcag	20
SEQ ID NO: 146 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 146 tggcccccggag gcctgcgacg	20
SEQ ID NO: 147 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 147 tctttcatca aacaccctca	20
SEQ ID NO: 148 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 148 ggaggagggcc tgcccgatca	20
SEQ ID NO: 149 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 149 ggggccgctc agcagcgaaa	20
SEQ ID NO: 150 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 150 gcgcaggaggc ccgggttcgac	20
SEQ ID NO: 151 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 151 ggtttccgggg acactggaaa	20

---

-continued

---

SEQ ID NO: 152	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 152	
agggagttc aagatgtcat	20
SEQ ID NO: 153	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 153	
ccggggctgg tatttgtca	20
SEQ ID NO: 154	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 154	
atttcacaat tctcggtgga	20
SEQ ID NO: 155	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 155	
cactcgccgc ctggggcgcc	20
SEQ ID NO: 156	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 156	
aggccccaga aaccgtggcc	20
SEQ ID NO: 157	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 157	
ccgcggggct gcccttgta	20
SEQ ID NO: 158	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 158	
gcggggccaac agttcttcaa	20
SEQ ID NO: 159	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 159	
tctccagccg cttccactgt	20
SEQ ID NO: 160	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 160	
taaaagtgaga tgttctgggc	20
SEQ ID NO: 161	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA

---

-continued

---

SEQUENCE: 161 agctttcaac tgcttcgccc	organism = synthetic construct  20
SEQ ID NO: 162 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 162 ttcggcatt tccagtggtcc	 20
SEQ ID NO: 163 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 163 gttggaaagct gcttggcac	 20
SEQ ID NO: 164 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 164 gccatgtcc tctttgtctt	 20
SEQ ID NO: 165 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 165 aattctcggt ggatggggc	 20
SEQ ID NO: 166 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 166 cgccgcgc acgggtcaga	 20
SEQ ID NO: 167 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 167 gggactgtcc gggggcgtccg	 20
SEQ ID NO: 168 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 168 gcaaaagggtt tccgggacac	 20
SEQ ID NO: 169 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 169 atgtttgtct gaaatgttgt	 20
SEQ ID NO: 170 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 170 ttacaatgtg gaaccggggc	 20

---

-continued

---

SEQ ID NO: 171	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 171	
caagggcagc cccgcggtca	20
SEQ ID NO: 172	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 172	
caaacacct catggactgc	20
SEQ ID NO: 173	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 173	
gtccctacag tggaggcgcc	20
SEQ ID NO: 174	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 174	
cctcttggc atttcggggt	20
SEQ ID NO: 175	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 175	
cagcactcgc ggcctggcgc	20
SEQ ID NO: 176	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 176	
aacaacattt cagacaaaaca	20
SEQ ID NO: 177	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 177	
gtttgatgaa agacacaaaaa	20
SEQ ID NO: 178	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 178	
gggcctggat tccacggcca	20
SEQ ID NO: 179	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 179	
caccggaca gcccccgtac	20
SEQ ID NO: 180	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA

---

-continued

---

SEQUENCE: 180	organism = synthetic construct
tgaagaactg ttggcccgcc	20
SEQ ID NO: 181	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 181	
ggcggggcta tggcgtgagt	20
SEQ ID NO: 182	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 182	
cccgccccca agccgaggcc	20
SEQ ID NO: 183	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 183	
gcacctggccc aggtgcgcgg	20
SEQ ID NO: 184	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 184	
tggcttac aatgtggAAC	20
SEQ ID NO: 185	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 185	
tgttgttag tttcagtttG	20
SEQ ID NO: 186	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 186	
tcgggtggatg gtggctggta	20
SEQ ID NO: 187	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 187	
aggccatggc cgtggaaatcc	20
SEQ ID NO: 188	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 188	
gggtgaatgt acatcccttt	20
SEQ ID NO: 189	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 189	
tccctgegcc gcagcggcag	20

---

-continued

---

SEQ ID NO: 190	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 190	
ggttcgcgag ggcacagcgc	20
SEQ ID NO: 191	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 191	
ggagatcttg gccgcgtcgc	20
SEQ ID NO: 192	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 192	
tggcaaaactg gctgggggtgc	20
SEQ ID NO: 193	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 193	
acgagatggt ggccggggcta	20
SEQ ID NO: 194	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 194	
tcctgtcgtc ctcgcgcgcc	20
SEQ ID NO: 195	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 195	
aactgttggc ccggcaggcc	20
SEQ ID NO: 196	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 196	
tatgtccgaa atgtactata	20
SEQ ID NO: 197	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 197	
tgtccagccg cccgcgtgtcc	20
SEQ ID NO: 198	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 198	
aggcggtgg agaagaagcc	20
SEQ ID NO: 199	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA

---

-continued

---

SEQUENCE: 199 cgcgcccttt ccacaccttggc	organism = synthetic construct  20
SEQ ID NO: 200 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 200 ctccaccagggt gtctcgggccc	20
SEQ ID NO: 201 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 201 tggggcctcc accgggtgtct	20
SEQ ID NO: 202 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 202 gttataaaatt tcacaatttct	20
SEQ ID NO: 203 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 203 ggcggtcgctc tcaccctcgcc	20
SEQ ID NO: 204 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 204 ggggctgccct ttgtcacggc	20
SEQ ID NO: 205 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 205 tccccagtag cggggctgtc	20
SEQ ID NO: 206 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 206 aactccctta tagtacattt	20
SEQ ID NO: 207 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 207 cgccgcgacgcg gcagaggccca	20
SEQ ID NO: 208 FEATURE source	moltype = DNA length = 20 Location/Qualifiers 1..20 mol_type = other DNA organism = synthetic construct
SEQUENCE: 208 cacagcgtatg aattcagttt	20

---

-continued

---

SEQ ID NO: 209	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 209	
gctgcggcgc agggactgtc	20
SEQ ID NO: 210	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 210	
ctcggtccagc actcgccggcc	20
SEQ ID NO: 211	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 211	
cttggccgcg tcgcaggcct	20
SEQ ID NO: 212	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 212	
ggcgctgtgc cctcgcaac	20
SEQ ID NO: 213	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 213	
ccagaacatc tcactttaaa	20
SEQ ID NO: 214	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 214	
gtccatcat aaaaatgtatt	20
SEQ ID NO: 215	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 215	
gtacatcctc tttggcattt	20
SEQ ID NO: 216	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 216	
ccatttaaag tgagatgttc	20
SEQ ID NO: 217	moltype = DNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 217	
cccctctgg ccacggttc	20
SEQ ID NO: 218	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA

---

-continued

---

```

SEQUENCE: 218      organism = synthetic construct
gctcagcagc gaaaagggtga gccttgaccg cggggggct gcc          43

SEQ ID NO: 219      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 219      organism = synthetic construct
caccggaca gccccgctac tggggAACAG tggtggatgt cca          43

SEQ ID NO: 220      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 220      organism = synthetic construct
gcgacgccc caagatctcc accaccacgt cggcgaggaa gcc          43

SEQ ID NO: 221      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 221      organism = synthetic construct
tatgtcccaag cgagggttag agcgacgccc aggaggccga gag          43

SEQ ID NO: 222      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 222      organism = synthetic construct
agtttcaaga tgtcatggc tcttacaatg tggtggaaacc ggg          43

SEQ ID NO: 223      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 223      organism = synthetic construct
ccgaaatgcc aaagaggatg tacattcacc cggcgacag ccc          43

SEQ ID NO: 224      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 224      organism = synthetic construct
taccttgag gttcgatgtc cctacagtgg aggaggccgc tgg          43

SEQ ID NO: 225      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 225      organism = synthetic construct
atccagccca gaacatctca cttaaatgg aggagggaaac att          43

SEQ ID NO: 226      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 226      organism = synthetic construct
gctggtaagg ccgaccccgaa aatgccaag aggaggatgt aca          43

SEQ ID NO: 227      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 227      organism = synthetic construct
aggccgagag caaagaggag catggcccg aggaggcctg cga          43

```

---

-continued

---

SEQ ID NO: 228	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 228	
tggccgtgga atccaggccc gagacaccgg tggtgaggc ccc	43
SEQ ID NO: 229	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 229	
aagctgcttg ttcaactggag gactcatcg aggaggccc att	43
SEQ ID NO: 230	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 230	
cgaaggcgtt gaaagctgct tttcaactgg aggaggactc atc	43
SEQ ID NO: 231	moltype = DNA length = 33
FEATURE	Location/Qualifiers
source	1..33
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 231	
gagggttcgtt gtccctacag tggtgaggc ggc	33
SEQ ID NO: 232	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 232	
cacagaaaac agctcaccct gcagtccatg aggagggtgt ttg	43
SEQ ID NO: 233	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 233	
cgacgcccag gccgagagca aagaggagca tggtgcccc gag	43
SEQ ID NO: 234	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 234	
cgactttgga catccactgt tccccactag cgccggggct gtc	43
SEQ ID NO: 235	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 235	
atattgaact ccatgcacaa ataccagccc cggcggttcc aca	43
SEQ ID NO: 236	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 236	
gactgtcggtt ataaatttca caattctcg tggtggtatgg tgg	43
SEQ ID NO: 237	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA

---

-continued

---

```

SEQUENCE: 237      organism = synthetic construct
cgcgccggcc gcgtccgtct gcaccgttag cggccgcgcg aag          43

SEQ ID NO: 238      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 238      organism = synthetic construct
ctcccacccg tgtctcgggc ctggattca cggccgcac ggc          43

SEQ ID NO: 239      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 239      organism = synthetic construct
aggaagagcg ggtgcccgtt gaagaactgt tggtgcccg cca          43

SEQ ID NO: 240      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 240      organism = synthetic construct
gttataaatt tcacaattct cggtagatgg tggtgctgg taa          43

SEQ ID NO: 241      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 241      organism = synthetic construct
ctttgagggtt cgatgtccct acagtggagg cggccgtgg aga          43

SEQ ID NO: 242      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 242      organism = synthetic construct
ccggtgtggagg ccccgaaaac cgtggccagg aggaggggac cca          43

SEQ ID NO: 243      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 243      organism = synthetic construct
tgagtttcag tttgtggaaa gtgacgactt tggtgacat cca          43

SEQ ID NO: 244      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 244      organism = synthetic construct
ccgtgaaatc caggccccgag acaccgtgg aggaggccc aga          43

SEQ ID NO: 245      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 245      organism = synthetic construct
cggacgcccc ggacagtccc tgcgcgcag cggccgcaga ggc          43

SEQ ID NO: 246      moltype = DNA  length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 246      organism = synthetic construct
aagaggagca tggccccgag gcctgcgcacg cggccgcacaa gat          43

```

---

-continued

---

SEQ ID NO: 247	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 247	ccttttgg tctttcatca aacaccctca tggtggaactg cag
	43
SEQ ID NO: 248	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 248	ccaccacgtc ggaggagccc tgccgtgaca aggagggcag ccc
	43
SEQ ID NO: 249	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 249	cgcgtgtccc gggcccgctc agcagcgaaa aggaggtgag cct
	43
SEQ ID NO: 250	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 250	gcgcggagga gcgcaggagc ccggttcgcg aggagggcac agc
	43
SEQ ID NO: 251	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 251	ttttgc当地 ggttccggg acactggaaa tggtggccga aga
	43
SEQ ID NO: 252	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 252	atgtactata agggagttc aagatgtcat tggtggctct tac
	43
SEQ ID NO: 253	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 253	caatgtggaa ccggggctgg tatttgtca tggtggagtt caa
	43
SEQ ID NO: 254	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 254	gtcggtataat atttcacaat tctcggtgga tggtgggtgc tgg
	43
SEQ ID NO: 255	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 255	tctcggtccag cactcgccgc ctggcgccgg aggaggagc cag
	43
SEQ ID NO: 256	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA

---

-continued

---

```

SEQUENCE: 256      organism = synthetic construct
acacccgggtgg aggccccaga aaccgtggcc aggaggaggg gac          43

SEQ ID NO: 257      moltype = DNA length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 257      tgagccttga ccgcggggct gcccttgtca cggccggcagg gct          43

SEQ ID NO: 258      moltype = DNA length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 258      cccggggcctg gcggggccaac agttttcaa cggccggcac ccg          43

SEQ ID NO: 259      moltype = DNA length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 259      ccaggcctct ttcaggccg cctccactgt aggagggaca tcg          43

SEQ ID NO: 260      moltype = DNA length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 260      ttccctccatt taaagtgaga ttttctggc tggtggtataa aaa          43

SEQ ID NO: 261      moltype = DNA length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 261      gtgaacaagg agctttcaac tgcttcgccc aggaggcttc ttc          43

SEQ ID NO: 262      moltype = DNA length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 262      ctttttctc ttcgccatt tccagtgtcc cggcggaaac ctt          43

SEQ ID NO: 263      moltype = DNA length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 263      gggcgaagca gttgaaagct gcttggcac tggtggtgggaa ctc          43

SEQ ID NO: 264      moltype = DNA length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 264      aggccctcggg gcatatgtcc tctttgtct cggccggcctc ggc          43

SEQ ID NO: 265      moltype = DNA length = 43
FEATURE           Location/Qualifiers
source            1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 265      taaatttcac aatttcgtggt ggatgggtggc tggtggttaag gcc

```

---

-continued

---

SEQ ID NO: 266	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 266	
aggaggcctt cgcgcgcgtc acgggtcaga cggcgacgc ggc	43
SEQ ID NO: 267	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 267	
ctgcggcgca gggactgtcc gggcgatcc cgccggccac cct	43
SEQ ID NO: 268	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 268	
caaccctttt gaaaaagggtt tccgggacac tggtgaaat ggc	43
SEQ ID NO: 269	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 269	
ttacaatcc atgttgtct gaaatgttgt tggtggttag ttt	43
SEQ ID NO: 270	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 270	
tcatggctc ttacaatgtc gaaccggggc tggtggtatt tgt	43
SEQ ID NO: 271	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 271	
cctggcgtga caagggcagc ccccggtca aggaggctca cct	43
SEQ ID NO: 272	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 272	
tgtcttcat caaacaccct catggactgc aggagggtga gct	43
SEQ ID NO: 273	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 273	
gaggttcgtat gtccctacag tggaggcgc tggtgagaaa gaa	43
SEQ ID NO: 274	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 274	
gaatgtacat cctctttggc atttcggtt cggcgccct acc	43
SEQ ID NO: 275	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA

---

-continued

---

```

SEQUENCE: 275          organism = synthetic construct
ccatctcgta cagcactcgcc ggcctggcg cggcgaggaa gcg           43

SEQ ID NO: 276          moltype = DNA  length = 43
FEATURE
source
1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 276
gaaaactcacc aacaacattt cagacaaaaca tggtggtt gta           43

SEQ ID NO: 277          moltype = DNA  length = 43
FEATURE
source
1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 277
ccatgagggt gtttgatgaa agacacaaaa aggaggagaa tgg           43

SEQ ID NO: 278          moltype = DNA  length = 43
FEATURE
source
1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 278
ccggtgttc gggcctggat tccacggca tggtgccctc tgc           43

SEQ ID NO: 279          moltype = DNA  length = 43
FEATURE
source
1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 279
gatgtacatt cacccggaca gccccgtac tggtggggaa cag           43

SEQ ID NO: 280          moltype = DNA  length = 43
FEATURE
source
1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 280
gggtgcccgt tgaagaactg ttggccgcg aggaggcccc ggg           43

SEQ ID NO: 281          moltype = DNA  length = 43
FEATURE
source
1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 281
acgagatggt ggcggggcta tggcggtagt cggcgccgca cgc           43

SEQ ID NO: 282          moltype = DNA  length = 43
FEATURE
source
1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 282
ccccggccagg cccggggcgca agccgaggcc aggaggcagg ggg           43

SEQ ID NO: 283          moltype = DNA  length = 43
FEATURE
source
1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 283
caggcagggg gccctgggcc aggtgcgcgg cggcgccgca gtc           43

SEQ ID NO: 284          moltype = DNA  length = 43
FEATURE
source
1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 284
aagatgtcat tggctttac aatgtgaaac cggcgccgca ggt           43

```

---

-continued

---

SEQ ID NO: 285	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 285	
tctgaaatgt tgttggtag tttcagttt tggtggaaag tga	43
SEQ ID NO: 286	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 286	
ttcacaaatc tcggtgatg gtggctggta aggaggccga ccc	43
SEQ ID NO: 287	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 287	
gcagcggcag aggccatggc cgttgcattc aggaggcccg aga	43
SEQ ID NO: 288	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 288	
ggggctgtcc ggggtaatgt acatcctttt tgggtggcatt tcc	43
SEQ ID NO: 289	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 289	
ccccggacac tccctgcgcc gcagcggcag aggaggccat ggc	43
SEQ ID NO: 290	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 290	
gcaggagccc ggttcgcgag ggcacagcgc cggcggccaa ggt	43
SEQ ID NO: 291	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 291	
acgtgggtt ggagatctt gccgcgtcgc aggaggcctc ggg	43
SEQ ID NO: 292	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 292	
gcgcgcgcgc tggcaaactg gctgggggtgc aggaggaaga gcg	43
SEQ ID NO: 293	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 293	
cgagtgttgg acgagatggt ggccgggtca tggtggcgtg agt	43
SEQ ID NO: 294	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA

---

-continued

---

SEQUENCE: 294	organism = synthetic construct
cgaaccgggc tcctgcgctc ctccgcgccc aggaggccgc gag	43
SEQ ID NO: 295	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 295	organism = synthetic construct
cccggttgaag aactgttggc cgcgcaggcc cgccggggcg aag	43
SEQ ID NO: 296	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 296	organism = synthetic construct
ggggaaacaag tatgtccgaa atgtactata aggagggagt ttc	43
SEQ ID NO: 297	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 297	organism = synthetic construct
ggcgacgttt tgtccagccg cccgtgtcc cgccggggcc gct	43
SEQ ID NO: 298	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 298	organism = synthetic construct
cctacagtggtt aggcggctgg agaagaagcc tgggtggcga agc	43
SEQ ID NO: 299	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 299	organism = synthetic construct
cgggagcgcg cgccgccttt ccaccttggc cgccggcgct gtg	43
SEQ ID NO: 300	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 300	organism = synthetic construct
tttctggggc ctccaccgggt gtctcgggcc tggtggattc cac	43
SEQ ID NO: 301	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 301	organism = synthetic construct
ccacgggttc tggggcctcc accgggtgtct cgccgggcct gga	43
SEQ ID NO: 302	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 302	organism = synthetic construct
gatgactgtc gttataaatt tcacaattct cggcggtgga tgg	43
SEQ ID NO: 303	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 303	organism = synthetic construct
tctcggtc ggcgtcgctc tcaccctcgcc tggtgggaca taa	43

---

-continued

---

SEQ ID NO: 304	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 304	
ccttgaccgc ggggctgccc ttgtcacggc aggagggctc ctc	43
SEQ ID NO: 305	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 305	
catccactgt tccccagtag cggggctgtc cggcggtga atg	43
SEQ ID NO: 306	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 306	
gacatcttga aactccctta tagtacattt cggcgacat act	43
SEQ ID NO: 307	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 307	
acagtccctg cgccgcagcg gcagaggcca tggtggccgt gga	43
SEQ ID NO: 308	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 308	
ggtatgcagt cacagegatg aattcagttt cggcgaaa caa	43
SEQ ID NO: 309	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 309	
ggcctctgcc gtcggcgcc agggactgtc cggcgaaaa tcc	43
SEQ ID NO: 310	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 310	
ccgcccacat ctgcgtccagc actcgccgc tggtggcgca gga	43
SEQ ID NO: 311	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 311	
tggtggagat ctggcccgcc tcgcaggct cggcgaaaa atg	43
SEQ ID NO: 312	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 312	
caccttggcc ggcgctgtgc cctcgcaac cggcgaaaa ctg	43
SEQ ID NO: 313	moltype = DNA length = 43
FEATURE	Location/Qualifiers
source	1..43
	mol_type = other DNA

---

-continued

---

```

SEQUENCE: 313          organism = synthetic construct
tttatccagc ccagaacatc tcacttaaa tggtgaggaa aac           43

SEQ ID NO: 314          moltype = DNA length = 43
FEATURE
source
1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 314
cagctataat gtccatcaat aaaatgtatt tggtggttt ttt           43

SEQ ID NO: 315          moltype = DNA length = 43
FEATURE
source
1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 315
ccgggtgaat gtacatccctc tttggcattt cggcggggtc ggc           43

SEQ ID NO: 316          moltype = DNA length = 43
FEATURE
source
1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 316
aatgttctt ccatttaaag tgagatgttc tggtgggctg gat           43

SEQ ID NO: 317          moltype = DNA length = 43
FEATURE
source
1..43
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 317
tggcatgggt cccctctgg ccacggttc tggtggggcc tcc           43

SEQ ID NO: 318          moltype = DNA length = 20
FEATURE
source
1..20
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 318
gagcacctca cttaaacgg                                     20

SEQ ID NO: 319          moltype = DNA length = 20
FEATURE
source
1..20
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 319
catcatggat cagtttagtgg                                     20

SEQ ID NO: 320          moltype = DNA length = 20
FEATURE
source
1..20
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 320
aaagagaatg tatatacaccc                                     20

SEQ ID NO: 321          moltype = DNA length = 20
FEATURE
source
1..20
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 321
tcccaagcga tcacgcaacg                                     20

```

---

What is claimed is:

1. A composition comprising a nucleic acid that down-regulates expression of TBX3 (T-box transcription factor 3) or a variant thereof.
2. The composition of claim 1, wherein the nucleic acid that downregulates expression of TBX3 comprises a siRNA,

a cluster regularly interspaced short palindromic repeats (CRISPR) related nucleic acid, a single guide RNA (sgRNA), a CRISPR-RNA (crRNA), or a trans-activating crRNA (tracrRNA).

3. The composition of claim 2 wherein the nucleic acid is a small interfering RNA (siRNA) molecule.

4. A composition comprising a plasmid or a viral vector, wherein the plasmid or the viral vector comprises a nucleic acid encoding the siRNA molecule of claim 3.
5. The composition of any one of claim 3 or 4, wherein the siRNA molecule comprises a nucleotide sequence that is 2 to 30 nucleotides in length and is at least 80% homologous to at least 2 to 30 contiguous nucleotides of a human TBX3 cDNA sequence.
6. The composition of claim 5, wherein the human TBX3 cDNA sequence comprises SEQ ID NO: 1.
7. The composition of any one of claim 3 or 4, wherein the siRNA molecule targets the open reading frame or the 5' or 3' UTRs of the TBX3 gene.
8. The composition of any one of claims 3-7, wherein the siRNA molecule comprises at least one sense sequence, at least one antisense sequence, or at least one sense sequence and at least one antisense sequence.
9. The composition of any one of claims 3-8, wherein the siRNA molecule comprises a nucleotide sequence SEQ ID NOs: 2-115 or any combination thereof.
10. The composition of claim 9, wherein the at least one sense sequence comprises SEQ ID NOs: 2-58.
11. The composition of claim 9, wherein the at least one antisense sequence comprises SEQ ID NOs: 59-115.
12. The composition of claim 2, wherein the nucleic acid that downregulates expression of TBX3 is a sgRNA molecule optionally selected from a group consisting of SEQ ID NO: 118-217.
13. A composition comprising a plasmid or a viral vector, wherein the plasmid or viral vector comprises a first nucleic acid encoding the sgRNA molecule of claim 12 and optionally a second nucleic acid encoding an RNA guided nuclelease.
14. The composition of claim 13, wherein the RNA guided nuclease is a Cas endonuclease.
15. The composition of any one of claims 3-11 wherein the siRNA molecule specifically downregulates gene expression of at least one variant of TBX3.
16. The composition of any one of claims 12-14, wherein the sgRNA molecule specifically downregulates gene expression of at least one variant of TBX3.
17. The composition of any one of claim 15 or 16, wherein the at least one variant of TBX3 is associated with a liver disease.
18. The composition of claim 17, wherein the liver disease comprises fatty liver disease (FLD), alcohol-related liver disease (ARLD), non-alcoholic fatty liver disease (NAFLD), non-alcoholic steatohepatitis (NASH), end stage liver disease (cirrhosis) from any etiology, liver cancer, or any combination thereof.
19. The composition of any one of claims 1-3 wherein the nucleic acid molecule is conjugated to least one targeting ligand.
20. The composition of claim 19, wherein the at least one targeting ligand comprises a liver targeting ligand.
21. The composition of claim 20, wherein the liver targeting ligand comprises at least one N-acetylgalactosamine (GalNAc) conjugate.
22. The composition of claim 21, wherein the nucleic acid molecule is conjugated to about one to about three GalNAc conjugates.
23. The composition of any one of claims 1-3 or 12, wherein the nucleic acid that specifically downregulates expression of TBX3 comprises at least one chemical modification.
24. The composition of claim 23, wherein the nucleic acid comprises a modification at least one ribosugar moiety of its nucleotide sequence.
25. The composition of claim 24, wherein at least one ribosugar moiety is modified with 2'-O-methyl (2'OMe), 2'-deoxy-2'-fluoro (2F), 2'-deoxy, 5-C-methyl, 2'-O-(2-methoxyethyl) (MOE), 4'-thio, 2'-amino, 2'-C-allyl, or any combination thereof.
26. The composition of either claim 24 or claim 25, wherein less than about 10% to about 70% of ribosugar moieties of the total nucleotide sequence is modified.
27. A pharmaceutical composition comprising any one of the compositions of claims 1-26 and at least one pharmaceutically acceptable carrier.
28. The pharmaceutical composition of claim 27, further comprising a nanoparticle.
29. The pharmaceutical composition of either claim 27 or claim 28, further comprising a lipid.
30. A method for treating a subject in need thereof, the method comprising administrating a therapeutically effective amount of the composition of any one of claims 1-26 or the pharmaceutical composition of any one of claims 27-29 to the subject in need thereof.
31. The method of claim 30, wherein the subject in need thereof, is a human subject having or suspected of having a liver disease.
32. The method of claim 31, wherein the liver disease comprises fatty liver disease (FLD), alcohol-related liver disease (ARLD), non-alcoholic fatty liver disease (NAFLD), non-alcoholic steatohepatitis (NASH), end stage liver disease (cirrhosis) from any etiology, liver cancer, or any combination thereof.
33. The method of any one of claims 30-32, wherein the method of administering comprises parenteral administration.
34. The method of any one of claims 30-33, wherein administration of a therapeutically effective amount of the composition of any one of claims 1-26 or the pharmaceutical composition of any one of claims 27-29 increases life expectancy of the subject compared to an untreated subject with identical disease condition and predicted outcome.
35. The method of any one of claims 30-33, wherein administration of a therapeutically effective amount of the composition of any one of claims 1-26 or the pharmaceutical composition of any one of claims 27-29 improves liver function of the subject compared to an untreated subject with identical disease condition and predicted outcome.
36. The method of any one of claims 30-33, wherein administration of a therapeutically effective amount of the composition of any one of claims 1-26 or the pharmaceutical composition of any one of claims 27-29 attenuates liver fibrosis in the subject compared to an untreated subject with identical disease condition and predicted outcome.
37. The method of any one of claims 30-33, wherein administration of a therapeutically effective amount of the composition of any one of claims 1-26 or the pharmaceutical composition of any one of claims 27-29 prevents additional liver fibrosis in the subject compared to an untreated subject with identical disease condition and predicted outcome.

**38.** A kit comprising:

- a. a container holding the composition of any one of claims **1-26** or the pharmaceutical composition of any one of claims **27-29**;
- b. a pharmaceutical administrative means; and
- c. an instruction.

\* \* \* \*