

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
10 May 2007 (10.05.2007)

PCT

(10) International Publication Number
WO 2007/053184 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/US2006/021062

(22) International Filing Date: 31 May 2006 (31.05.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/686,135 31 May 2005 (31.05.2005) US

(71) Applicant (for all designated States except US): **COLD SPRING HARBOR LABORATORY** [US/US]; One Bungtown Road, Cold Spring Harbor, NY 11724 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **DICKINS, Ross** [AU/US]; 1 Bungtown Road, James Building, Cold Spring Harbor, NY 11724 (US). **HANNON, Gregory, J.** [US/US]; 34 Griffith Lane, Huntington, NY 11743 (US). **LOWE, Scott, W.** [US/US]; One Bungtown Road, Cold Spring Harbor, NY 11724 (US).

(74) Agent: **VINCENT, Matthew, P.**; Fish and Neave IP Group of, Ropes & Gray LLP, One International Place, Boston, MA 02110-2624 (US).

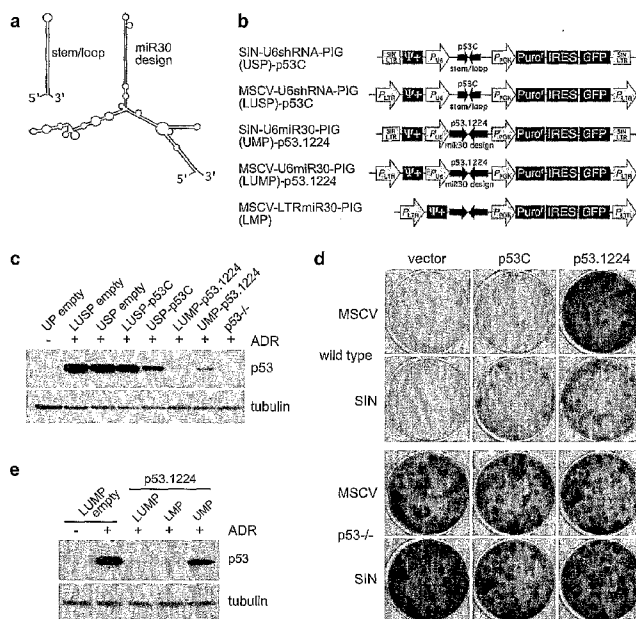
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS FOR PRODUCING MICRORNAS



(57) Abstract: The invention relates to recombinant vectors for inducible and/or tissue specific expression of double-stranded RNA molecules that interfere with the expression of a target gene. In certain embodiments, the invention relates to the use of Tet (tetracycline)-responsive RNA Polymerase II (Pol II) promoters (e.g., TetON or TetOFF) to direct inducible knockdown in certain cells of an integrated or an endogenous gene, such as p53. The invention also relates to a method for producing transgenic animals (e.g., mice) expressing inducible (such as tetracycline-regulated), reversible, and/or tissue-specific double-stranded RNA molecules that interfere with the expression of a target gene.

WO 2007/053184 A2

METHODS FOR PRODUCING MICRORNAS

Reference to Related Application

This application claims the benefit of the filing date of U.S. Provisional Application Serial No. 60/ 686,135, entitled "METHODS FOR PRODUCING
5 MICRORNAS," and filed on May 31, 2005. The teachings of the entire referenced application is incorporated herein by reference.

Government Support

Work described herein was funded, in whole or in part, by Mouse Models of Human Cancer Consortium Grant No. 25480211. The United States government has
10 certain rights in the invention.

Background of the Invention

RNA interference (RNAi) has been used to silence the expression of a target gene. RNAi is a sequence-specific post-transcriptional gene silencing mechanism triggered by double-stranded RNA (dsRNA). It causes degradation of mRNAs
15 homologous in sequence to the dsRNA. The mediators of the degradation are 21-23-nucleotide small interfering RNAs (siRNAs) generated by cleavage of longer dsRNAs (including hairpin RNAs) by DICER, a ribonuclease III-like protein. Molecules of siRNA typically have 2-3-nucleotide 3' overhanging ends resembling the RNase III processing products of long dsRNAs that normally initiate RNAi.
20 When introduced into a cell, they assemble an endonuclease complex (RNA-induced silencing complex), which then guides target mRNA cleavage. As a consequence of degradation of the targeted mRNA, cells with a specific phenotype of the suppression of the corresponding protein product are obtained (*e.g.*, reduction of tumor size, metastasis, angiogenesis, and growth rates).

25 The small size of siRNAs, compared with traditional antisense molecules, prevents activation of the dsRNA-inducible interferon system present in mammalian cells. This helps avoid the nonspecific phenotypes normally produced by dsRNA larger than 30 base pairs in somatic cells. See, *e.g.*, Elbashir *et al.*, *Methods* 26:199-

213 (2002); McManus and Sharp, Nature Reviews 3:737- 747 (2002); Hannon, Nature 418:244-251 (2002); Brummelkamp *et al.*, Science 296:550-553 (2002); Tuschl, Nature Biotechnology 20:446-448 (2002); U.S. Application US2002/0086356 A1; WO 99/32619; WO 01/36646; and WO 01/68836.

5 Summary of the Invention

One aspect of the invention provides an artificial nucleic acid construct comprising an RNA Polymerase II (Pol II) promoter operably linked to a coding sequence for expressing a precursor molecule for an siRNA, the siRNA inhibiting the expression of a target gene, wherein the nucleic acid construct directs the expression of the precursor molecule and/or the siRNA, and substantially inhibits the expression of the target gene when the artificial nucleic acid construct is stably integrated into a host cell genome.

In certain embodiments, the Pol II promoter is an inducible promoter, a tissue-specific promoter, or a developmental stage-specific promoter. For example, the inducible promoter may be a tetracyclin-responsive promoter, including commercially available TetON promoter (the transcription from which promoter is activated at the presence of tetracyclin (tet), doxycycline (Dox), or tet analog), or the TetOFF promoter (the transcription from which promoter is turned off at the presence of tetracyclin (tet), doxycycline (Dox), or a tet analog), *e.g.*, those from Clontech, Inc.

In other embodiments, the inducible promoter may be selected from: a promoter operably linked to a lac operator (LacO), a LoxP-stop-LoxP system promoter, or a GeneSwitchTM or T-RExTM system promoter (Invitrogen), or equivalents thereof with identical or substantially similar mechanisms.

In yet other embodiments, the Pol II promoter can be any art-recognized Pol II promoters, such as an LTR promoter or a CMV promoter.

In certain embodiments, the precursor molecule may be a precursor microRNA, such as an artificial miR comprising coding sequence for the siRNA for the target gene. For example, the miR may comprise a backbone design of microRNA-30 (miR-30). Alternatively, the miR may comprise a backbone design of miR-15a, -16, -19b, -20, -23a, -27b, -29a, -30b, -30c, -104, -132s, -181, -191, -223.

See US 2005/0075492A1 (incorporated herein by reference).

In other embodiments, the precursor molecule may be a short hairpin RNA (shRNA).

The constructs of the instant invention is highly potent, and a single
5 integrated copy of the subject nucleic acid construct is sufficient for substantially inhibiting the expression of the target gene. "Substantially inhibiting" as used herein includes inhibiting at least about 20%, or about 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99% or close to 100% of the expression or mRNA and/or protein of the target gene.

10 The constructs of the invention may further comprise an enhancer for the Pol II promoter.

The constructs of the invention may further comprise a reporter gene under the control of a second promoter, such as a luciferase, a fluorescent protein (*e.g.* GFP, RFP, YFP, BFP, *etc.*), or an enzyme, or any other art-recognized reporter
15 whose physical presence and/or activity can be readily assessed using an art-recognized method.

In certain embodiments, the second promoter and the reporter gene can be downstream of (3'-to) the coding sequence for the precursor molecule. In other
20 embodiments, the reporter gene is translated from an internal ribosomal entry site (IRES) between a second promoter and the reporter gene.

In other embodiments, the coding sequence for expressing the precursor molecule may be embedded or inserted into the 5'-UTR (5'-untranslated region), 3'-UTR, or an intron of the reporter gene.

The constructs of the invention may further comprise at least one selectable
25 marker, such as puromycin, zeocin, hygromycin, or neomycin, *etc.*

The constructs of the invention may further comprise a Pol III promoter upstream of the coding sequence for expressing the precursor molecule.

The constructs of the invention can be used to inhibit the expression of a number of different target genes. In certain embodiments, the target gene is
30 associated with a disease condition such as cancer or infectious disease. For example, the target gene may be over-expressed or abnormally active in the disease. In addition, the target gene may be an oncogene or an antagonist / inhibitor or

dominant negative mutation of a tumor suppressor gene.

Another aspect of the invention provides a cell comprising any of the subject nucleic acid constructs.

In certain embodiments, the cell may be a mammalian cell.

5 In certain embodiments, the cell may be a tissue culture cell (*e.g.*, a primary cell, or a cell from an established cell line), a cell *in vivo*, or a cell manipulated *ex vivo*.

If the Pol II promoter is an inducible promoter, the cell may further comprise an additional construct for expressing an activator or an inhibitor of the inducible promoter. For example, if the inducible promoter is a tet-responsive promoter, the
10 additional construct may encode tTA or rtTA. If the inducible promoter is a LacO-responsive promoter, the additional construct may encode LacI. If the inducible promoter is a LoxP-stop-LoxP system promoter, the additional construct may encode a Cre recombinase, which may be under the transcriptional control of an
15 inducible promoter, a developmental stage-specific promoter, or a tissue-specific promoter.

Another aspect of the invention provides a non-human mammal comprising any of the subject cells described above. In certain embodiments, the non-human mammal may be a chimeric mammal some of whose somatic or germ cells are
20 subject cells as described above. Alternatively, the non-human mammal may be a transgenic mammal all of whose somatic or germ cells are subject cells described above.

Another aspect of the invention provides a method for making a subject chimeric non-human mammal as described above, comprising introducing a
25 construct according to any of the subject nucleic acid constructs into an embryonic stem (ES) cell and generating a chimeric mammal from the ES cell.

Another aspect of the invention provides a method for making a subject transgenic non-human mammal described above, comprising mating a subject chimeric non-human mammal described above with another animal from the same
30 species.

Another aspect of the invention provides a method for inhibiting the expression of a target gene of interest in a cell, comprising introducing a subject

construct into the cell, wherein the siRNA molecule derived from the precursor molecule is specific for the target gene.

In certain embodiments, the method further comprises inhibiting at least one additional target gene(s) of interest in the cell by introducing at least one additional
5 constructs according to any one of the subject nucleic acid constructs into the cell, wherein each of the siRNA molecules derived from the precursor molecules are specific for the additional target genes, respectively.

Another aspect of the invention provides a method for treating a gene-mediated disease, comprising introducing into an individual having the disease a
10 construct according to any of the subject nucleic acid constructs, where the siRNA derived from the precursor molecule is specific for the gene mediating the disease.

Another aspect of the invention provides a method of validating a gene as a potential target for treating a disease, comprising: (1) introducing a construct according to any one of the subject nucleic acid constructs described herein into a
15 cell associated with the disease, wherein the siRNA molecule derived from the precursor molecule is specific for the gene; (2) assessing the effect of inhibiting the expression of the gene on one or more disease-associated phenotype; wherein a positive effect on at least one disease-associated phenotype is indicative that the gene is a potential target for treating the disease.

In certain embodiments, the gene is over-expressed or abnormally active in
20 disease cells or tissues. Alternatively, the gene may be downstream of and is activated by a second gene over-expressed or abnormally active in disease cells or tissues. In addition, the product of the gene antagonizes an suppressor of a second gene over-expressed or abnormally active in disease cells or tissues.

In certain embodiments, the cell may be a tissue culture cell, such as a
25 primary cell isolated from diseased tissues, or from an established cell line derived from diseased tissues.

In other embodiments, the cell is within diseased tissues, and step (2) above comprises evaluating one or more symptoms of the disease.

In certain embodiments, the cell may be one from a transgenic animal, such
30 as one comprising any of the subject nucleic acid constructs.

For example, in a transgenic animal with a transgene comprising any of the

subject nucleic acid constructs, the transgene may encode a precursor molecule, which, upon processing, generates a siRNA specific for the candidate target gene. Preferably, the expression of the precursor molecule is inducible, reversible, and/or tissue-specific.

5 In certain embodiments, the method further comprises assessing the side effect, if any, of knocking down the expression of the target gene in one or more tissues / organs other than the diseased tissue, wherein the target gene is a valid target if the side effect, if any, is acceptable to a person of skill in the respective art (e.g., when validating a drug target, such side effects resulting from impairment of
10 the target gene function in other tissues must be acceptable to a physician or veterinarian).

 In certain embodiments, the expression of the gene may be inducibly inhibited by a subject construct, or inducibly activated by turning down the expression of a subject construct.

15 It is also contemplated that all embodiments of the invention, including those specifically described for different aspects of the invention, can be combined with any other embodiments of the invention as appropriate.

 Other features and advantages of the invention will be apparent from the following detailed description and claims.

20 Brief Description of the Drawings

 Figure 1 shows effective knockdown via single copy expression of miR30-based shRNAs from a retroviral LTR promoter. (A) Schematic representation of predicted RNA folds for simple stem/loop and miR30 design shRNAs. Note extensive predicted folding for the ~300 nt pre-miR30 RNA. Folds were generated
25 using mfold. (B) Retroviral vectors used to deliver shRNAs to mammalian cells. Provirus layouts are shown to indicate promoter activity of the integrated virus. Active promoters are shown as open arrows, with two inverted black arrows representing shRNA stem sequences. (C) Western blot analysis for p53 expression of NIH3T3 cells transduced with the retroviral vectors shown in B and selected in
30 puromycin. A tubulin blot is shown as a loading control. (D) Colony formation assay for the cells shown in (C). Cells were seeded in 6 well plates at 2500

cells/well, and allowed to grow for 10 days before harvesting. (E) Western blot analysis for p53 expression in NIH3T3 cells transduced at less than 5% efficiency (assessed by GFP FACS; not shown) with the retroviral vectors shown in (B). A tubulin blot is shown as a loading control. Similar results were obtained in other cell types including wild type and p19ARF-null MEFs (data not shown).

Figure 2 shows that RNA polymerase II-driven shRNAs can effectively promote tumorigenesis and chemotherapy resistance *in vivo*. (A) Kaplan–Meier curve showing mouse survival following adoptive transfer of E μ -Myc HSCs infected with LTR-driven Bim shRNAs. (B) Western blot showing reduced BimEL and BimL expression in E μ -Myc lymphomas expressing Bim shRNAs. Control: archived tumors arising from E μ -Myc HSCs (on either a wild type, ARF^{+/-} or p53^{+/-} background; not shown) were used as controls for Bim expression. (C) Kaplan–Meier curves showing tumor-free survival (left) and overall survival (right) for mice harboring p19ARF-null lymphomas infected with either the LMP-p53.1224 retrovirus or vector control. Tumor-bearing mice were given a single 10 mg/kg dose of adriamycin at day zero. (D) Flow cytometry analysis of GFP expression in lymphoma cells harvested from the mice in (A). Representative histograms show the percent of GFP-positive cells at the time of treatment (left) and after tumor relapse (right).

Figure 3 shows stable and regulatable shRNA expression from a tet-responsive RNA polymerase II promoter. (A) Provirus layout of the SIN-TREmiR30-PIG (TMP) retroviral vector. (B) Western blot analysis of Rb expression in HeLa-tTA cells infected with TMP-Rb.670. Cells were treated with 100 ng/mL Dox for 4 days prior to harvesting. Control uninfected HeLa-tTA cells treated with Dox are also shown. (C) Rb expression in homogeneous cultures derived from single-cell clones of HeLa-tTA cells infected at single copy with TMP-Rb.670. Cells were cultured in normal, Dox-free medium prior to harvesting. con = control uninfected HeLa-tTA cells. (D) Dox dose/response analysis of Rb expression in HeLa-tTA clone Rb.670C. Cells were cultured for 8 days in the indicated Dox concentration prior to harvesting. Control uninfected HeLa-tTA cells (con) cultured with or without Dox are also shown. Note the presence of a non-specific band in the GFP immunoblot, running just below GFP. (E) Rb expression in HeLa-tTA clone

Rb.670C cells over time in response to shifting into or out of Dox. Cells were cultured without Dox (left panels) or in 100 ng/mL Dox (right panels) for eight days prior to shifting them into 100 ng/mL Dox or Dox-free medium, respectively. Again, note the presence of a faint non-specific band in the GFP immunoblot. Similar results were observed for all Rb.670 clones showing good Rb knockdown in Dox-free medium (C, above), with some clonal variation in kinetics. In addition, similar results were obtained using other mishRNAs targeting Rb and PTEN (data not shown). (F) Rb expression in homogeneous cultures derived from single-cell clones of U2OS-rtTA cells infected at around 1% efficiency with TMP-Rb.670. Cells were cultured in 1000 ng/mL Dox for several days prior to harvesting. (G) Dox dose/response analysis of Rb expression in U2OS-rtTA clone Rb.670R5 cells. Cells were cultured for 8 days in the indicated Dox concentration prior to harvesting. Control uninfected U2OS-rtTA cells cultured with or without Dox are also shown. (H) Rb expression in U2OS-rtTA clone Rb.670R5 cells in response to Dox treatment. Cells were cultured without Dox for eight days prior to shifting them into 1000 ng/mL Dox. Note that Rb.670R5 cells express some GFP in Dox-free medium, and Rb levels are slightly decreased compared with controls, indicating slightly leaky expression from the TRE-CMV promoter in this particular clone.

Figure 4 shows reversible p53 knockdown in primary MEFs. (A) Colony formation assays of wild type MEFs doubly infected with TMP-p53.1224 and tTA. Cells were seeded in 6 well plates at 5000 cells/well, and grown for 8 days before harvesting. Upper wells contained Dox-free medium, whereas lower wells contained 100 ng/mL Dox. Positive control p53-null MEFs are shown, as are negative control wild type MEFs infected with TMP-p53.1224 alone or with TMP-PTEN.1010 plus tTA. (B) Western blot analysis of p53 and GFP expression in cells expanded from a single-cell clone of wild type MEFs infected with TMP-p53.1224 and tTA (WtT cells). Cells were cultured in 100 ng/mL Dox for various times prior to harvesting. (C) Morphology and GFP fluorescence of WtT cells originally plated at colony formation density, and cultured in Dox-free medium (upper panels) or 100 ng/mL Dox (lower panels). Right panel: SA- β -gal staining of WtT cells cultured in Dox-free medium (upper) or 100 ng/mL Dox (lower). (D) Left panel: Colony formation assay for WtT cells cultured for 8 days in 100 ng/mL Dox, then seeded in Dox free

medium (upper well) or 100 ng/mL Dox (lower well). Right panel: Colony formation assay of cells equivalent to those in the upper well of the left panel (formerly Dox-treated, dormant WtT cells after extended culture in Dox-free medium). Cells were seeded and harvested as in (A). (E) Morphology, GFP fluorescence, and SA- β -gal staining of WtT cells infected with Ras and cultured in normal medium (upper panels) or 100 ng/mL Dox (lower panels). (F) Western blot analysis of p53 and GFP expression in WtT cells infected with Ras. Cells were cultured in 100 ng/mL Dox for various times prior to harvesting.

Figure 5 shows regulated p53 knockdown in tumors. (A) GFP and standard imaging of representative tumor-bearing nude mice, with Dox treatment commencing at day 0 (lower panels). Untreated controls are shown (upper panels). (B) Representative tumor growth curves for WtT-Ras tumors in an untreated mouse (open squares), or a mouse treated for 10 days with Dox (filled circles indicate Dox treatment) commencing at day 0. Each data point is the average volume of 2 tumors for a single mouse. Similar results were obtained for 8 different WtT-Ras clones, with slightly differing kinetics. (C) Representative tumor growth curves for WtT-E1A/Ras tumors in an untreated mouse (open squares), or a mouse treated for 7 days with Dox (filled circles indicate Dox treatment) commencing at day 0. Each data point is the average volume of 2 tumors for a single mouse. Insets show GFP status of a single tumor at various times. (D) Histological analysis of cell morphology and apoptosis in representative nude mouse tumors harvested from untreated mice or mice treated with Dox for several days. (E) Western blot analysis of p53 and GFP expression in representative WtT-Ras and WtT-E1A/Ras tumors harvested from untreated mice or mice treated with Dox for several days. Cultured WtT-Ras cells treated with Dox are shown as a control.

Figure 6 shows an siRNA northern blot of tissues isolated from animals of various genotypes, probed with a labelled oligonucleotide that hybridizes to the p53.1224 siRNA. The individual lanes are: **M**: Molecular weight marker; **1**: LAP-tTA liver; **2**: LAP-tTA;TRE-1224 liver; **3**: LAP-tTA;TRE-1224 liver after 4 days Doxycycline administration; **4**: LAP-tTA spleen; **5**: LAP-tTA;TRE-1224 spleen; **6**: LAP-tTA;TRE-1224 spleen after 4 days Doxycycline administration; **7**: Eu-myc;TRE-1224 mouse 4-1 spleen; **8**: Eu-myc;TRE-1224 mouse 6-4 spleen; **9**: Eu-

myc;Eu-tTA;TRE-1224 mouse #1 spleen; **10**: Eu-myc;Eu-tTA;TRE-1224 mouse #2 spleen; **11**: Eu-myc;Eu-tTA;TRE-1224 mouse #1-2 spleen; **12 & 13**: Spleen (**12**) and lymph node (**13**) from a tumor-bearing nude mouse recipient of Eu-myc lymphoma cells; **14 & 15**: Spleen (**14**) and lymph node (**15**) from a tumor-bearing nude mouse recipient of Eu-myc;Eu-tTA;TRE-1224 lymphoma cells; **16 & 17**: Spleen (**16**) and lymph node (**17**) from a tumor-bearing nude mouse recipient of Eu-myc;Eu-tTA;TRE-1224 lymphoma cells, after 14 days Doxycycline administration.

Detailed Description of the Invention

I. *Overview*

10 RNA interference (RNAi) is normally triggered by double stranded RNA (dsRNA) or endogenous microRNA precursors (pre-miRNAs). Since its discovery, RNAi has emerged as a powerful genetic tool for suppressing gene expression in mammalian cells. Stable gene knockdown can be achieved by expression of synthetic short hairpin RNAs (shRNAs), traditionally from RNA polymerase III
15 promoters.

The instant invention generally relates to the use of RNA Polymerase II promoters to express microRNA (miRNA) precursors and/or short hairpin RNAs (shRNAs), either *in vitro*, *ex vivo*, or *in vivo*, especially from as few as one single stably integrated expression construct. The single expression construct may be stably
20 transfected / infected into a target cell, or may be a germline transgene. Transgenic animals with the subject RNAi constructs, which may be regulated to express mishRNA in an inducible, reversible, and/or tissue-specific manner, can be used to establish valuable animal models for certain disease, such as those associated with loss-of-function of certain target genes. The ability to control both the timing (*e.g.*,
25 at certain developmental stages) and location (*e.g.*, tissue-specific) of target gene knock-down, including the ability to reverse the course of induction / inactivation, renders the subject system a powerful tool to study gene function and disease progression. Such animal models or cells thereof may also be used for drug screening or validation.

30 In certain embodiments, Pol II promoters controls the transcription of the subject miRNA / shRNA coding sequence. In general, any Pol II compatible

promoters may be used for the instant invention.

In certain embodiments, various inducible Pol II promoters may be used to direct precursor miRNA / shRNA expression. Exemplary inducible Pol II promoters include the tightly regulatable Tet system (either TetOn or TetOFF), and a number
5 of other inducible expression systems known in the art and/or described herein. The tet systems allows incremental and reversible induction of precursor miRNA / shRNA expression *in vitro* and *in vivo*, with no or minimal leakiness in precursor miRNA / shRNA expression. Such inducible system is advantages over the existing unidirectional Cre-lox strategies. Other systems of inducible expression may also be
10 used with the instant constructs and methods.

In certain embodiments, expression of the subject miRNA / shRNA may be under the control of a tissue specific promoter, such as a promoter that is specific for: liver, pancreas (exocrine or endocrine portions), spleen, esophagus, stomach, large or small intestine, colon, GI tract, heart, lung, kidney, thymus, parathyroid,
15 pineal gland, pituitary gland, mammary gland, salivary gland, ovary, uterus, cervix (*e.g.*, neck portion), prostate, testis, germ cell, ear, eye, brain, retina, cerebellum, cerebrum, PNS or CNS, placenta, adrenal cortex or medulla, skin, lymph node, muscle, fat, bone, cartilage, synovium, bone marrow, epithelial, endothelial, vesicular, nervous tissues, *etc.* The tissue specific promoter may also be specific for
20 certain disease tissues, such as cancers. See Fukazawa *et al.*, *Cancer Research* **64**: 363-369, 2004 (incorporated herein by reference).

Any tissue specific promoters may be used in the instant invention. Merely to illustarte, Chen *et al.* (Nucleic Acid Research, Vol. 34, database issue, pages D104-D107, 2006) described TiProD, the Tissue-specific Promoter Database
25 (incorporated herein by reference). Specifically, TiProD is a database of human promoter sequences for which some functional features are known. It allows a user to query individual promoters and the expression pattern they mediate, gene expression signatures of individual tissues, and to retrieve sets of promoters according to their tissue-specific activity or according to individual Gene Ontology
30 terms the corresponding genes are assigned to. The database have defined a measure for tissue-specificity that allows the user to discriminate between ubiquitously and specifically expressed genes. The database is accessible at tiprod.cbi.pku dot

edu.cn:8080/index.html. It covers most (if not all) the tissues described above.

In certain embodiments, if the reversibly inducible systems of the invention are used, the subject shRNAs are not designed to target the promoter regions of a target gene to avoid irreversible TGS.

5 In certain embodiments, artificial miRNA constructs based on, for example, miR30 (microRNA 30), may be used to express precursor miRNA / shRNA from single / low copy stable integration in cells *in vivo*, or through germline transmission
10 in transgenic animals. For example, Silva *et al.* (*Nature Genetics* **37**: 1281-88, 2005, incorporated herein by reference) have described extensive libraries of pri-miR-30-based retroviral expression vectors that can be used to down-regulate almost all known human (at least 28,000) and mouse (at least 25,000) genes (see RNAi Codex, a single database that curates publicly available RNAi resources, and provides the most complete access to this growing resource, allowing investigators to see not only released clones but also those that are soon to be released, available at
15 <http://codex.cshl dot edu>). Although such libraries are driven by Pol III promoters, they can be easily converted to the subject Pol II-driven promoters (see Methods in Dickins *et al.*, *Nat. Genetics* **37**: 1289-95, 2005; also see page 1284 in Silva *et al.*, *Nat. Genetics* **37**: 1281-89, 2005).

In certain embodiments, even a single copy of stably integrated precursor
20 miRNA / shRNA construct results in effective knockdown of a target gene.

In certain embodiments, the inducible Tet system, coupled with the low-copy integration feature of invention, allows more flexible screening applications, such as in screening for potentially lethal shRNAs or synthetic lethal shRNAs.

In certain embodiments, the subject precursor miRNA cassette may be
25 inserted within a gene encoded by the subject vector. For example, the subject precursor miRNA coding sequence may be inserted with an intron, the 5'- or 3'-UTR of a reporter gene such as GFP, *etc.*

In certain embodiments, cultured cells, such as wild type mouse fibroblasts or primary cells can be switched from proliferative to senescent states simply
30 through regulated knockdown of p53 using the subject constructs and methods.

The constructs and methods of the invention is advantageous in several respects.

In one respect, stable precursor miRNA / shRNA expression may be effected through retroviral or lentiviral delivery of the miRNA / shRNAs, which is shown to be effective at single copy per cell. This allows very effective stable gene expression regulation at extremely low copy number per cell (*e.g.* one per cell), thus vastly
5 advantageous over systems requiring the introduction of a large copy number of constructs into the target cell by, for example, transient transfection.

Compare to transfection where there are multiple copies (such as multiple episomal copies) of the shRNA construct, and the LTR is active, the instant system is preferable for stable expression of the shRNA.

10 Using the instant system, Applicants have discovered rapid and coordinated entry into senescence upon re-establishment of wild type p53 expression in p53 defective cells. Such an observation would not have been possible using previous technologies.

Another useful feature of the invention is that it is compatible with an
15 established miR30 miRNA / shRNA library, which contains designed miRNA / shRNA constructs targeting almost all human and mouse genes. Any specific member of the library can be readily cloned (such as by PCR) into the vectors of the instant invention for Pol II-driven regulated and stable expression.

Other vector designs with different promoters have shown dependence on
20 position of transcriptional start and stop sites. The subject method / system apparently has no such stringent requirements.

Applicants have also discovered that promoter interference between Pol II and Pol III promoters may prevent efficient transcription of encoded shRNA, while the use of miRNA precursor has largely overcome this problem.

25 Another aspect of the invention provides a method for drug target validation. The outcome of inhibiting the function of a gene, especially the associated effect *in vivo*, is usually hard to predict. Gene knock-out experiments offer valuable data for this purpose, but is expensive, time consuming, and potentially non-informative since many genes are required for normal development, such that loss-of-function
30 mutation in such genes causes embryonic lethality. Using the methods of the instant invention, especially the inducible expression regulation system of the invention, any potential drug target / candidate gene for therapeutic intervention may be tested

first by selectively up- and/or down-regulating their expression *in vitro*, *ex vivo*, or *in vivo*, and determining the effect of such regulated expression, especially *in vivo* effects on an organism. If disruption of the normal expression pattern of a candidate gene shows desired phenotypes *in vitro* and/or *in vivo*, the candidate gene is chosen
5 as a target for therapeutic intervention. Various candidate compounds can then be screened to identify inhibitors or activators of such validated targets.

Another aspect of the invention provides a method to determine the effect of coordinated expression regulation of two or more genes. For example, miRNA / shRNA constructs for two more target genes may be introduced into a target cell
10 (*e.g.*, by stable integration) or an organism (*e.g.*, by viral vector infection or transgenic techniques), and their expression may be individually or coordinately regulated using the inducible and/or tissue specific or developmental specific promoters according to the instant invention. Since different inducible promoters are available, the expression of the two or more target genes may be regulated either in
15 the same or opposite direction (*e.g.*, both up- or down-regulating, or one up one down, *etc.*). Such experiments can provide useful information regarding, *inter alia*, genetic interaction between related genes.

In certain embodiments, the instant invention allows highly efficient knockdown of a target gene from a single (retroviral) integration event, thus
20 providing a highly efficient means for certain screening applications. For example, the instant system and methods may be used to test potentially lethal miRNA / shRNAs or synthetic lethal miRNA / shRNAs.

The invention also provide a method to treat certain cancer, especially those cancer overexpressing Ras pathway genes (*e.g.*, Ras itself) and having impaired p53
25 function, comprising introducing into such cells an active p53 gene or gene product to induce senescence and/or apoptosis, thereby killing the cancer cells, or at least inhibit cancer progression and/or growth.

The general feature of the invention having been described, the following section provides certain illustrative aspects of the invention that may be combined in
30 specific embodiments. Other similar or equivalent art-recognized methods may also be readily adapted for use in the instant invention.

II. *MicroRNA and RNAi Design*

DNA vectors that express perfect complementary short hairpins RNAs (shRNAs) are commonly used to generate functional siRNAs. However, the efficacy of gene silencing mediated by different short-hairpin derived siRNAs may be
5 inconsistent, and a substantial number of short-hairpin siRNA expression vectors can trigger an anti-viral interferon response (*Nature Genetics* 34: 263, 2003). Moreover, siRNA short-hairpins are typically processed symmetrically, in that both the functional siRNA strand and its complement strand are incorporated into the RISC complex. Entry of both strands into the RISC can decrease the efficiency of
10 the desired regulation and increase the number of off-target mRNAs that are influenced. In comparison, endogenous microRNA (miRNA) processing and maturation is a fairly efficient process that is not expected to trigger an anti-viral interferon response. This process involves sequential steps that are specified by the information contained in miRNA hairpin and its flanking sequences.

15 MicroRNAs (miRNAs) are endogenously encoded ~22-nt-long RNAs that are generally expressed in a highly tissue- or developmental-stage-specific fashion and that post-transcriptionally regulate target genes. More than 200 distinct miRNAs having been identified in plants and animals, these small regulatory RNAs are believed to serve important biological functions by two prevailing modes of action:
20 (1) by repressing the translation of target mRNAs, and (2) through RNA interference (RNAi), that is, cleavage and degradation of mRNAs. In the latter case, miRNAs function analogously to small interfering RNAs (siRNAs). Importantly, miRNAs are expressed in a highly tissue-specific or developmentally regulated manner and this regulation is likely key to their predicted roles in eukaryotic development and
25 differentiation. Analysis of the normal role of miRNAs will be facilitated by techniques that allow the regulated over-expression or inappropriate expression of authentic miRNAs *in vivo*, whereas the ability to regulate the expression of siRNAs will greatly increase their utility both in cultured cells and *in vivo*. Thus one can design and express artificial microRNAs based on the features of existing
30 microRNA genes, such as the gene encoding the human miR-30 microRNA. These miR30-based shRNAs have complex folds, and, compared with simpler stem/loop style shRNAs, are more potent at inhibiting gene expression in transient assays.

miRNAs are first transcribed as part of a long, largely single-stranded primary transcript (Lee *et al.*, *EMBO J.* **21**: 4663–4670, 2002). This primary miRNA transcript is generally, and possibly invariably, synthesized by RNA polymerase II (pol II) and therefore is normally polyadenylated and may be spliced. It contains an
5 ~80-nt hairpin structure that encodes the mature ~22-nt miRNA as part of one arm of the stem. In animal cells, this primary transcript is cleaved by a nuclear RNaseIII-type enzyme called Drosha (Lee *et al.*, *Nature* **425**: 415–419, 2003) to liberate a hairpin miRNA precursor, or pre-miRNA, of ~65 nt, which is then exported to the cytoplasm by exportin-5 and the GTP-bound form of the Ran cofactor (Yi *et al.*,
10 *Genes Dev.* **17**: 3011–3016, 2003). Once in the cytoplasm, the pre-miRNA is further processed by Dicer, another RNaseIII enzyme, to produce a duplex of ~22 bp that is structurally identical to an siRNA duplex (Hutvagner *et al.*, *Science* **293**: 834–838, 2001). The binding of protein components of the RNA-induced silencing complex (RISC), or RISC cofactors, to the duplex results in incorporation of the mature,
15 single-stranded miRNA into a RISC or RISC-like protein complex, whereas the other strand of the duplex is degraded (Bartel, *Cell* **116**: 281–297, 2004).

The miR-30 architecture can be used to express miRNAs or siRNAs from pol II promoter-based expression plasmids. See also Zeng *et al.*, *Methods in Enzymology* **392**: 371–380, 2005 (incorporated herein by reference).

20 Figure 2B of Zeng (*supra*) shows the predicted secondary structure of the miR-30 precursor hairpin (“the miR-30 cassette”). Boxed are extra nucleotides that were added originally for subcloning purposes (Zeng and Cullen, *RNA* **9**: 112–123, 2003; Zeng *et al.*, *Mol. Cell* **9**: 1327–1333, 2002). They represent XhoI–BglII sites at the 50 end and BamHI–XhoI sites at the 30 end. These appended nucleotides
25 extend the minimal miR-30 precursor stem shown by several basepairs, similar to the *in vivo* situation where the primary miR-30 precursor is transcribed from its genomic locus (Lee *et al.*, *Nature* **425**: 415–419, 2003), and an extended stem of at least 5 bp is essential for efficient miR-30 production. Based on the numbering in Fig. 2B, mature miR-30 is encoded by nucleotides 44 to 65 and anti-miR-30 by
30 nucleotides 3 to 25 of this precursor. In the simplest expression setting, the cytomegalovirus (CMV) immediate early enhancer/promoter may be used to transcribe the miR-30 cassette. The cassette is preceded by a leader sequence of

approximately 100 nt and followed by approximately 170 nt before the polyadenylation site (Zeng *et al.*, *Mol. Cell* 9: 1327–1333, 2002). These lengths are arbitrary and can be longer or shorter. Mature 22-nt miR-30 can be made from such constructs.

5 Several other authentic miRNAs have been over-expressed by using analogous RNA pol II-based expression vectors or even pol III-dependent promoters (Chen *et al.*, *Science* 303: 83–86, 2004; Zeng and Cullen, *RNA* 9: 112–123, 2003). Expression simply requires the insertion of the entire predicted miRNA precursor stem-loop structure into the expression vector at an arbitrary location. Because the
10 actual extent of the precursor stem loop can sometimes be difficult to accurately predict, it is generally appropriate to include ~50 bp of flanking sequence on each side of the predicted ~80-nt miRNA stem-loop precursor to be sure that all cis-acting sequences necessary for accurate and efficient Drosha processing are included (Chen
et al., *Science* 303: 83–86, 2004).

15 In an exemplary embodiment, to make the miR-30 expression cassette, the sequence from +1 to 65 (excluding the 15-nt terminal loop of the miR-30 cassette, Fig. 2B of Zeng) may be replaced as follows: the sequence from nucleotides 39 to 61, which is perfectly complementary to a target gene sequence, will act as the active strand during RNAi. The sequence from nucleotides 2 to 23 is thus designed
20 to preserve the double-stranded stem in the miR-30-target cassette, but nucleotide +1 is now a C, to create a mismatch with nucleotide 61, a U, just like nucleotides 1 and 65 in the miR-30 cassette (Fig. 2B). Because the 30 arm of the stem (miR-30-target) is the active component for RNAi, changes in the 50 arm of the stem will not affect RNAi specificity. A 2-nt bulge may be present in the stem region of the authentic
25 miR-30 precursor (Fig. 2B of Zeng). A break in the helical nature of the RNA stem may help ward off nonspecific effects, such as induction of an interferon response (Bridge *et al.*, *Nat. Genet.* 34: 263–264, 2003) in expressing cells. This may be why miRNA precursors almost invariably contain bulges in the predicted stem. The miR-30 cassette in Fig. 2A of Zeng is then substituted with the miR-30-target cassette,
30 and the resulting expression plasmid can be transfected into target cells.

 The use of pol II promoters, especially when coupled with an inducible expression system (such as the TetOFF system of Clontech) offers flexibility in

regulating the production of miRNAs in cultured cells or *in vivo*. Selection of stable cell lines leads to less leaky expression in the absence of the activator or presence of doxycycline, and therefore a stronger induction.

In certain embodiments, it would be advantageous if the antisense strand, for example, of the above miR-30-target construct is preferentially made as a mature miRNA, because its opposite strand does not have any known target. The relative basepairing stability at the 50 ends of an siRNA duplex is a strong determinant of which strand will be incorporated into RISC and hence be active in RNAi; the strand whose 50 end has a weaker hydrogen bonding pattern is preferentially incorporated into RISC, the RNAi effector complex (Khvorova *et al.*, *Cell* **115**: 209–216, 2003; Schwarz *et al.*, *Cell* **115**: 208–299, 2003). This same principle can also be applied to the design of DNA vector-based siRNA expression strategies, including the one described here. However, for artificial miRNAs, the fact that the internal cleavage sites by Drosha and Dicer cannot be precisely predicted at present adds a degree of uncertainty as a 1- or 2-nt shift in the cleavage site can generate rather different hydrogen bonding patterns at the 50 ends of the resulting duplex, thus changing which strand of the duplex intermediate is incorporated into RISC. This is in contrast to the situation with synthetic siRNA duplexes, which have defined ends. On the other hand, any minor heterogeneity at the ends of an artificial miRNA duplex intermediate might not be a problem, as the miRNAs would still be perfectly complementary to their target.

The role of internal loop, stem length, and the surrounding sequences on the expression of miRNAs from miR-30-derived cassettes may also be systematically examined to optimize expression of the miR-based shRNA. Such analyses may suggest design elements that would maximize the yield of the intended RNA products. On the other hand, some heterogeneity could be inevitable. In addition to the 50-end rule, specific residues at some positions within an siRNA may also enhance siRNA function (Reynolds *et al.*, *Nat. Biotech.* **22**: 326–330, 2004).

In general, picking a target region with more than 50% AU content and designing a weak 50 end base pair on the antisense strand would be a good starting point in the design of any artificial miRNA/siRNA expression plasmid (Khvorova *et al.*, *Cell* **115**: 209–216, 2003; Reynolds *et al.*, *Nat. Biotech.* **22**: 326–330, 2004;

Schwarz *et al.*, *Cell* **115**: 208–299, 2003).

In certain embodiments, expression of the miR-30 cassette may be in the antisense orientation, especially when the cassette is to be used in lentiviral or retroviral vectors. This is partly because miRNA processing may result in the
5 degradation of the remainder of the primary miRNA transcript.

In other embodiments, vectors may contain inserts expressing more than one miRNAs. In such constructs, the fact that each miRNA stem-loop precursor is independently excised from the primary transcript by Drosha cleavage to give rise to a pre-miRNA allows simultaneous expression of several artificial or authentic
10 miRNAs by a tandem array on a precursor RNA transcript.

Genome wide libraries of shRNAs based on the miR30 precursor RNA have also been generated. Each member of such libraries target specific human or mouse genes, and may be readily converted to the vectors / expression systems of the instant invention. The following section describes the design of such libraries.

15 Paddison *et al.* (*Nature Methods* **1(2)**: 163-67, 2004, incorporated herein by reference) have described a genome-wise library of shRNAs based on the miR30 precursor RNA, which may be adapted for use in the instant invention. The described vector pSHAG-MAGIC2 (pSM2) is roughly equivalent to pSHAG-MAGIC1 as described in Paddison *et al. Methods Mol. Biol.* **265**: 85-100 (2004),
20 incorporated herein by reference. The few notable exceptions include: the new cloning strategy is based on the use of a single oligonucleotide that contains the hairpin and common 5' and 3' ends as a PCR template (see Fig. 2 of Paddison, *Nature Methods* **1(2)**: 163-67, 2004). The resulting PCR product is then cloned into the hairpin cloning site of the pSM2 vector, which drives miR-30-styled hairpins by
25 the human U6 promoter. Inserts from this library may be excised (see Example below) and cloned into the instant vectors for Pol II-driven expression of the same miR-30-styled hairpins. This allows the instant methods to be coupled with the existing library of miR-30-style constructs that contains most human and mouse genes.

30 Paddison also describes the detailed methods for designing 22-nucleotide sequences (targeting a target gene) that can be inserted into the precursor miRNA, PCR protocols for amplification, and relevant critical steps and trouble-shootings,

etc. (all incorporated herein by reference).

MicroRNAs (including the siRNA products and artificial microRNAs as well as endogenous microRNAs) have potential for use as therapeutics as well as research tools, *e.g.* analyzing gene function. As a general method, the mature
5 microRNA (miR) of the invention, especially those non-miR-30 based microRNA constructs of the invention may also be produced according to the following description.

In certain embodiments, the methods for efficient expression of microRNA involve the use of a precursor microRNA molecule having a microRNA sequence in
10 the context of microRNA flanking sequences. The precursor microRNA is composed of any type of nucleic acid based molecule capable of accommodating the microRNA flanking sequences and the microRNA sequence. Examples of precursor microRNAs and the individual components of the precursor (flanking sequences and microRNA sequence) are provided herein. The invention, however, is not limited to
15 the examples provided. The invention is based, at least in part, on the discovery of an important component of precursor microRNAs, that is, the microRNA flanking sequences. The nucleotide sequence of the precursor and its components may vary widely.

In one aspect a precursor microRNA molecule is an isolated nucleic acid
20 including microRNA flanking sequences and having a stem-loop structure with a microRNA sequence incorporated therein. An "isolated molecule" is a molecule that is free of other substances with which it is ordinarily found in nature or *in vivo* systems to an extent practical and appropriate for its intended use. In particular, the molecular species are sufficiently free from other biological constituents of host
25 cells or if they are expressed in host cells they are free of the form or context in which they are ordinarily found in nature. For instance, a nucleic acid encoding a precursor microRNA having homologous microRNA sequences and flanking sequences may ordinarily be found in a host cell in the context of the host cell genomic DNA. An isolated nucleic acid encoding a microRNA precursor may be
30 delivered to a host cell, but is not found in the same context of the host genomic DNA as the natural system. Alternatively, an isolated nucleic acid is removed from the host cell or present in a host cell that does not ordinarily have such a nucleic acid

sequence. Because an isolated molecular species of the invention may be admixed with a pharmaceutically-acceptable carrier in a pharmaceutical preparation or delivered to a host cell, the molecular species may comprise only a small percentage by weight of the preparation or cell. The molecular species is nonetheless isolated in that it has been substantially separated from the substances with which it may be associated in living systems.

An "isolated precursor microRNA molecule" is one which is produced from a vector having a nucleic acid encoding the precursor microRNA. Thus, the precursor microRNA produced from the vector may be in a host cell or removed from a host cell. The isolated precursor microRNA may be found within a host cell that is capable of expressing the same precursor. It is nonetheless isolated in that it is produced from a vector and, thus, is present in the cell in a greater amount than would ordinarily be expressed in such a cell.

The term "nucleic acid" is used to mean multiple nucleotides (*i.e.* molecules comprising a sugar (*e.g.* ribose or deoxyribose) linked to a phosphate group and to an exchangeable organic base, which is either a substituted pyrimidine (*e.g.* cytosine (C), thymidine (T) or uracil (U)) or a substituted purine (*e.g.* adenine (A) or guanine (G)). The term shall also include polynucleosides (*i.e.* a polynucleotide minus the phosphate) and any other organic base containing polymer. Purines and pyrimidines include but are not limited to adenine, cytosine, guanine, thymidine, inosine, 5-methylcytosine, 2-aminopurine, 2-amino-6-chloropurine, 2,6-diaminopurine, hypoxanthine, and other naturally and non-naturally occurring nucleobases, substituted and unsubstituted aromatic moieties. Other such modifications are well known to those of skill in the art. Thus, the term nucleic acid also encompasses nucleic acids with substitutions or modifications, such as in the bases and/or sugars.

"MicroRNA flanking sequence" as used herein refers to nucleotide sequences including microRNA processing elements. MicroRNA processing elements are the minimal nucleic acid sequences which contribute to the production of mature microRNA from precursor microRNA. Often these elements are located within a 40 nucleotide sequence that flanks a microRNA stem-loop structure. In some instances the microRNA processing elements are found within a stretch of nucleotide sequences of between 5 and 4,000 nucleotides in length that flank a

microRNA stem-loop structure.

Thus, in some embodiments the flanking sequences are 5-4,000 nucleotides in length. As a result, the length of the precursor molecule may be, in some instances at least about 150 nucleotides or 270 nucleotides in length. The total length of the precursor molecule, however, may be greater or less than these values. In other
5 embodiments the minimal length of the microRNA flanking sequence is 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200 and any integer there between. In other
embodiments the maximal length of the microRNA flanking sequence is 2,000,
2,100, 2,200, 2,300, 2,400, 2,500, 2,600, 2,700, 2,800, 2,900, 3,000, 3,100, 3,200,
10 3,300, 3,400, 3,500, 3,600, 3,700, 3,800, 3,900 4,000 and any integer there between.

The microRNA flanking sequences may be native microRNA flanking sequences or artificial microRNA flanking sequences. A native microRNA flanking sequence is a nucleotide sequence that is ordinarily associated in naturally existing systems with microRNA sequences, *i.e.*, these sequences are found within the
15 genomic sequences surrounding the minimal microRNA hairpin *in vivo*. Artificial microRNA flanking sequences are nucleotides sequences that are not found to be flanking to microRNA sequences in naturally existing systems. The artificial microRNA flanking sequences may be flanking sequences found naturally in the context of other microRNA sequences. Alternatively they may be composed of
20 minimal microRNA processing elements which are found within naturally occurring flanking sequences and inserted into other random nucleic acid sequences that do not naturally occur as flanking sequences or only partially occur as natural flanking sequences.

The microRNA flanking sequences within the precursor microRNA
25 molecule may flank one or both sides of the stem-loop structure encompassing the microRNA sequence. Thus, one end (*i.e.*, 5') of the stem-loop structure may be adjacent to a single flanking sequence and the other end (*i.e.*, 3') of the stem-loop structure may not be adjacent to a flanking sequence. Preferred structures have flanking sequences on both ends of the stem-loop structure. The flanking sequences
30 may be directly adjacent to one or both ends of the stem-loop structure or may be connected to the stem-loop structure through a linker, additional nucleotides or other molecules.

A “stem-loop structure” refers to a nucleic acid having a secondary structure that includes a region of nucleotides which are known or predicted to form a double strand (stem portion) that is linked on one side by a region of predominantly single-stranded nucleotides (loop portion). The terms “hairpin” and “fold-back” structures are also used herein to refer to stem-loop structures. Such structures are well known in the art and the term is used consistently with its known meaning in the art. The actual primary sequence of nucleotides within the stem-loop structure is not critical to the practice of the invention as long as the secondary structure is present. As is known in the art, the secondary structure does not require exact base-pairing. Thus, the stem may include one or more base mismatches. Alternatively, the base-pairing may be exact, *i.e.* not include any mismatches.

In some instances the precursor microRNA molecule may include more than one stem-loop structure. The multiple stem-loop structures may be linked to one another through a linker, such as, for example, a nucleic acid linker or by a microRNA flanking sequence or other molecule or some combination thereof.

In an alternative embodiment, useful interfering RNAs can be designed with a number of software programs, *e.g.*, the OligoEngine siRNA design tool available at www.oligoengine.com. The siRNAs of this invention may range about, *e.g.*, 19-29 basepairs in length for the double-stranded portion. In some embodiments, the siRNAs are hairpin RNAs having an about 19-29 bp stem and an about 4-34 nucleotide loop. Preferred siRNAs are highly specific for a region of the target gene and may comprise any about 19-29 bp fragment of a target gene mRNA that has at least one, preferably at least two or three, bp mismatch with a nontarget gene-related sequence. In some embodiments, the preferred siRNAs do not bind to RNAs having more than 3 mismatches with the target region.

III. *Expression Vectors and Host Cells*

The invention also includes vectors for producing precursor microRNA molecules. Generally these vectors include a sequence encoding a precursor microRNA and (*in vivo*) expression elements. The expression elements include at least one promoter, such as a Pol II promoter, which may direct the expression of the operably linked microRNA precursor (*e.g.* the shRNA encoding sequence). The

vector or primary transcript is first processed to produce the stem-loop precursor molecule. The stem-loop precursor is then processed to produce the mature microRNA.

RNA polymerase III (Pol III) transcription units normally encode the small nuclear RNA U6 (see Tran *et al.*, *BMC Biotechnology* **3**: 21, 2003, incorporate herein by reference), or the human RNase P RNA Hi. However, RNA polymerase II (Pol II) transcription units (*e.g.*, units containing a CMV promoter) is preferred for use with inducible expression. It will be appreciated that in the vectors of the invention, the subject shRNA encoding sequence may be operably linked to a variety of other promoters.

In some embodiments, the promoter is a type II tRNA promoter such as the tRNA^{Va} promoter and the tRNA^{met} promoter. These promoters may also be modified to increase promoter activity. In addition, enhancers can be placed near the promoter to enhance promoter activity. Pol II enhancer may also be used for Pol III promoters. For example, an enhancer from the CMV promoter can be placed near the U6 promoter to enhance U6 promoter activity (Xia *et al.*, *Nuc Acids Res* **31**, 2003).

In certain embodiments, the subject Pol II promoters are inducible promoters. Exemplary inducible Pol II systems are available from Invitrogen, *e.g.*, the GeneSwitchTM or T-RExTM systems; from Clontech (Palo Alto, CA), *e.g.*, the TetON and TetOFF systems.

An exemplary Tet-responsive promoter is described in WO 04/056964A2 (incorporated herein by reference). See, for example, Figure 1 of WO 04/056964A2. In one construct, a Tet operator sequence (TetOp) is inserted into the promoter region of the vector. TetOp is preferably inserted between the PSE and the transcription initiation site, upstream or downstream from the TATA box. In some embodiments, the TetOp is immediately adjacent to the TATA box. The expression of the subject shRNA encoding sequence is thus under the control of tetracycline (or its derivative doxycycline, or any other tetracycline analogue). Addition of tetracycline or Dox relieves repression of the promoter by a tetracycline repressor that the host cells are also engineered to express.

In the TetOFF system, a different tet transactivator protein is expressed in the

tetOFF host cell. The difference is that Tet / Dox, when bind to an activator protein, is now required for transcriptional activation. Thus such host cells expressing the activator will only activate the transcription of an shRNA encoding sequence from a TetOFF promoter at the presence of Tet or Dox.

5 An alternative inducible promoter is a lac operator system, as illustrated in Figure 2A of WO 04/056964 A2 (incorporated by reference). Briefly, a Lac operator sequence (LacO) is inserted into the promoter region. The LacO is preferably inserted between the PSE and the transcription initiation site, upstream or downstream of the TATA box. In some embodiments, the LacO is immediately
10 adjacent to the TATA box. The expression of the RNAi molecule (shRNA encoding sequence) is thus under the control of IPTG (or any analogue thereof). Addition of IPTG relieves repression of the promoter by a Lac repressor (*i.e.*, the LacI protein) that the host cells are also engineered to express. Since the Lac repressor is derived from bacteria, its coding sequence may be optionally modified to adapt to the codon
15 usage by mammalian transcriptional systems and to prevent methylation. In some embodiments, the host cells comprise (i) a first expression construct containing a gene encoding a Lac repressor operably linked to a first promoter, such as any tissue or cell type specific promoter or any general promoter, and (ii) a second expression construct containing the dsRNA-coding sequence operably linked to a second
20 promoter that is regulated by the Lac repressor and IPTG. Administration of IPTG results in expression of dsRNA in a manner dictated by the tissue specificity of the first promoter.

Yet another inducible system, a LoxP-stop-LoxP system, is illustrated in Figures 3A-3E of WO 04/056964 A2 (incorporated by reference). The RNAi vector
25 of this system contains a LoxP-Stop-LoxP cassette before the hairpin or within the loop of the hairpin. Any suitable stop sequence for the promoter can be used in the cassette. One version of the LoxP Stop-LoxP system for Pol II is described in, *e.g.*, Wagner *et al.*, *Nucleic Acids Research* **25**:4323-4330, 1997. The "Stop" sequences (such as the one described in Wagner, sierra, or a run of five or more T nucleotides)
30 in the cassette prevent the RNA polymerase III from extending an RNA transcript beyond the cassette. Upon introduction of a Cre recombinase, however, the LoxP sites in the cassette recombine, removing the Stop sequences and leaving a single

LoxP site. Removal of the Stop sequences allows transcription to proceed through the hairpin sequence, producing a transcript that can be efficiently processed into an open-ended, interfering dsRNA. Thus, expression of the RNAi molecule is induced by addition of Cre.

5 In some embodiments, the host cells contain a Cre-encoding transgene under the control of a constitutive, tissue-specific promoter. As a result, the interfering RNA can only be inducibly expressed in a tissue-specific manner dictated by that promoter. Tissue-specific promoters that can be used include, without limitation: a tyrosinase promoter or a TRP2 promoter in the case of melanoma cells and
10 melanocytes; an MMTV or WAP promoter in the case of breast cells and/or cancers; a Villin or FABP promoter in the case of intestinal cells and/or cancers; a RIP promoter in the case of pancreatic beta cells; a Keratin promoter in the case of keratinocytes; a Probasin promoter in the case of prostatic epithelium; a Nestin or GFAP promoter in the case of CNS cells and/or cancers; a Tyrosine Hydroxylase,
15 S100 promoter or neurofilament promoter in the case of neurons; the pancreas-specific promoter described in Edlund *et al.*, *Science* **230**: 912-916, 1985; a Clara cell secretory protein promoter in the case of lung cancer; and an Alpha myosin promoter in the case of cardiac cells.

 Cre expression also can be controlled in a temporal manner, *e.g.*, by using an
20 inducible promoter, or a promoter that is temporally restricted during development such as Pax3 or Protein O (neural crest), Hoxal (floorplate and notochord), Hoxb6 (extraembryonic mesoderm, lateral plate and limb mesoderm and midbrain-hindbrain junction), Nestin (neuronal lineage), GFAP (astrocyte lineage), Lck (immature thymocytes). Temporal control also can be achieved by using an
25 inducible form of Cre. For example, one can use a small molecule controllable Cre fusion, for example a fusion of the Cre protein and the estrogen receptor (ER) or with the progesterone receptor (PR). Tamoxifen or RU486 allow the Cre-ER or Cre-PR fusion, respectively, to enter the nucleus and recombine the LoxP sites, removing the LoxP Stop cassette. Mutated versions of either receptor may also be used. For
30 example, a mutant Cre-PR fusion protein may bind RU486 but not progesterone. Other exemplary Cre fusions are a fusion of the Cre protein and the glucocorticoid receptor (GR). Natural GR ligands include corticosterone, cortisol, and aldosterone.

Mutant versions of the GR receptor, which respond to, *e.g.*, dexamethasone, triamcinolone acetonide, and/or RU38486, may also be fused to the Cre protein.

In certain embodiments, additional transcription units may be present 3' to the shRNA portion. For example, an internal ribosomal entry site (IRES) may be positioned downstream of the shRNA insert, the transcription of which is under the control of a second promoter, such as the PGK promoter. The IRES sequence may be used to direct the expression of a operably linked second gene, such as a reporter gene (*e.g.*, a fluorescent protein such as GFP, BFP, YFP, *etc.*, an enzyme such as luciferase (Promega), *etc.*). The reporter gene may serve as an indication of infection / transfection, and the efficiency and/or amount of mRNA transcription of the shRNA – IRES – reporter cassette / insert. Optionally, one or more selectable markers (such as puromycin resistance gene, neomycin resistance gene, hygromycin resistance gene, zeocin resistance gene, *etc.*) may also be present on the same vector, and are under the transcriptional control of the second promoter. Such markers may be useful for selecting stable integration of the vector into a host cell genome.

Certain exemplary vectors useful for expressing the precursor microRNAs are shown in the examples. Thus the invention encompasses the nucleotide sequence of such vectors as well as variants thereof.

In general, variants typically will share at least 40% nucleotide identity with any of the described vectors, in some instances, will share at least 50% nucleotide identity; and in still other instances, will share at least 60% nucleotide identity. The preferred variants have at least 70% sequence homology. More preferably the preferred variants have at least 80% and, most preferably, at least 90% sequence homology to the described sequences.

Variants with high percentage sequence homology can be identified, for example, using stringent hybridization conditions. The term “stringent conditions”, as used herein, refers to parameters with which the art is familiar. More specifically, stringent conditions, as used herein, refer to hybridization at 65°C in hybridization buffer (3.5 × SSC, 0.02% Ficoll, 0.02% polyvinyl pyrrolidone, 0.02% bovine serum albumin, 2.5 mM NaH₂PO₄ (pH 7), 0.5% SDS, 2 mM EDTA). SSC is 0.15M sodium chloride/0.15M sodium citrate, pH 7; SDS is sodium dodecyl sulphate; and EDTA is ethylenediaminetetraacetic acid. After hybridization, the membrane to

which the DNA is transferred is washed at $2 \times$ SSC at room temperature and then at $0.1 \times$ SSC / $0.1 \times$ SDS at 65°C . There are other conditions, reagents, and so forth which can be used, which result in a similar degree of stringency. Such variants may be further subject to functional testing such that variants that substantially preserve
5 the desired / relevant function of the original vectors are selected / identified.

The "*in vivo* expression elements" are any regulatory nucleotide sequence, such as a promoter sequence or promoter-enhancer combination, which facilitates the efficient expression of the nucleic acid to produce the precursor microRNA. The
10 *in vivo* expression element may, for example, be a mammalian or viral promoter, such as a constitutive or inducible promoter or a tissue specific promoter. Constitutive mammalian promoters include, but are not limited to, polymerase II promoters as well as the promoters for the following genes: hypoxanthine phosphoribosyl transferase (HPTR), adenosine deaminase, pyruvate kinase, and β -actin. Exemplary viral promoters which function constitutively in eukaryotic cells
15 include, for example, promoters from the simian virus, papilloma virus, adenovirus, human immunodeficiency virus (HIV), Rous sarcoma virus, cytomegalovirus, the long terminal repeats (LTR) of moloney leukemia virus and other retroviruses, and the thymidine kinase promoter of herpes simplex virus. Other constitutive promoters are known to those of ordinary skill in the art. The promoters useful as *in vivo*
20 expression element of the invention also include inducible promoters. Inducible promoters are expressed in the presence of an inducing agent. For example, the metallothionein promoter is induced to promote transcription in the presence of certain metal ions. Other inducible promoters are known to those of ordinary skill in the art.

25 One useful inducible expression system that can be adapted for use in the instant invention is the Tet-responsive system, including both the TetON and TetOFF embodiments.

TetOn system is a commercially available inducible expression system from Clontech Inc. This is of particular interest because current siRNA expression
30 systems utilize pol III promoters, which are difficult to adapt for inducible expression. The Clontech TetON system includes the pRev-TRE vector, which can be packaged into retrovirus and used to infect a Tet-On cell line expressing the

reverse tetracycline-controlled transactivator (rtTA). Once introduced into the TetON host cell, the shRNA insert can then be inducibly expressed in response to varying concentrations of the tetracycline derivate doxycycline (Dox).

5 In general, the *in vivo* expression element shall include, as necessary, 5' non-transcribing and 5' non-translating sequences involved with the initiation of transcription. They optionally include enhancer sequences or upstream activator sequences as desired.

Vectors include, but are not limited to, plasmids, phagemids, viruses, other vehicles derived from viral or bacterial sources that have been manipulated by the
10 insertion or incorporation of the nucleic acid sequences for producing the precursor microRNA, and free nucleic acid fragments which can be attached to these nucleic acid sequences. Viral and retroviral vectors are a preferred type of vector and include, but are not limited to, nucleic acid sequences from the following viruses: retroviruses, such as: Moloney murine leukemia virus; Murine stem cell virus,
15 Harvey murine sarcoma virus; murine mammary tumor virus; Rous sarcoma virus; adenovirus; adeno-associated virus; SV40-type viruses; polyoma viruses; Epstein-Barr viruses; papilloma viruses; herpes viruses; vaccinia viruses; polio viruses; lentiviruses; and RNA viruses such as any retrovirus. One can readily employ other unnamed vectors known in the art.

20 Viral vectors are generally based on non-cytopathic eukaryotic viruses in which non-essential genes have been replaced with the nucleic acid sequence of interest. Non-cytopathic viruses include retroviruses, the life cycle of which involves reverse transcription of genomic viral RNA into DNA with subsequent proviral integration into host cellular DNA. Retroviruses have been approved for human
25 gene therapy trials. Genetically altered retroviral expression vectors have general utility for the high-efficiency transduction of nucleic acids *in vivo*. Standard protocols for producing replication-deficient retroviruses (including the steps of incorporation of exogenous genetic material into a plasmid, transfection of a packaging cell lined with plasmid, production of recombinant retroviruses by the
30 packaging cell line, collection of viral particles from tissue culture media, and infection of the target cells with viral particles) are provided in Kriegler, M., "Gene Transfer and Expression, A Laboratory Manual," W.H. Freeman Co., New York

(1990) and Murry, E. J. Ed. "Methods in Molecular Biology," vol. 7, Humana Press, Inc., Clifton, N.J. (1991).

Exemplary vectors are disclosed herein and in US 2005/0075492 A2 (incorporated herein by reference) and WO 04/056964 A2 (incorporated herein by reference).

The invention also encompasses host cells transfected with the subject vectors, especially host cell lines with stably integrated shRNA constructs. In certain embodiments, the subject host cell contains only a single copy of the integrated construct expressing the desired shRNA (optionally under the control of an inducible and/or tissue specific promoter). Host cells include for instance, cells (such as primary cells) and cell lines, e.g. prokaryotic (e.g., *E. coli*), and eukaryotic (e.g., dendritic cells, CHO cells, COS cells, yeast expression systems and recombinant baculovirus expression in insect cells, etc.). Exemplary cells include: NIH3T3 cells, MEFs, 293 or 293T cells, CHO cells, hematopoietic stem/progenitor cells, cancer cells, etc.

IV. *Methods of Using*

In certain aspects, methods of the invention comprise contacting and introducing into a target cell with a subject vector capable of expressing a precursor microRNA as described herein, to regulate the expression of a target gene in the cell. The vector produces the microRNA transcript, which is then processed into precursor microRNA in the cell, which is then processed to produce the mature functional microRNA which is capable of altering accumulation of a target protein in the target cell. Accumulation of the protein may be effected in a number of different ways. For instance the microRNA may directly or indirectly affect translation or may result in cleavage of the mRNA transcript or even effect stability of the protein being translated from the target mRNA. MicroRNA may function through a number of different mechanisms. The methods and products of the invention are not limited to any one mechanism. The method may be performed *in vitro*, e.g., for studying gene function, *ex vivo* or *in vivo*, e.g. for therapeutic purposes.

An "ex vivo" method as used herein is a method which involves isolation of a

cell from a subject, manipulation of the cell outside of the body, and reimplantation of the manipulated cell into the subject. The *ex vivo* procedure may be used on autologous or heterologous cells, but is preferably used on autologous cells. In preferred embodiments, the *ex vivo* method is performed on cells that are isolated
5 from bodily fluids such as peripheral blood or bone marrow, but may be isolated from any source of cells. When returned to the subject, the manipulated cell will be programmed for cell death or division, depending on the treatment to which it was exposed. *Ex vivo* manipulation of cells has been described in several references in the art, including Engleman, E. G., 1997, *Cytotechnology*, 25:1; Van Schooten, W.,
10 *et al.*, 1997, *Molecular Medicine Today*, June, 255; Steinman, R. M., 1996, *Experimental Hematology*, 24, 849; and Gluckman, J. C., 1997, *Cytokines, Cellular and Molecular Therapy*, 3:187. The *ex vivo* activation of cells of the invention may be performed by routine *ex vivo* manipulation steps known in the art. *In vivo* methods are also well known in the art. The invention thus is useful for therapeutic
15 purposes and also is useful for research purposes such as testing in animal or *in vitro* models of medical, physiological or metabolic pathways or conditions.

The *ex vivo* and *in vivo* methods are performed on a subject. A "subject" shall mean a human or non-human mammal, including but not limited to, a dog, cat, horse, cow, pig, sheep, goat, primate, rat, and mouse, *etc.*

20 In some instances the mature microRNA is expressed at a level sufficient to cause at least a 2-fold, or in some instances, a 10 fold reduction in accumulation of the target protein. The level of accumulation of a target protein may be assessed using routine methods known to those of skill in the art. For instance, protein may be isolated from a target cell and quantitated using Western blot analysis or other
25 comparable methodologies, optionally in comparison to a control. Protein levels may also be assessed using reporter systems or fluorescently labeled antibodies. In other embodiments, the mature microRNA is expressed at a level sufficient to cause at least a 2, 5, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, or 100 fold reduction in accumulation of the target protein. The "fold reduction" may be assessed using
30 any parameter for assessing a quantitative value of protein expression. For instance, a quantitative value can be determined using a label *i.e.* fluorescent, radioactive linked to an antibody. The value is a relative value that is compared to a control or a

known value.

Different microRNA sequences have different levels of expression of mature microRNA and thus have different effects on target mRNA and/or protein expression. For instance, in some cases a microRNA may be expressed at a high
5 level and may be very efficient such that the accumulation of the target protein is completely or near completely blocked. In other instances the accumulation of the target protein may be only reduced slightly over the level that would ordinarily be expressed in that cell at that time under those conditions in the absence of the mature
10 microRNA. Complete inhibition of the accumulation of the target protein is not essential, for example, for therapeutic purposes. In many cases partial or low inhibition of accumulation may produce a preferred phenotype. The actual amount that is useful will depend on the particular cell type, the stage of differentiation, conditions to which the cell is exposed, the modulation of other target proteins, *etc.*

The microRNAs may be used to knock down gene expression in vertebrate
15 cells for gene-function studies, including target-validation studies during the development of new pharmaceuticals, as well as the development of human disease models and therapies, and ultimately, human gene therapies.

The methods of the invention are useful for treating any type of “disease”, “disorder” or “condition” in which it is desirable to reduce the expression or
20 accumulation of a particular target protein(s). Diseases include, for instance, but are not limited to, cancer, infectious disease, cystic fibrosis, blood disorders, including leukemia and lymphoma, spinal muscular dystrophy, early-onset Parkinsonism (Waisman syndrome) and X-linked mental retardation (MRX3).

Cancers include but are not limited to biliary tract cancer; bladder cancer;
25 breast cancer; brain cancer including glioblastomas and medulloblastomas; cervical cancer; choriocarcinoma; colon cancer including colorectal carcinomas; endometrial cancer; esophageal cancer; gastric cancer; head and neck cancer; hematological neoplasms including acute lymphocytic and myelogenous leukemia, multiple myeloma, AIDS-associated leukemias and adult T-cell leukemia lymphoma;
30 intraepithelial neoplasms including Bowen's disease and Paget's disease; liver cancer; lung cancer including small cell lung cancer and non-small cell lung cancer; lymphomas including Hodgkin's disease and lymphocytic lymphomas;

neuroblastomas; oral cancer including squamous cell carcinoma; osteosarcomas; ovarian cancer including those arising from epithelial cells, stromal cells, germ cells and mesenchymal cells; pancreatic cancer; prostate cancer; rectal cancer; sarcomas including leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma, 5 synovial sarcoma and osteosarcoma; skin cancer including melanomas, Kaposi's sarcoma, basocellular cancer, and squamous cell cancer; testicular cancer including germinal tumors such as seminoma, non-seminoma (teratomas, choriocarcinomas), stromal tumors, and germ cell tumors; thyroid cancer including thyroid adenocarcinoma and medullar carcinoma; transitional cancer and renal cancer 10 including adenocarcinoma and Wilms tumor.

An infectious disease, as used herein, is a disease arising from the presence of a foreign microorganism in the body. A microbial antigen, as used herein, is an antigen of a microorganism. Microorganisms include but are not limited to, infectious virus, infectious bacteria, and infectious fungi.

15 Examples of infectious virus include but are not limited to: *Retroviridae* (*e.g.* human immunodeficiency viruses, such as HIV-1 (also referred to as HTLV-III, LAV or HTLV-III/LAV, or HIV-III; and other isolates, such as HIV-LP; *Picornaviridae* (*e.g.* polio viruses, hepatitis A virus; enteroviruses, human Coxsackie viruses, rhinoviruses, echoviruses); *Calciviridae* (*e.g.* strains that cause 20 gastroenteritis); *Togaviridae* (*e.g.* equine encephalitis viruses, rubella viruses); *Flaviridae* (*e.g.* dengue viruses, encephalitis viruses, yellow fever viruses); *Coronaviridae* (*e.g.* coronaviruses); *Rhabdoviridae* (*e.g.* vesicular stomatitis viruses, rabies viruses); *Coronaviridae* (*e.g.* coronaviruses); *Rhabdoviridae* (*e.g.* vesicular stomatitis viruses, rabies viruses); *Filoviridae* (*e.g.* ebola viruses); 25 *Paramyxoviridae* (*e.g.* parainfluenza viruses, mumps virus, measles virus, respiratory syncytial virus); *Orthomyxoviridae* (*e.g.* influenza viruses); *Bungaviridae* (*e.g.* Hantaan viruses, bunga viruses, phleboviruses and Nairo viruses); *Arena viridae* (hemorrhagic fever viruses); *Reoviridae* (*e.g.* reoviruses, orbiviruses and rotaviruses); *Birnaviridae*; *Hepadnaviridae* (Hepatitis B virus); *Parvoviridae* 30 (parvoviruses); *Papovaviridae* (papilloma viruses, polyoma viruses); *Adenoviridae* (most adenoviruses); *Herpesviridae* (herpes simplex virus (HSV) 1 and 2, varicella zoster virus, cytomegalovirus (CMV), herpes virus); *Poxviridae* (variola viruses,

vaccinia viruses, pox viruses); and *Iridoviridae* (e.g. African swine fever virus); and unclassified viruses (e.g. the etiological agents of Spongiform encephalopathies, the agent of delta hepatitis (thought to be a defective satellite of hepatitis B virus), the agents of non-A, non-B hepatitis (class 1=internally transmitted; class
 5 2=parenterally transmitted (i.e. Hepatitis C); Norwalk and related viruses, and astroviruses).

Examples of infectious bacteria include but are not limited to: *Helicobacter pylori*, *Borelia burgdorferi*, *Legionella pneumophila*, *Mycobacteria sps* (e.g. *M. tuberculosis*, *M. avium*, *M. intracellulare*, *M. kansasii*, *M. goodii*), *Staphylococcus aureus*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Listeria monocytogenes*,
 10 *Streptococcus pyogenes* (Group A *Streptococcus*), *Streptococcus agalactiae* (Group B *Streptococcus*), *Streptococcus* (viridans group), *Streptococcus faecalis*, *Streptococcus bovis*, *Streptococcus* (anaerobic sps.), *Streptococcus pneumoniae*, pathogenic *Campylobacter sp.*, *Enterococcus sp.*, *Haemophilus influenzae*, *Bacillus antracis*, *Corynebacterium diphtheriae*, *Corynebacterium sp.*, *Erysipelothrix rhusiopathiae*, *Clostridium perfringens*, *Clostridium tetani*, *Enterobacter aerogenes*,
 15 *Klebsiella pneumoniae*, *Pasturella multocida*, *Bacteroides sp.*, *Fusobacterium nucleatum*, *Streptobacillus moniliformis*, *Treponema pallidum*, *Treponema pertenuis*, *Leptospira*, *Rickettsia*, and *Actinomyces israelii*.

Examples of infectious fungi include: *Cryptococcus neoformans*,
 20 *Histoplasma capsulatum*, *Coccidioides immitis*, *Blastomyces dermatitidis*, *Chlamydia trachomatis*, *Candida albicans*. Other infectious organisms (i.e., protists) include: *Plasmodium* such as *Plasmodium falciparum*, *Plasmodium malariae*, *Plasmodium ovale*, and *Plasmodium vivax* and *Toxoplasma gondii*.

The vectors of this invention can be delivered into host cells via a variety of
 25 methods, including but not limited to, liposome fusion (transposomes), infection by viral vectors, and routine nucleic acid transfection methods such as electroporation, calcium phosphate precipitation and microinjection. In some embodiments, the vectors are integrated into the genome of a transgenic animal (e.g., a mouse, a rabbit,
 30 a hamster, or a nonhuman primate). Diseased or disease-prone cells containing these vectors can be used as a model system to study the development, maintenance, or progression of a disease that is affected by the presence or absence of the interfering

RNA.

Expression of the miRNA/siRNA introduced into a target cell may be confirmed by art-recognized techniques, such as Northern blotting using a nucleic acid probe. For cell lines that are more difficult to transfect, more extracted RNA
5 can be used for analyses, optionally coupled with exposing the film longer. Once expression of the miRNA/siRNA is confirmed, the DNA construct can then be tested for RNAi efficacy against a cotransfected construct encoding the target protein or directly against an endogenous target. In the latter case, one preferably should have a clear idea of transfection efficiency and of the half-life of the target protein before
10 performing the experiment.

V. *Pharmaceutical Use and Methods of Administration*

In one aspect, the invention provides a method of administering any of the compositions described herein to a subject. When administered, the compositions
15 are applied in a therapeutically effective, pharmaceutically acceptable amount as a pharmaceutically acceptable formulation. As used herein, the term “pharmaceutically acceptable” is given its ordinary meaning. Pharmaceutically acceptable compounds are generally compatible with other materials of the formulation and are not generally deleterious to the subject. Any of the compositions
20 of the present invention may be administered to the subject in a therapeutically effective dose. A “therapeutically effective” or an “effective” as used herein means that amount necessary to delay the onset of, inhibit the progression of, halt altogether the onset or progression of, diagnose a particular condition being treated, or otherwise achieve a medically desirable result, *i.e.*, that amount which is capable
25 of at least partially preventing, reversing, reducing, decreasing, ameliorating, or otherwise suppressing the particular condition being treated. A therapeutically effective amount can be determined on an individual basis and will be based, at least in part, on consideration of the species of mammal, the mammal's age, sex, size, and health; the compound and/or composition used, the type of delivery system used; the
30 time of administration relative to the severity of the disease; and whether a single, multiple, or controlled-release dose regiment is employed. A therapeutically effective amount can be determined by one of ordinary skill in the art employing

such factors and using no more than routine experimentation.

The terms "treat," "treated," "treating," and the like, when used herein, refer to administration of the systems and methods of the invention to a subject, which may, for example, increase the resistance of the subject to development or further
5 development of cancers, to administration of the composition in order to eliminate or at least control a cancer or a infectious disease, and/or to reduce the severity of the cancer or infectious disease, or symptoms thereof. Such terms also include prevention of disease / condition in, for example, subjects / individuals predisposed to such diseases / conditions, or at high risk of developing such diseases / conditions.

10 When administered to a subject, effective amounts will depend on the particular condition being treated and the desired outcome. A therapeutically effective dose may be determined by those of ordinary skill in the art, for instance, employing factors such as those further described below and using no more than routine experimentation.

15 In administering the systems and methods of the invention to a subject, dosing amounts, dosing schedules, routes of administration, and the like may be selected so as to affect known activities of these systems and methods. Dosage may be adjusted appropriately to achieve desired drug levels, local or systemic, depending upon the mode of administration. The doses may be given in one or
20 several administrations per day. As one example, if daily doses are required, daily doses may be from about 0.01 mg/kg/day to about 1000 mg/kg/day, and in some embodiments, from about 0.1 to about 100 mg/kg/day or from about 1 mg/kg/day to about 10 mg/kg/day. Parental administration, in some cases, may be from one to several orders of magnitude lower dose per day, as compared to oral doses. For
25 example, the dosage of an active compound when parentally administered may be between about 0.1 micrograms/kg/day to about 10 mg/kg/day, and in some embodiments, from about 1 microgram/kg/day to about 1 mg/kg/day or from about 0.01 mg/kg/day to about 0.1 mg/kg/day.

In some embodiments, the concentration of the active compound(s), if
30 administered systemically, is at a dose of about 1.0 mg to about 2000 mg for an adult of 70 kg body weight, per day. In other embodiments, the dose is about 10 mg to about 1000 mg/70 kg/day. In yet other embodiments, the dose is about 100 mg to

about 500 mg/70 kg/day. Preferably, the concentration, if applied topically, is about 0.1 mg to about 500 mg/gm of ointment or other base, more preferably about 1.0 mg to about 100 mg/gm of base, and most preferably, about 30 mg to about 70 mg/gm of base. The specific concentration partially depends upon the particular
5 composition used, as some are more effective than others. The dosage concentration of the composition actually administered is dependent at least in part upon the particular physiological response being treated, the final concentration of composition that is desired at the site of action, the method of administration, the efficacy of the particular composition, the longevity of the particular composition,
10 and the timing of administration relative to the severity of the disease. Preferably, the dosage form is such that it does not substantially deleteriously affect the mammal. The dosage can be determined by one of ordinary skill in the art employing such factors and using no more than routine experimentation.

In the event that the response of a particular subject is insufficient at such
15 doses, even higher doses (or effectively higher doses by a different, more localized delivery route) may be employed to the extent that subject tolerance permits. Multiple doses per day are also contemplated in some cases to achieve appropriate systemic levels within the subject or within the active site of the subject. In some cases, dosing amounts, dosing schedules, routes of administration, and the like may
20 be selected as described herein, whereby therapeutically effective levels for the treatment of cancer are provided.

In certain embodiments where cancers are being treated, a composition of the invention may be administered to a subject who has a family history of cancer, or to a subject who has a genetic predisposition for cancer. In other embodiments, the
25 composition is administered to a subject who has reached a particular age, or to a subject more likely to get cancer. In yet other embodiments, the compositions is administered to subjects who exhibit symptoms of cancer (*e.g.*, early or advanced). In still other embodiments, the composition may be administered to a subject as a preventive measure. In some embodiments, the inventive composition may be
30 administered to a subject based on demographics or epidemiological studies, or to a subject in a particular field or career.

Administration of a composition of the invention to a subject may be

accomplished by any medically acceptable method which allows the composition to reach its target. The particular mode selected will depend of course, upon factors such as those previously described, for example, the particular composition, the severity of the state of the subject being treated, the dosage required for therapeutic efficacy, *etc.* As used herein, a “medically acceptable” mode of treatment is a mode
5 able to produce effective levels of the active compound(s) of the composition within the subject without causing clinically unacceptable adverse effects.

Any medically acceptable method may be used to administer a composition to the subject. The administration may be localized (*i.e.*, to a particular region,
10 physiological system, tissue, organ, or cell type) or systemic, depending on the condition being treated. For example, the composition may be administered orally, vaginally, rectally, buccally, pulmonary, topically, nasally, transdermally, through parenteral injection or implantation, via surgical administration, or any other method of administration where suitable access to a target is achieved. Examples of
15 parenteral modalities that can be used with the invention include intravenous, intradermal, subcutaneous, intracavity, intramuscular, intraperitoneal, epidural, or intrathecal. Examples of implantation modalities include any implantable or injectable drug delivery system. Oral administration may be preferred in some embodiments because of the convenience to the subject as well as the dosing
20 schedule. Compositions suitable for oral administration may be presented as discrete units such as hard or soft capsules, pills, cachettes, tablets, troches, or lozenges, each containing a predetermined amount of the active compound. Other oral compositions suitable for use with the invention include solutions or suspensions in aqueous or non-aqueous liquids such as a syrup, an elixir, or an emulsion. In another set of
25 embodiments, the composition may be used to fortify a food or a beverage.

Injections can be *e.g.*, intravenous, intradermal, subcutaneous, intramuscular, or interperitoneal. The composition can be injected interdermally for treatment or prevention of infectious disease, for example. In some embodiments, the injections can be given at multiple locations. Implantation includes inserting implantable drug
30 delivery systems, *e.g.*, microspheres, hydrogels, polymeric reservoirs, cholesterol matrixes, polymeric systems, *e.g.*, matrix erosion and/or diffusion systems and non-polymeric systems, *e.g.*, compressed, fused, or partially-fused pellets. Inhalation

includes administering the composition with an aerosol in an inhaler, either alone or attached to a carrier that can be absorbed. For systemic administration, it may be preferred that the composition is encapsulated in liposomes.

In general, the compositions of the invention may be delivered using a
5 bioerodible implant by way of diffusion, or more preferably, by degradation of the polymeric matrix. Exemplary synthetic polymers which can be used to form the biodegradable delivery system include: polyamides, polycarbonates, polyalkylenes, polyalkylene glycols, polyalkylene oxides, polyalkylene terephthalates, polyvinyl alcohols, polyvinyl ethers, polyvinyl esters, poly-vinyl halides,
10 polyvinylpyrrolidone, polyglycolides, polysiloxanes, polyurethanes and co-polymers thereof, alkyl cellulose, hydroxyalkyl celluloses, cellulose ethers, cellulose esters, nitro celluloses, polymers of acrylic and methacrylic esters, methyl cellulose, ethyl cellulose, hydroxypropyl cellulose, hydroxy-propyl methyl cellulose, hydroxybutyl methyl cellulose, cellulose acetate, cellulose propionate, cellulose acetate butyrate,
15 cellulose acetate phthalate, carboxylethyl cellulose, cellulose triacetate, cellulose sulphate sodium salt, poly(methyl methacrylate), poly(ethyl methacrylate), poly(butylmethacrylate), poly(isobutyl methacrylate), poly(hexylmethacrylate), poly(isodecyl methacrylate), poly(lauryl methacrylate), poly(phenyl methacrylate), poly(methyl acrylate), poly(isopropyl acrylate), poly(isobutyl acrylate),
20 poly(octadecyl acrylate), polyethylene, polypropylene, poly(ethylene glycol), poly(ethylene oxide), poly(ethylene terephthalate), poly(vinyl alcohols), polyvinyl acetate, poly vinyl chloride, polystyrene, polyvinylpyrrolidone, and polymers of lactic acid and glycolic acid, polyanhydrides, poly(ortho)esters, poly(butic acid), poly(valeric acid), and poly(lactide-cocaprolactone), and natural polymers such as
25 alginate and other polysaccharides including dextran and cellulose, collagen, chemical derivatives thereof (substitutions, additions of chemical groups, for example, alkyl, alkylene, hydroxylations, oxidations, and other modifications routinely made by those skilled in the art), albumin and other hydrophilic proteins, zein and other prolamines and hydrophobic proteins, copolymers and mixtures
30 thereof. In general, these materials degrade either by enzymatic hydrolysis or exposure to water *in vivo*, by surface or bulk erosion. Examples of non-biodegradable polymers include ethylene vinyl acetate, poly(meth)acrylic acid,

polyamides, copolymers and mixtures thereof.

Bioadhesive polymers of particular interest include bioerodible hydrogels described by H. S. Sawhney, C. P. Pathak and J. A. Hubell in *Macromolecules*, (1993) 26:581-587, the teachings of which are incorporated herein, polyhyaluronic
5 acids, casein, gelatin, glutin, polyanhydrides, polyacrylic acid, alginate, chitosan, poly(methyl methacrylates), poly(ethyl methacrylates), poly(butylmethacrylate), poly(isobutyl methacrylate), poly(hexylmethacrylate), poly(isodecyl methacrylate), poly(lauryl methacrylate), poly(phenyl methacrylate), poly(methyl acrylate), poly(isopropyl acrylate), poly(isobutyl acrylate), and poly(octadecyl acrylate).

10 In certain embodiments of the invention, the administration of the composition of the invention may be designed so as to result in sequential exposures to the composition over a certain time period, for example, hours, days, weeks, months or years. This may be accomplished, for example, by repeated
15 administrations of a composition of the invention by one of the methods described above, or by a sustained or controlled release delivery system in which the composition is delivered over a prolonged period without repeated administrations. Administration of the composition using such a delivery system may be, for example, by oral dosage forms, bolus injections, transdermal patches or subcutaneous implants. Maintaining a substantially constant concentration of the
20 composition may be preferred in some cases.

Other delivery systems suitable for use with the present invention include time-release, delayed release, sustained release, or controlled release delivery systems. Such systems may avoid repeated administrations in many cases, increasing convenience to the subject and the physician. Many types of release
25 delivery systems are available and known to those of ordinary skill in the art. They include, for example, polymer-based systems such as polylactic and/or polyglycolic acids, polyanhydrides, polycaprolactones, copolyoxalates, polyesteramides, polyorthoesters, polyhydroxybutyric acid, and/or combinations of these. Microcapsules of the foregoing polymers containing drugs are described in, for
30 example, U.S. Pat. No. 5,075,109. Other examples include nonpolymer systems that are lipid-based including sterols such as cholesterol, cholesterol esters, and fatty acids or neutral fats such as mono-, di- and triglycerides; hydrogel release systems;

liposome-based systems; phospholipid based-systems; silastic systems; peptide based systems; wax coatings; compressed tablets using conventional binders and excipients; or partially fused implants. Specific examples include, but are not limited to, erosional systems in which the composition is contained in a form within a matrix (for example, as described in U.S. Pat. Nos. 4,452,775, 4,675,189, 5,736,152, 4,667,013, 4,748,034 and 5,239,660), or diffusional systems in which an active component controls the release rate (for example, as described in U.S. Pat. Nos. 3,832,253, 3,854,480, 5,133,974 and 5,407,686). The formulation may be as, for example, microspheres, hydrogels, polymeric reservoirs, cholesterol matrices, or polymeric systems. In some embodiments, the system may allow sustained or controlled release of the composition to occur, for example, through control of the diffusion or erosion/degradation rate of the formulation containing the composition. In addition, a pump-based hardware delivery system may be used to deliver one or more embodiments of the invention.

Examples of systems in which release occurs in bursts includes, *e.g.*, systems in which the composition is entrapped in liposomes which are encapsulated in a polymer matrix, the liposomes being sensitive to specific stimuli, *e.g.*, temperature, pH, light or a degrading enzyme and systems in which the composition is encapsulated by an ionically-coated microcapsule with a microcapsule core degrading enzyme. Examples of systems in which release of the inhibitor is gradual and continuous include, *e.g.*, erosional systems in which the composition is contained in a form within a matrix and effusional systems in which the composition permeates at a controlled rate, *e.g.*, through a polymer. Such sustained release systems can be *e.g.*, in the form of pellets, or capsules.

Use of a long-term release implant may be particularly suitable in some embodiments of the invention. "Long-term release," as used herein, means that the implant containing the composition is constructed and arranged to deliver therapeutically effective levels of the composition for at least 30 or 45 days, and preferably at least 60 or 90 days, or even longer in some cases. Long-term release implants are well known to those of ordinary skill in the art, and include some of the release systems described above.

In some embodiments, the compositions of the invention may include

pharmaceutically acceptable carriers with formulation ingredients such as salts, carriers, buffering agents, emulsifiers, diluents, excipients, chelating agents, fillers, drying agents, antioxidants, antimicrobials, preservatives, binding agents, bulking agents, silicas, solubilizers, or stabilizers that may be used with the active

5 compound. For example, if the formulation is a liquid, the carrier may be a solvent, partial solvent, or non-solvent, and may be aqueous or organically based. Examples of suitable formulation ingredients include diluents such as calcium carbonate, sodium carbonate, lactose, kaolin, calcium phosphate, or sodium phosphate; granulating and disintegrating agents such as corn starch or algenic acid; binding

10 agents such as starch, gelatin or acacia; lubricating agents such as magnesium stearate, stearic acid, or talc; time-delay materials such as glycerol monostearate or glycerol distearate; suspending agents such as sodium carboxymethylcellulose, methylcellulose, hydroxypropylmethylcellulose, sodium alginate, polyvinylpyrrolidone; dispersing or wetting agents such as lecithin or other

15 naturally-occurring phosphatides; thickening agents such as cetyl alcohol or beeswax; buffering agents such as acetic acid and salts thereof, citric acid and salts thereof, boric acid and salts thereof, or phosphoric acid and salts thereof; or preservatives such as benzalkonium chloride, chlorobutanol, parabens, or thimerosal. Suitable carrier concentrations can be determined by those of ordinary

20 skill in the art, using no more than routine experimentation. The compositions of the invention may be formulated into preparations in solid, semi-solid, liquid or gaseous forms such as tablets, capsules, elixirs, powders, granules, ointments, solutions, depositories, inhalants or injectables. Those of ordinary skill in the art will know of other suitable formulation ingredients, or will be able to ascertain such, using only

25 routine experimentation.

Preparations include sterile aqueous or nonaqueous solutions, suspensions and emulsions, which can be isotonic with the blood of the subject in certain embodiments. Examples of nonaqueous solvents are polypropylene glycol, polyethylene glycol, vegetable oil such as olive oil, sesame oil, coconut oil, arachis

30 oil, peanut oil, mineral oil, injectable organic esters such as ethyl oleate, or fixed oils including synthetic mono or di-glycerides. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered

media. Parenteral vehicles include sodium chloride solution, 1,3-butandiol, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers (such as those based on Ringer's dextrose), and the like. Preservatives and other additives
5 may also be present such as, for example, antimicrobials, antioxidants, chelating agents and inert gases and the like. In addition, sterile, fixed oils are conventionally employed as a solvent or suspending medium. For this purpose any bland fixed oil may be employed including synthetic mono- or di-glycerides. In addition, fatty acids such as oleic acid may be used in the preparation of injectables. Carrier formulation
10 suitable for oral, subcutaneous, intravenous, intramuscular, *etc.* administrations can be found in Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pa. Those of skill in the art can readily determine the various parameters for preparing and formulating the compositions of the invention without resort to undue experimentation.

15 In some embodiments, the present invention includes the step of forming a composition of the invention by bringing an active compound into association or contact with a suitable carrier, which may constitute one or more accessory ingredients. The final compositions may be prepared by any suitable technique, for example, by uniformly and intimately bringing the composition into association with
20 a liquid carrier, a finely divided solid carrier or both, optionally with one or more formulation ingredients as previously described, and then, if necessary, shaping the product.

In some embodiments, the compositions of the present invention may be present as pharmaceutically acceptable salts. The term "pharmaceutically acceptable
25 salts" includes salts of the composition, prepared in combination with, for example, acids or bases, depending on the particular compounds found within the composition and the treatment modality desired. Pharmaceutically acceptable salts can be prepared as alkaline metal salts, such as lithium, sodium, or potassium salts; or as alkaline earth salts, such as beryllium, magnesium or calcium salts. Examples of
30 suitable bases that may be used to form salts include ammonium, or mineral bases such as sodium hydroxide, lithium hydroxide, potassium hydroxide, calcium hydroxide, magnesium hydroxide, and the like. Examples of suitable acids that may

be used to form salts include inorganic or mineral acids such as hydrochloric, hydrobromic, hydroiodic, hydrofluoric, nitric, carbonic, monohydrogencarbonic, phosphoric, monohydrogenphosphoric, dihydrogenphosphoric, sulfuric, monohydrogensulfuric, phosphorous acids and the like. Other suitable acids include
5 organic acids, for example, acetic, propionic, isobutyric, maleic, malonic, benzoic, succinic, suberic, fumaric, mandelic, phthalic, benzenesulfonic, p-tolylsulfonic, citric, tartaric, methanesulfonic, glucuronic, galacturonic, salicylic, formic, naphthalene-2-sulfonic, and the like. Still other suitable acids include amino acids such as arginate, aspartate, glutamate, and the like.

10 The invention also includes methods for quantitating a level of precursor microRNA expression. The method involves incorporating a precursor microRNA into a reporter system, transfecting a host cell with the reporter system, and detecting expression of a reporter gene product to quantitate the level of precursor microRNA expression. In some embodiments the reporter system includes a firefly luciferase
15 reporter gene.

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

VI. *Exemplary Uses*

Drug Target Validation

Good drugs are potent and specific; that is, ideally, they must have strong
25 effects on a specific biological pathway or tissue (such as the disease tissue), while having minimal effects on all other pathways or all other tissues (*e.g.*, healthy tissues). Confirmation that a compound inhibits the intended target (drug target validation) and the identification of undesirable secondary effects are among the main challenges in developing new drugs.

30 Modern drug screening typically requires tremendous amounts of time and financial resources. Ideally, before even committing to such an extensive drug development program to identify a drug, one would like to know whether the

intended drug target would even make a good target for treating a disease. That is, whether antagonizing the function of the intended target (such as a disease-associated oncogene or survival gene), would be sufficient / effective to treat the disease, and whether such treatment would bear an acceptable risk or side effect. For example, if a cancer is determined to be caused by an activating mutation in the Ras pathway, or caused by abnormal activity of a survival gene such as Bcl-2, the subject system can be used to generate animal models for drug target validation. Specifically, one can generate a transgenic mouse with the subject tet-responsive mishRNA expression, with the mishRNA targeting a gene that is a potential drug target (*i.e.*, Ras or Bcl-2 in this example). Tumors with various initiating lesions can then be made in the mouse, and the mishRNA can be switched on in the tumor (if, for example, a tet-ON regulator is used). Such mishRNA expression mimicks the action of a (yet to be identified) drug that would interfere with that target. If knocking down the target gene is effective to reverse or stall the course of the disease, the target gene is a valid target.

Optionally, the mishRNA transgene can be switched on in a number of tissues or organs, or even in the whole organism, in order to verify the potential side effects of the (yet to be identified) drug on other healthy tissues / organs.

Thus another aspect of the invention provides an animal useful for drug target validation, comprising a germline transgene encompassing the subject artificial nucleic acid, which transcription is driven by a Pol II promoter. The expression of the encoded precursor molecule (such as one based on the miR30-design) leads to an siRNA that targets a candidate drug target. Optionally, the precursor molecule is expressed in an inducible, reversible, and/or tissue-specific manner.

In a related aspect, the invention provides a method for drug target validation, comprising antagonizing the function of a candidate drug target (gene) using a subject cell or animal (*e.g.*, a transgenic animal) encompassing the subject artificial nucleic acid, either *in vitro* or *in vivo*, and assessing the ability of the encoded precursor molecule to reverse or stall the disease progress or a particular phenotype associated with a pathological condition. Optionally, the method further comprises assessing any side effects of inhibiting the function of the target gene on

one or more healthy organs / tissues.

Animal Disease Model

The subject nucleic acid constructs enables one to switch on or off a target gene or certain target genes (*e.g.*, by using crossing different lines of transgenic animals to generate multiple-transgenic animals) inducibly, reversibly, and/or in a tissue-specific manner. This would facilitate conditional knock-out or turning-on of any target gene(s) in a tissue-specific manner, and/or during a specific developmental stage (*e.g.*, embryonic, fetal, neonatal, postnatal, adult, etc.). Animals bearing such transgenes may be treated, such as by providing a tet analog in drinking water, to turn on or off certain genes to allow certain diseases to develop / manifest. Such system and methods are particularly useful, for example, to analyze the role of any known or suspected tumor suppressor genes in the maintenance of immortalized or transformed states, and in continued tumor growth *in vivo*.

In certain embodiments, the extent of gene knock-down may be controlled to achieve a desired level of gene expression. Such animals or cell (healthy or diseased) may be used to study disease progress, response to certain treatment, and/or screening for drug leads.

The ability of the subject system to use a single genomic copy of the Pol II promoter-driven miRNA cassette to control gene expression is particularly valuable for complex library screening.

The subject gene knock-down by expression of shRNA-mirs may be very similar to overexpression of protein-coding cDNAs. Thus any expression systems allowing targeted, regulated and tissue-specific expression, which have traditionally be limited to gene overexpression studies, may now be adapted for loss-of-function studies, especially when combined with the available genome-wide, sequence-verified banks of miR-30-based shRNAs targeting model organisms, such as human and mouse.

Examples

The invention now being generally described, it will be more readily understood by reference to the following examples, which are included merely for purposes of illustration of certain aspects and embodiments of the present invention,

and are not intended to limit the invention.

Introduction

In contrast to RNA Polymerase II promoters which are used by genes
5 encoding proteins, RNA Polymerase III (Pol III) promoters, such as U6 and H1
promoters, normally drive the transcription of several endogenous small nuclear
RNAs (snRNAs). For this reason, Pol III promoters have been widely adopted to
drive transcription of synthetic short hairpin RNAs (shRNAs) in cells and animals.
Applicants and others have used shRNAs driven by U6 promoters to achieve stable
10 knockdown of target genes. Delivery of Pol III promoter-shRNA cassettes by
retroviral transduction of mammalian cells results in stable suppression of target
gene expression.

However, shRNA driven by Pol III promoters has certain practical problems.
First of all, unlike Pol II promoters, the Pol III promoters do not lend themselves to
15 regulation. Secondly, such Pol III-driven shRNAs can be ineffective inhibitors of
their target mRNAs when expressed from a single-copy vector.

Here, Applicants have used certain RNA Polymerase II (Pol II) expression
systems to allow potent and regulatable RNAi in mammalian cells. Applicants have
shown that miR30 design shRNAs expressed from the LTR promoter of an
20 integrated retrovirus suppress target genes more effectively than when expressed
from an RNA polymerase III promoter, even when expressed from a single-copy in
the genome (*e.g.*, from a stably transfected or a transgenic copy). Furthermore,
regulated shRNA expression was also achieved by using inducible / reservable Pol II
promoters, such as a Tet-responsive Pol II promoter.

25

Example I RNA Polymerase III Promoters are Sufficient for Stable shRNA Expression

30 In order to identify a preferred retroviral vector for delivery of promoter-
shRNA cassettes into mammalian cells, Applicants compared two vectors based on
the murine stem cell virus (MSCV) and the self-inactivating (SIN) retroviral vector,

respectively. The 5'-long terminal repeat (5'-LTR) promoter of the SIN provirus is transcriptionally inactive, thus using the SIN 5'-UTR promoter in this construct (*e.g.*, USP-p53C, see below) serves as a control for the construct using the functional MSCV Pol II promoter (*e.g.*, LUSP-p53C, see below).

5 Into each vector, Applicants cloned the U6 RNA polymerase III promoter upstream of a sequence encoding p53C, a short hairpin RNA (shRNA) that targets murine p53. The shRNA p53C is predicted to fold into a simple, symmetrical shRNA with a 29-nucleotide stem and an eight nt loop (Figure 1A, left). Also in each vector, a puroR-IRES-GFP (PIG) cassette under the control of the PGK
10 promoter was operably linked downstream of the U6-shRNA cassette (Figure 1B).

Primary murine embryonic fibroblasts (MEFs) were infected with either the SIN-U6shRNA-PIG (UP)-p53C construct (USP-p53C), the MSCV-U6shRNA-PIG (LUP)-p53C construct (LUSP-p53C), or a control virus, and subject to puromycin selection to establish stably-infected cell lines.

15 After treatment with the DNA-damaging agent adriamycin to induce p53 expression, cells stably-integrated by the above constructs were harvested, and their p53 expression levels were assessed by Western blotting. Interestingly, Applicants found that p53 knockdown was far more effective in cells transduced with the SIN retrovirus (Figure 1C), indicating that the internal U6 Pol III promoter is sufficient
20 for expression of the p53C shRNA in mammalian cells (since the SIN Pol II promoter is inactive in the USP-p53C construct). Also surprisingly, our observations suggest that transcription from the upstream MSCV LTR promoter, a strong RNA polymerase II promoter, inhibited shRNA function and p53 knockdown in this context. This effect may be due to promoter interference between the LTR Pol II and
25 U6 Pol III promoters.

Similar results were obtained for several other shRNAs with simple stem/loop folds similar to p53C (data not shown), verifying the general applicability of using RNA polymerase III promoters alone for expression of this style (the simple stem-loop style) of shRNA.

30

**Example II LTR Pol II Promoter is More Effective than RNA
Polymerase III Promoter in Directing Integrated miR30-
Design shRNAs Suppression of Target Genes**

This example demonstrates that synthetic shRNAs with folds designed to
5 mimic endogenous microRNA (miRNA) precursors can effectively inhibit target
gene expression. To illustrate, Applicants used the exemplary miR30-design
shRNAs (designated microRNA-based shRNAs, or mishRNAs) to demonstrate
stable suppression of gene expression in mammalian cells, which strategy can be
generally applied to other microRNA (miRNA) precursors. Specifically, Applicants
10 recovered a mishRNA referred to as p53.1224 (so named because the predicted
siRNA begins at nucleotide 1224 of the p53 cDNA) from the mishRNA library (a
genome wide miR30-based shRNA library).

As shown above, standard stem-loop shRNAs are most effectively expressed
from RNA polymerase III (Pol III) promoters such as the U6 promoter. Applicants
15 sub-cloned a U6 promoter-p53.1224 cassette into a murine stem cell virus (MSCV)
and a self-inactivating (SIN) retroviral vector, thus generating two vectors designed
to express miR-based shRNAs (as opposed to the stem-loop shRNA): MSCV/LTR-
U6miR30-PIG (LUMP)-p53.1224 and SIN-U6miR30-PIG (UMP)-p53.1224 (Fig.
1B). One difference between the mishRNA and the standard stem-loop shRNA is
20 that the miR30 precursor RNA is approximately 300 nt in length and is predicted to
fold into an extensive secondary structure (Figure 1A, right).

Applicants have previously constructed similar vectors expressing a standard
stem-loop shRNA targeting p53 (p53C), producing MSCV/LTR-U6shRNA-PIG
(LUSP)-p53C or SIN-U6shRNA-PIG (USP)-p53C (see above and Fig. 1B). All four
25 constructs were introduced into early passage murine embryonic fibroblasts (MEFs).
The resulting cell populations were assessed for p53 knockdown after adriamycin
treatment (a DNA damaging agent that stabilizes p53), and the ability to form
colonies when plated at low density (a functional readout of p53 loss).

Contrast to what was observed for the simple stem-loop shRNA, the MSCV-
30 based p53.1224 mishRNA (LUMP-p53.1224) driven by a functional Pol II promoter
was more effective at suppressing p53 than its SIN-based counterpart (UMP-
p53.1224) devoid of of a functional Pol II promoter, producing nearly undetectable

p53 levels as assessed by immunoblotting (Fig. 1C, compare lanes 6, 7, and 8). As shown above, for the standard stem-loop shRNA, the SIN-based p53C shRNA (USP-p53C) was more effective at suppressing p53 than its MSCV-derived counterpart (LUSP-p53C), verifying that the U6 promoter is sufficient for stem/loop shRNA expression (Fig. 1C). The ability of each vector to suppress p53 correlated precisely with its ability to stimulate colony formation at low density, with cells expressing the MSCV-based p53.1224 vector producing as many colonies as p53-null cells (Fig. 1D).

Southern blotting using a GFP probe verified that these differences were not due to variation in retroviral copy number (data not shown).

This vector preference was also observed for several other mishRNAs and stem-loop shRNAs targeting diverse gene products (data not shown). Thus, in general, mishRNAs can be remarkably potent when stably expressed from retroviral vectors, particularly those with a functional 5'-LTR (with a Pol II promoter). In the examples shown herein, this system achieved near-complete (if not complete) target gene knockdown.

Example III Pol II Promoter Contributes to Functional shRNA Production

The more potent knockdown produced by mishRNAs expressed from the MSCV vector compared with the SIN vector implies that the 5'-LTR contributes to optimal mishRNA expression. To determine whether the 5'-LTR promoter, a strong Pol II promoter, is sufficient for effective gene knockdown using mishRNAs, Applicants introduced the p53.1224 shRNA into an MSCV vector lacking a U6 promoter (MSCV/LTRmiR30-PIG (LMP) (Fig. 1B). To facilitate comparison, Applicants introduced this vector and its LUMP and UMP counterparts into NIH3T3 cells at a low multiplicity of infection (<5% efficiency) such that the vast majority of transduced cells should contain single proviral integrations. Remarkably, both vectors harboring the MSCV LTR (LUMP-p53.1224 and LMP-p53.1224) suppressed p53 expression extremely efficiently, and were far superior to UMP-p53.1224, which expresses p53.1224 from the U6 promoter alone (Fig. 1E). Similar results were obtained in other cell types including wild type and p19ARF-null MEFs

III promoter) formed tumors much more rapidly than animals reconstituted with stem cells expressing a control vector (Fig. 2A, $P < 0.05$). Importantly, lymphomas arising in animals transduced with shBim were GFP-positive, expressed low levels of Bim (Fig. 2B), and displayed a mature (IgM⁺) B cell phenotype uniquely characteristic of Bim null lymphomas (data not shown). Thus, *in vivo* loss of function phenotypes can be recapitulated using miR30-design shRNAs expressed from Pol II promoters.

Example V Identification and Characterization of Genes that Modify Drug Responses *in vivo*

siRNAs have been used to identify modulators of drug action, but are not suitable for long-term assays or *in vivo* systems. The miR30-based vectors described above enable the identification and characterization of genes that modify drug responses *in vivo*.

As an illustrative example, Applicants examined the ability of a miR30-design p53 shRNA to promote chemoresistance in E μ -myc lymphomas, which respond poorly to therapy in the absence of p53. Applicants introduced LMS-p53.1224 (co-expressing GFP) into chemosensitive E μ -myc lymphoma cells at ~10% infection efficiency and transplanted the mixed cell populations into syngeneic recipient mice. Upon lymphoma manifestation, animals were treated with the chemotherapeutic drug adriamycin and monitored for tumor response. Strikingly, mice harboring lymphomas expressing LMS-p53.1224 did not regress following adriamycin treatment and showed significantly reduced overall survival relative to control tumor-bearing mice (Fig. 2C). This indicates that the LMS-p53.1224 construct effectively knocked-down p53 expression in tumor cells, resulting in their poor response (or chemoresistance) to therapy. Furthermore, the percentage of GFP positive cells dramatically increased in lymphomas harboring p53.1224 but not the control vector, indicating a selective advantage for p53.1224 expressing cells (Fig. 2D).

These results demonstrated that sufficient p53 knockdown may promote *in vivo* chemoresistance. Such an animal mode (or tumor cells therein) may also be used to screen (*in vivo* or *in vitro*) for compounds that can overcome

chemoresistance in p53 negative cells.

Together, these data indicate that mishRNAs expressed from Pol II promoters are suitable for a variety of *in vivo* applications, with strong potential for transgenic animals, tissue specific gene knockdowns and *in vivo* forward genetic screens.

Example VI Pol II Promoter-Driven Inducible and Reversible shRNA Production from Low-Copy Stable Integration

RNAi inhibits gene function without altering DNA sequence, therefore its effects are potentially reversible. Given our findings that low copy Pol II promoters can effectively drive mishRNAs from a single integrated construct, Applicants adapted the traditional inducible protein expression systems, such as the tetracycline (tet)-regulated Pol II promoter TRE-CMV, to achieve inducible stable expression of mishRNAs.

Many inducible promoters are known in the art in the context of protein expression. These inducible systems can all be adapted to express the mishRNAs of the subject invention. In one illustrative example, the TRE-CMV promoter consists of a tandem array of tet transactivator binding sites fused to a minimal CMV promoter. Transactivator protein tTA transactivates the TRE-CMV promoter in the absence of the tetracycline analog doxycycline (Dox). This promoter system has been shown to be highly effective for conditional expression of protein-coding cDNAs both *in vitro* and *in vivo*. Thus when adapted for use in the subject invention, shRNA expression can mediate target gene knockdown in the absence of Dox both *in vitro* and *in vivo*.

Using a SIN vector backbone, Applicants cloned a mishRNA targeting human Rb (Rb.670) downstream of the TRE-CMV promoter, producing SIN-TREmiR30-PIG, or TMP-Rb.670; Fig. 3A). HeLa cells stably expressing the tet transactivator protein tTA (tet-off) were infected with TMP-Rb.670 at single copy in the absence of Dox. Rb levels in these cell populations were slightly decreased compared with uninfected controls, indicating potential shRNA production from the TRE-CMV promoter (Fig. 3B). Indeed, when single cell clones were generated from this population, 6 of 13 showed excellent Rb knockdown (Fig. 3C and data not

shown), demonstrating that the TRE-CMV promoter can effectively drive shRNA expression at low copy number.

To examine inducible regulation of shRNA expression, Rb.670C cells, which showed significant Rb knockdown in Dox-free medium (Fig. 3C), were cultured in various Dox concentrations for many days. Cell growth and viability were indistinguishable at all Dox concentrations. However, Applicants observed a clear dose-dependency of Rb expression, with maximum Rb knockdown achieved in low Dox concentrations and vice versa (Fig. 3D). At Dox concentrations of less than 0.005 ng/mL, Dox produced minimal Rb expression. However, cells grown in 0.008 ng/mL Dox showed slight de-repression of Rb. Normal Rb expression was restored in cells cultured in approximately 0.05 ng/mL Dox and higher, suggesting that shRNA expression is not leaky at these Dox concentrations.

Thus, Dox concentration can tightly control the extent of stable gene knockdown. This effect was also observed in time-course studies, where Applicants observed normalization of Rb expression upon Dox addition, and rapid Rb knockdown upon Dox removal (Fig. 3E), demonstrating the reversibility of the induced shRNA expression. Remarkably, in all cases GFP and Rb levels were inversely correlated (Fig. 3D and 3E), with intermediate GFP expression observed between 0.002 and 0.008 ng/mL Dox. As GFP and shRNA are produced from the same transcript, GFP expression may be regarded as a surrogate marker of shRNA production.

A great advantage of tet-regulated systems is that expression from the TRE-CMV promoter can be either induced (tet-ON) or repressed (tet-OFF) by Dox, depending on which tet transactivator protein is used. To test a tet-ON shRNA expression system, Applicants utilized U2OS cells stably expressing the reverse tTA (rtTA) protein, which in contrast to tTA, requires Dox to activate transcription.

Applicants also isolated a clone (Rb.670R5; Fig. 3F) of U2OS cells infected with TMP-Rb.670 and stably expressing the reverse tTA (rtTA; tet-on) protein. As predicted, Dox concentration and Rb knockdown were positively correlated in these cells (Fig. 3G and 3H).

At Dox concentrations of less than 0.005 ng/mL, Dox produced minimal Rb expression. However, cells grown in 0.008 ng/mL Dox showed slight de-repression

of Rb. Normal Rb expression was restored in cells cultured in approximately 0.05 ng/mL Dox and higher, suggesting that shRNA expression is not leaky at these Dox concentrations. As GFP protein is translated from an IRES, it can be produced from transcripts originating from both the PGK and TRE-CMV promoters. As GFP is not
5 detected in cells grown in high Dox concentrations, it appears that GFP production from the PGK promoter transcript is very weak. Our results suggest that the majority of GFP in untreated Rb.670C cells arises from the CMV-TRE transcript, production of which is blocked by Dox in a dose-dependent manner. As the TRE-CMV transcript also carries the miR30-based shRNA fold, GFP expression may be
10 regarded as a surrogate marker of shRNA production.

Using the same Tet-responsive system, good protein expression regulation was also achieved in several other cell clones, including those expressing a PTEN-miR30 construct.

These observation verifies that low copy delivery of the TMP vector (also
15 lacking a Pol III promoter) allows regulated mishRNA expression in either tet-on or tet-off configurations, and altering Dox concentration in this system allows tight control of the extent of stable gene knockdown.

Example VII Reversible Induction of Pol II-Driven Tet-Responsive p53 20 shRNA Production in Primary Cells

The instant regulatable shRNA expression is not only operable in immortalized cell lines, but also functional to regulate suppression of genes (*e.g.*, the tumor suppressor gene p53) in primary cells.

For example, inactivation of the tumor suppressor p53 immortalizes wild
25 type MEFs, and transforms MEFs over-expressing oncogenic Ras. Early passage MEFs were co-transduced with TMP-p53.1224 and a retrovirus expressing the tTA (tet-off) protein. Many doubly infected MEFs (designated wild type/tTA/TMP-p53.1224, or WtT) showed stable p53 knockdown when cultured in Dox-free medium. WtT cells plated at low density formed colonies comparable in size and
30 number to those formed by p53-null MEFs (Fig. 4A), suggesting that p53 was functionally inactivated in most cells. Colony formation of WtT cells cultured in Dox in parallel was similar to that of control cells (Fig. 4A), suggesting normal p53

expression. p53-null MEF growth was unaffected by Dox, ruling out non-specific effects (Fig. 4A).

Applicants also isolated several WtT clones and examined their p53 regulation in response to Dox. p53 expression in WtT cells increased rapidly and GFP expression was lost upon Dox treatment (Fig. 4B and 4C). WtT clones cultured in Dox failed to form colonies when plated at low density. Instead, by day 8 of Dox treatment, all cells had a flattened morphology characteristic of senescent cells, and many were positive for senescence-associated β -galactosidase (SA- β -gal; Fig. 4C). This dormant phenotype was stable for weeks of continuous culture in Dox. Therefore, p53.1224 shRNA expression can be tightly regulated by Dox treatment in wild type MEFs doubly infected with tTA and TMP-p53.1224.

The rapid and coordinated senescence response observed when endogenous p53 expression was restored in MEFs immortalized by p53 knockdown was reversed upon Dox removal (Fig. 4D, left panel, upper well), in agreement with previous observations in other MEF culture systems. Control cells continually cultured in Dox remained dormant (Fig. 4D, left panel, lower well). Newly proliferating cells (Fig. 4D, left panel, upper well) remained responsive to p53 re-expression, as they lost GFP expression and failed to form colonies when re-plated in Dox (Fig. 4D, right panel, lower well).

These results suggest that wild type MEFs can be reversibly switched between cycling and senescent states simply by regulating p53 knockdown. WtT cells transformed by infection with activated Ras (Fig. 4E, upper panels) also became morphologically senescent and SA- β -gal positive when treated with Dox (Fig. 4E, lower panels), with p53 and GFP expression changes similar to that of parental WtT cells (Fig. 4F).

Furthermore, Applicants conclude that restoration of p53 expression in wild type MEFs immortalized by stable p53 knockdown causes a rapid and coordinated senescence response. This demonstrates that at least in cancers with p53 loss-of-function mutations, cancers can be treated by restoring p53 expression to induce senescence. This technique can also be extended to test any potential target genes whose functions are lost in diseases, such as in cancer. Specifically, the system of the instant invention may be used to test whether loss-of-function of a candidate

gene causes certain disease state, and whether restoring such target gene function in diseased tissues can reverse the disease status, or at least slow down disease progression.

5 Example VIII Reversible *in vivo* Gene Knockdown Using Tet-Responsive Promoter

Tet-regulated over-expression systems have revolutionized the study of the role of oncogenes in tumor survival *in vivo*. Tet-regulated RNAi holds similar promise for regulated knockdown of tumor suppressor genes. To illustrate this concept, Applicants injected WtT-Ras MEFs subcutaneously into the flanks of nude mice formed visible, rapidly growing and strongly GFP positive tumors after approximately 2 weeks, verifying that these cells were functionally transformed (Fig. 5A; upper panels). To inactivate p53.1224 shRNA in established tumors, mice were treated with Dox via their drinking water. After only 2 days of Dox treatment, tumor GFP intensity was markedly diminished compared with untreated mice, and after 4 days tumors were almost GFP negative (Fig. 5A). Remarkably, tumor growth slowed upon Dox treatment, and tumors began shrinking after approximately 4 to 6 days (Fig. 5B). Animals treated with Dox for 10 days often showed continued tumor regression and became tumor-free (Fig. 5B). This regression was p53-dependent, as tumors derived from p53-null MEFs infected with tTA, TMP-53.1224 and Ras lost GFP expression but continued to grow when treated with Dox (data not shown). Similar results were obtained for several WtT-Ras clones and WtT clones infected with E1A/Ras, with variable tumor growth rates and regression kinetics (data not shown).

Dox-treated animals with regressing tumors were taken off Dox treatment after various times. In many cases, usually when initial tumor size was less, mice became tumor-free and remained so for weeks. However, removing Dox from animals with larger regressing tumors or after a briefer Dox treatment often allowed renewed GFP expression and tumor growth (Fig. 5C). Interestingly, tumors isolated from Dox-treated animals contained cells with unusually compact nuclei, and widespread apoptosis was seen compared with untreated controls (Fig. 5D), suggesting that tumor regression was at least in part due to p53-dependent apoptosis.

Indeed, as predicted, p53 expression was dramatically elevated in tumors from Dox-treated animals (Fig. 5E).

In summary, by adapting a standard Pol II promoter-driven tet-responsive promoter normally used for inducible protein expression, Applicants for the first time have demonstrated inducible and reversible target gene knockdown *in vivo*. p53 re-expression in tumors caused regression associated with widespread apoptosis, in contrast to the senescence observed when p53 was re-expressed in the same cells in culture. These findings highlight the ability of this technology for the study of many aspects of biology, including identification and/or validation of potential drug targets in animal models. The tet system has obvious advantages over unidirectional Cre-lox strategies, and many key reagents, such as tissue-specific tet transactivator mice, are readily available.

In summary, expression of miRNA-design short hairpin RNAs (shRNAs) allows stable, post-transcriptional suppression of gene activity, which is optionally reversible. Applicants have developed a new retroviral vector system that uses RNA polymerase II promoters to express shRNAs based on the human miR30 precursor. Single copy expression of shRNAs from this vector yields potent and stable gene knockdown in cultured cells and *in vivo*. Expression of an shRNA targeting p53 using this system mimics complete p53 loss and renders tumor cells chemoresistant *in vivo*. By improving standard tet-inducible promoters for shRNA expression, we show stable, incremental, and reversible gene knockdown of various target genes in tet-on or tet-off configurations. Interestingly, cultured wild type mouse fibroblasts can be switched from proliferative to senescent states simply through regulated knockdown of p53. We find that tumors derived from wild type mouse fibroblasts transformed by Ras overexpression and p53 knockdown regress upon p53 re-activation *in vivo*, suggesting that ongoing suppression of p53 is essential for tumor maintenance in this context. This system proves useful for studying potential therapeutic targets in cancer, and in most other biological systems.

All vectors described in these experiments are compatible with genome-wide, sequence verified banks of miR30 shRNAs (or any other similar banks of miR shRNAs) targeting human and mouse genes, creating a formidable resource for

diverse, large scale RNAi studies in mammalian systems.

METHODS

The following methods and reagents were used in the Examples above.

- 5 These are merely for illustrative purpose, and are by no means limiting. Other comparable minor variations can be readily made without undue experimentation for adapting to specific problems.

Vector construction.

- 10 The retroviral vector MSCV-PIG has an EcoRI site in the polylinker and another between the Puro^R cassette and the IRES sequence. To facilitate cloning into the polylinker, the second site was destroyed using a PCR-based strategy: a PCR product was generated using MSCV-PIG template, forward primer 5'-TCTAGGCGCCGGAATTAGATCTCTCG-3' (SEQ ID NO: 1), and reverse primer
15 5'-CCTGCAATTGGATGCATGGGGTCGTGC-3' (SEQ ID NO: 2), and digested with BglII and MfeI. This fragment was cloned into MSCV-PIG digested with BglII/EcoRI, yielding MSCV-PIGdRI. MSCV-U6miR30-PIG was made by ligating the 762 bp BamHI-MfeI "U6-miR30 context" cassette from pSM2 into MSCV-PIGdRI digested with BglII/EcoRI. MSCV-LTRmiR30-PIG was made by ligating
20 the 256 bp Sall-MfeI "miR30 context" cassette from pSM2 into MSCV-PIGdRI digested with XhoI/EcoRI. MSCV-LTRmiR30-SV40GFP (LMS) was made in two steps. Firstly, a ~1.2 kb EcoRI-ClaI SV40GFP fragment from pBabeGFP was ligated into MSCV-puro (Clontech) digested with EcoRI/ClaI, forming MSCV-SV40GFP. This was digested with XhoI/EcoRI, and the 256 bp Sall-MfeI "miR30
25 context" cassette from pSM2 was inserted, forming MSCV-LTRmiR30-SV40GFP. SIN-PIGdRI was made by ligating the 2524 bp EcoRI-Sall fragment from MSCV-PIGdRI into pQCXIX (Clontech) digested with EcoRI/XhoI. SIN-TREmiR30-PIG was constructed in two steps. Firstly, a PCR product spanning the TRE-CMV promoter was generated using template plasmid pQTXIX (a kind gift from Abba
30 Malina, generated by cloning the XbaI-EcoRI TRE-CMV promoter fragment from pUHD10.3 into pQCXIX (Clontech) digested with XbaI/EcoRI), using the primers 5'-GAATTGAAGATCT

GGGGGATCGATC-3' (SEQ ID NO: 3) and 5'-
CATCAATTGCTAGAATTCTGGTTGCT
CGAGAGGCTGGATCGGTCCCGGTGTCTTC-3' (SEQ ID NO:4). This PCR
product was digested with BglII / MfeI and ligated into SIN-PIGdRI digested with
5 BglII / EcoRI (removing its CMV promoter), forming SIN-TRE-PIG. SIN-
TREmiR30-PIG was completed by ligating the 256 bp Sall-MfeI "miR30 context"
cassette from pSM2 into SIN-TRE-PIG digested with XhoI / EcoRI. DNA fragments
encoding various mishRNA folds were generated using oligonucleotide template
PCR as described previously, or subcloned as 110 bp XhoI / EcoRI fragments from
10 the pSM2 mishRNA library. Oligonucleotides were designed at
katahdin.cshl.org:9331/siRNA/
RNAi.cgi?type=shRNA (incorporated herein by reference). PCR products were
digested with XhoI/EcoRI and ligated into the unique XhoI / EcoRI sites within the
"miR30 context" in the vectors described above.

15

Cell Culture and Expression Analysis

Cells were grown in DMEM containing 10% fetal bovine serum at 37°C with
7.5% CO₂. Doxycycline (Clontech) was dissolved in water and generally used at a
final concentration of 100 ng/mL. Medium containing Dox was refreshed every two
20 days. Infections and colony formation assays were carried out as previously
described. SA-β-gal activity was detected as previously described, with sample
equilibration and X-gal staining done at pH 5.5. For western blotting analysis,
Laemmli buffer protein lysates were run on SDS-PAGE, and transferred to
Immobilon PVDF membrane (Millipore). Antibodies were anti-p53 (1:1000 IMX25,
25 Vector Laboratories), anti-PTEN (1:1000 486, a kind gift from Michael Myers),
anti-GFP (1:5000, Clontech), anti-tubulin (1:5000 B-5-1-2, Sigma), anti-actin
(1:5000, Sigma), and anti-Rb (1:1000 G3-245, Pharmingen with 1:100 XZ-55 and
C36 hybridoma supernatants).

30 Lymphoma studies.

Eμ-myc lymphomagenesis and drug treatment studies were performed as
previously described (Schmitt, 2000; Hemann, 2003). Chemosensitive lymphoma

cells were isolated from tumors arising in mice transplanted with E μ -myc; p19^{ARF}^{+/-} HSCs, which invariably lose the wild type p19^{ARF} allele while retaining wild type p53.

5 Nude mouse studies

Approximately 10⁶ transformed cells were injected subcutaneously into the two rear flanks of nude mice. Mice were treated with 0.2 mg/mL Dox in a 0.5 % sucrose solution in light-proof bottles, refreshed every four days. Tumor volume (mm³) was calculated as (length \times width² \times $\pi/6$). For analysis of protein expression, 10 tumors were snap-frozen and pulverised in liquid nitrogen using a mortar and pestle. Powdered tumor was lysed in Laemmli buffer and western analysis was performed as above. For histology, tumor tissue was fixed for 24 hours in 4% formaldehyde in PBS prior to embedding and sectioning. Apoptosis was measured by TUNEL assay (In situ Cell Death Detection Kit, POD; Roche).

15 Results described herein above are published in *Nat Genet.* 37(11): 1289-95, 2005 (Dickins *et al.*, 2005). Other related work is published in *Nat Genet.* 37(11): 1281-88, 2005 (Silva *et al.*, 2005). The entire contents of these publications, including the related online (supplemental) information and contents of the publications cited therein are incorporated herein by reference. The subject system 20 can also be used in lentiviral, pre-miR-30based siRNA expression vectors, such as those including a tetracyclin-responsive Pol II promoter and thus can be used to tightly regulate the expression of target genes in transduced cells. See Stegmeier *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**: 13212-17, 2005 (incorporated herein in its entirety).

25

Example IX *In vivo* Transgenic Animal Model for Tissue-Specific and Inducible Target Gene Knockdown

This example demonstrates knockdown of a target gene, *e.g.*, p53, in a tissue-specific, inducible and/or reversible manner, in a germline (transgenic) animal 30 model.

To achieve regulated transgene expression in germline transgenic mice, two lines of transgenic mice were generated: one expressing a tet transactivator protein

(either tTA/tet-off or rtTA/tet-on), optionally in a tissue-specific manner using tissue-specific promoters; and another harboring a tetracycline-responsive (TRE) promoter driving the transgene of interest. Crossing these two lines yielded double transgenic mice that expressed the transgene, in a Dox-regulatable manner (either
5 tet-on or tet-off), in cells where the transactivator (tTA or rtTA) was expressed.

Alternatively, the tTA or rtTA construct may be introduced (*e.g.* via infection or transfection, *etc.*) into cells of a transgenic animal bearing TRE-miRNA-expression cassette.

For example, to demonstrate that tet-regulated miR30-based shRNA
10 expression can be achieved in animals, Applicants generated several transgenic founder lines harboring a TRE-p53.1224 shRNA cassette (using standard pronuclear injection protocols). To test shRNA activity in these animals, Applicants isolated MEFs (mouse embryonic fibroblasts) from F2 transgenic embryos and wild-type controls, infected them with a retrovirus expressing the tTA (tet-off) protein, and
15 assessed p53 knockdown after p53 induction by adriamycin. Specifically, primary MEFs derived from embryos from a cross between wild-type B6 females mated to TRE-p53.1224 founder lines A and B, were infected with either tTA-IRES-Neo or tTA-IRES-GFP retrovirus. Then tTA-IRES-Neo MEFs were selected for 7 days in G418 prior to harvesting in order to eliminate uninfected cells. The tTA-IRES-GFP
20 MEFs were unselected, though the MEFs were infected at high percentage. All cells were adriamycin treated prior to harvesting.

Of the two founder lines tested so far, one (line A) showed striking p53 knockdown in transgene-positive cells (results not shown). This knockdown was similar to that seen when the p53.1224 shRNA was expressed from a retroviral LTR promoter (*supra*; Dickins *et al.*, *Nat Genet.* 37(11): 1289-95, 2005). Importantly,
25 p53 induction was normal in uninfected transgene-positive cells (*e.g.*, by comparing founder line A MEFs either uninfected or infected with tTA-IRES-Neo. All cells were adriamycin treated prior to harvesting. Results not shown). This demonstrates that shRNA expression and p53 knockdown is tTA-dependent and not leaky.

Moreover, as expected, these MEF lines showed a rapid re-expression of p53
30 upon Doxycycline treatment, indicating that shRNA expression was tightly controlled by Doxycycline (results not shown).

To our knowledge, the above experiments for the first time demonstrated that tetracycline effectively regulated shRNA expression in a germline transgenic setting. This enables one to reversibly switch any endogenous gene on or off, simply by administering a reversible activator or inhibitor of a transcriptional regulator, such as
5 Doxycycline (or other Tet homolog), preferably via drinking water. This technology is especially powerful in examining gene function *in vivo*, for example, during embryonic or postnatal development, tumorigenesis, or after treatment of tumors with chemotherapeutic drugs.

As indicated above, Applicants have also crossed the TRE-p53.1224
10 transgenic mice to established transgenic lines that express the tTA (tet-off) protein in a tissue-specific manner. As expected, Applicants detected p53.1224 siRNA in the liver of LAP-tTA; TRE-p53.1224 double transgenic mice, where tTA expression was restricted to the liver (lane 2 of Figure 6). After 4 days of doxycycline administration, some attenuation of siRNA production was observed (lane 3 of
15 Figure 6). Applicants have also been assessing p53 knockdown in the liver of these mice, in order to determine whether longer term doxycycline administration will further or completely block siRNA production. Note that the spleens of these mice were devoid of siRNA (see lanes 4-6 of Figure 6), consistent with liver-specific expression of the siRNA.

20 The system can also be used to generate animal models for studying the effect of turning on/off certain target genes in the progression of certain diseases, such as cancer.

For example, the E μ -myc mouse is prone to developing lymphoma, which is accelerated further by loss of p53 function. To model this process using tet-regulated
25 p53 knockdown *in vivo*, Applicants crossed E μ -myc mice to TRE-p53.1224 mice and E μ -tTA mice, which expressed tTA specifically in the B cell compartment. As myc and tTA should be expressed coordinately in B cells of the E μ -myc; E μ -tTA; TRE-p53.1224 triple transgenic mice, Applicants expected reversible p53 knockdown in oncogene-expressing cells.

30 Consistent with this prediction, in a spleen (a tissue enriched for lymphoma cells) isolated from lymphoma-laden triple transgenic mice, Applicants observed highly abundant p53.1224 siRNA, at levels known to promote p53 knockdown and

tumor progression (lanes 9-11 of Figure 6). In contrast, spleens from E μ -myc; TRE-1224 double transgenic mice do not express the siRNA, indicating that the TRE-1224 transgene requires tTA for expression.

Lymphoma cells isolated from these triple transgenic mice were then
5 transplanted into several recipient nude mice to allow controlled p53 re-activation. Specifically, tumor-bearing recipient mice were treated with Doxycycline. Survival of heavily tumor-bearing transplant recipients was extended by many days when doxycycline was administered via the drinking water. Furthermore, p53.1224 siRNA expression was completely suppressed in the lymph nodes and spleen of these
10 treated mice, indicating effective switching of shRNA expression *in vivo*.

These results demonstrated that Applicants can produce transgenic mice where miR30-based shRNA production was tissue-specific, and can be inducibly and reversibly regulated simply by administering or omitting doxycycline in the drinking water.

15

The practice of aspects of the present invention may employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the
20 literature. See, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); *DNA Cloning*, Volumes I and II (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis *et al.* U.S. Patent No: 4,683,195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Transcription And*
25 *Translation* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor
30 Laboratory); *Methods In Enzymology*, Vols. 154 and 155 (Wu *et al.* eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds., 1986); *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

All patents, patent applications and references cited herein are incorporated in their entirety by reference.

Equivalents

5 Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

CLAIMS:

1. An artificial nucleic acid construct comprising an RNA Polymerase II (Pol II) promoter operably linked to a coding sequence for expressing a precursor molecule for an siRNA, said siRNA inhibiting the expression of a target gene, wherein the nucleic acid construct directs the expression of the precursor molecule and/or the siRNA, and substantially inhibits the expression of the target gene when stably integrated into a host cell genome.
2. The nucleic acid construct of claim 1, wherein said Pol II promoter is an inducible promoter, a tissue-specific promoter, and/or a developmental stage-specific promoter.
3. The nucleic acid of claim 1 or 2, wherein the inducible promoter is a tetracyclin-responsive promoter.
4. The nucleic acid construct of claim 3, wherein the tetracyclin-responsive promoter is a TetON promoter, the transcription from which promoter is activated at the presence of tetracyclin (tet), doxycycline (Dox), or a tet analog.
5. The nucleic acid construct of claim 3, wherein the tetracyclin-responsive promoter is a TetOFF promoter, the transcription from which promoter is turned off at the presence of tetracyclin (tet), doxycycline (Dox), or a tet analog.
6. The nucleic acid construct of claim 2, wherein the inducible promoter is selected from: a promoter operably linked to a lac operator (LacO), a LoxP-stop-LoxP system promoter, or a GeneSwitchTM or T-RExTM system promoter (Invitrogen).
7. The nucleic acid construct of claim 1 or 2, wherein the Pol II promoter is an LTR promoter or a CMV promoter.
8. The nucleic acid construct of any of claims 1-7, wherein the precursor molecule is a precursor microRNA.
9. The nucleic acid construct of claim 8, wherein the precursor microRNA

(miR) is an artificial miR comprising coding sequence for said siRNA for said target gene.

10. The nucleic acid construct of claim 9, wherein the miR comprises a backbone design of microRNA-30 (miR-30).
- 5 11. The nucleic acid construct of claim 9, wherein the miR comprises a backbone design of miR-15a, -16, -19b, -20, -23a, -27b, -29a, -30b, -30c, -104, -132s, -181, -191, -223.
12. The nucleic acid construct of any of claims 1-7, wherein the precursor molecule is a short hairpin RNA (shRNA).
- 10 13. The nucleic acid construct of any of claims 1-12, wherein a single integrated copy of the nucleic acid construct is sufficient for substantially inhibiting the expression of the target gene.
14. The nucleic acid construct of any of claims 1-13, further comprising an enhancer for the Pol II promoter.
- 15 15. The nucleic acid construct of any of claims 1-14, further comprising a reporter gene under the control of a second promoter.
16. The nucleic acid construct of claim 15, wherein the reporter gene is a luciferase, a fluorescent protein (*e.g.* GFP, RFP, YFP, BFP, *etc.*), or an enzyme.
- 20 17. The nucleic acid construct of claim 15, wherein the second promoter and the reporter gene is downstream of (3'-to) the coding sequence for the precursor molecule.
18. The nucleic acid construct of any of claims 15-17, wherein the reporter gene is translated from an internal ribosomal entry site (IRES) between a second
25 promoter and the reporter gene.
19. The nucleic acid construct of any of claims 1-18, further comprising at least one selectable marker.
20. The nucleic acid construct of claim 19, wherein the selectable marker is puromycin, zeocin, hygromycin, or neomycin.

21. The nucleic acid construct of any of claims 1-20, further comprising a reporter gene, wherein the coding sequence for expressing the precursor molecule is embedded or inserted into the 5'-UTR (untranslated region), 3'-UTR, or an intron of the reporter gene.
- 5 22. The nucleic acid construct of any of claims 1-21, further comprising a Pol III promoter upstream of the coding sequence for expressing the precursor molecule.
23. The nucleic acid construct of any of claims 1-22, wherein the target gene is associated with a disease condition selected from cancer or infectious
10 disease.
24. The nucleic acid construct of claim 23, wherein the target gene is over-expressed or abnormally active in the disease.
25. The nucleic acid construct of claim 23 or 24, wherein the target gene is an oncogene or an antagonist / inhibitor or dominant negative mutation of a
15 tumor suppressor gene.
26. A cell comprising the nucleic acid construct of any of claims 1-25.
27. The cell of claim 26, which is a mammalian cell.
28. The cell of claim 26 or 27, which is a tissue culture cell, a cell *in vivo*, or a cell manipulated *ex vivo*.
- 20 29. The cell of any of claims 26-28, wherein the Pol II promoter is an inducible promoter, and wherein the cell further comprises an additional construct for expressing an activator or an inhibitor of the inducible promoter.
30. The cell of claim 29, wherein the inducible promoter is a tet-responsive promoter, and wherein the additional construct encodes tTA or rtTA.
- 25 31. The cell of claim 29, wherein the inducible promoter is a LacO-responsive promoter, and wherein the additional construct encodes LacI.
32. The cell of claim 29, wherein the inducible promoter is a LoxP-stop-LoxP system promoter, and wherein the additional construct encodes Cre recombinase.

33. The cell of claim 32, wherein the Cre recombinase is under the transcriptional control of an inducible promoter, a developmental stage-specific promoter, or a tissue-specific promoter.
34. A non-human mammal comprising the cell according to any of claims 26-33.
- 5 35. The non-human mammal of claim 34, which is a chimeric mammal some of whose somatic or germ cells are a cell according to any of claims 26-33.
36. The non-human mammal of claim 34, which is a transgenic mammal all of whose somatic or germ cells are cells according to any of claims 26-33.
37. A method for inhibiting the expression of a target gene of interest in a cell,
10 comprising introducing a construct according to any of claims 1-25 into the cell, wherein the siRNA molecule derived from the precursor molecule is specific for the target gene.
38. The method of claim 37, further comprising inhibiting at least one additional target gene(s) of interest in the cell by introducing at least one additional
15 constructs according to any of claims 1-25 into the cell, wherein each of the siRNA molecules derived from the precursor molecules are specific for the additional target genes, respectively.
39. A method for treating a gene-mediated disease, comprising introducing into an individual having the disease a construct according to any of claims 1-25,
20 where the siRNA derived from the precursor molecule is specific for the gene mediating the disease.
40. A method of validating a gene as a potential target for treating a disease, comprising:
- 25 (1) introducing a construct according to any of claims 1-25 into a cell associated with the disease, wherein the siRNA molecule derived from the precursor molecule is specific for the gene;
- (2) assessing the effect of inhibiting the expression of the gene on one or more disease-associated phenotype;
- wherein a positive effect on at least one disease-associated phenotype is
30 indicative that the gene is a potential target for treating the disease.

41. The method of claim 40, wherein the gene is over-expressed or abnormally active in disease cells or tissues.
42. The method of claim 40 or 41, wherein the gene is downstream of and is activated by a second gene over-expressed or abnormally active in disease
5 cells or tissues.
43. The method of any of claims 40-42, wherein the product of the gene antagonizes an suppressor of a second gene over-expressed or abnormally active in disease cells or tissues.
44. The method of any of claims 40-43, wherein the cell is a tissue culture cell.
- 10 45. The method of claim 44, wherein the tissue culture cell is a primary cell isolated from diseased tissues, or from an established cell line derived from diseased tissues.
46. The method of any of claims 40-45, wherein the cell is within diseased tissues, and step (2) comprises evaluating one or more symptoms of the
15 disease.
47. The method of any of claims 40-46, wherein the expression of the gene is inducibly inhibited by a construct according to any of claims 1-25.
48. The method of any of claims 40-47, wherein the expression of the gene is inducibly activated by turning down the expression of a construct according
20 to any of claims 1-25.

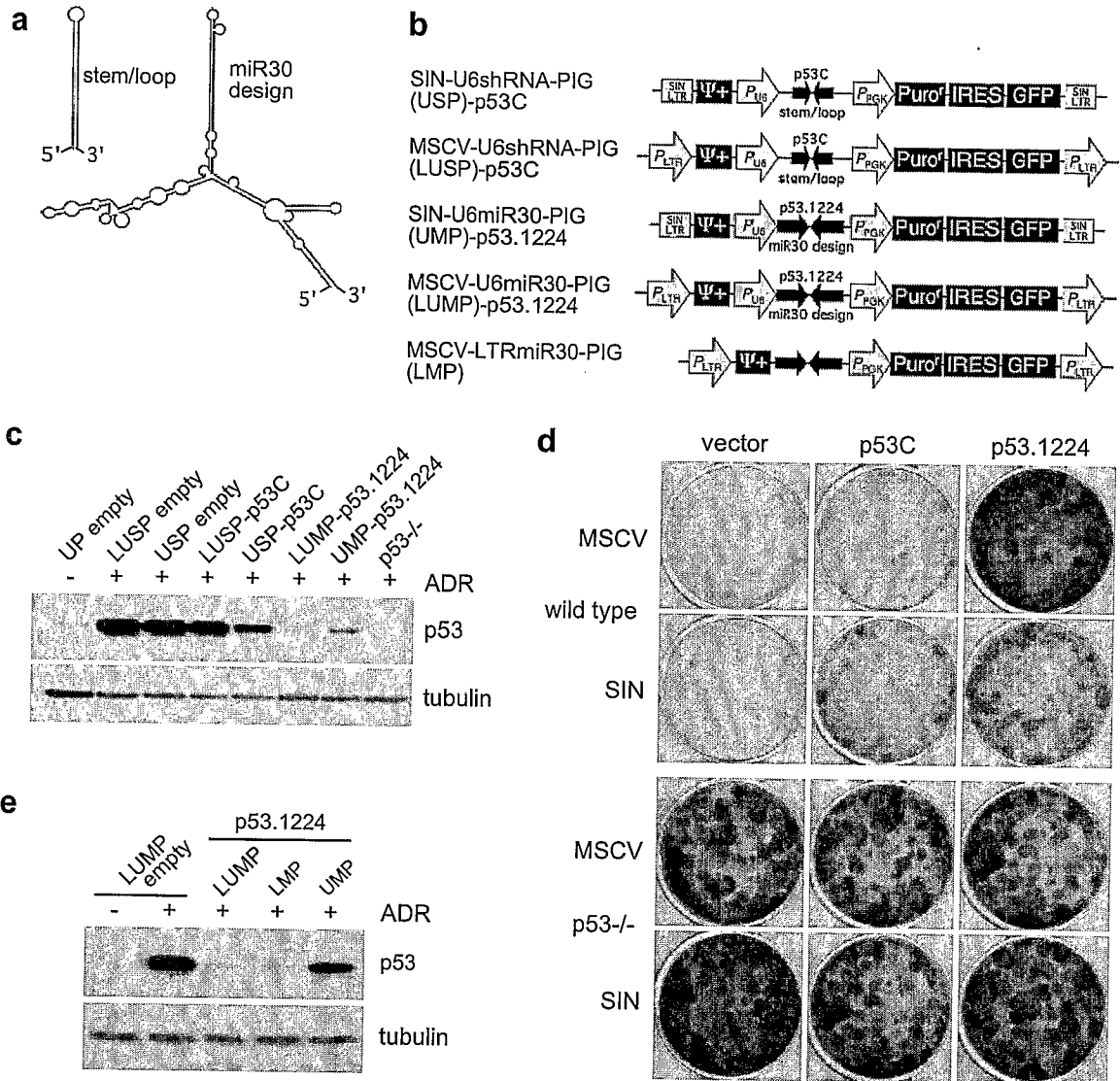


Figure 1

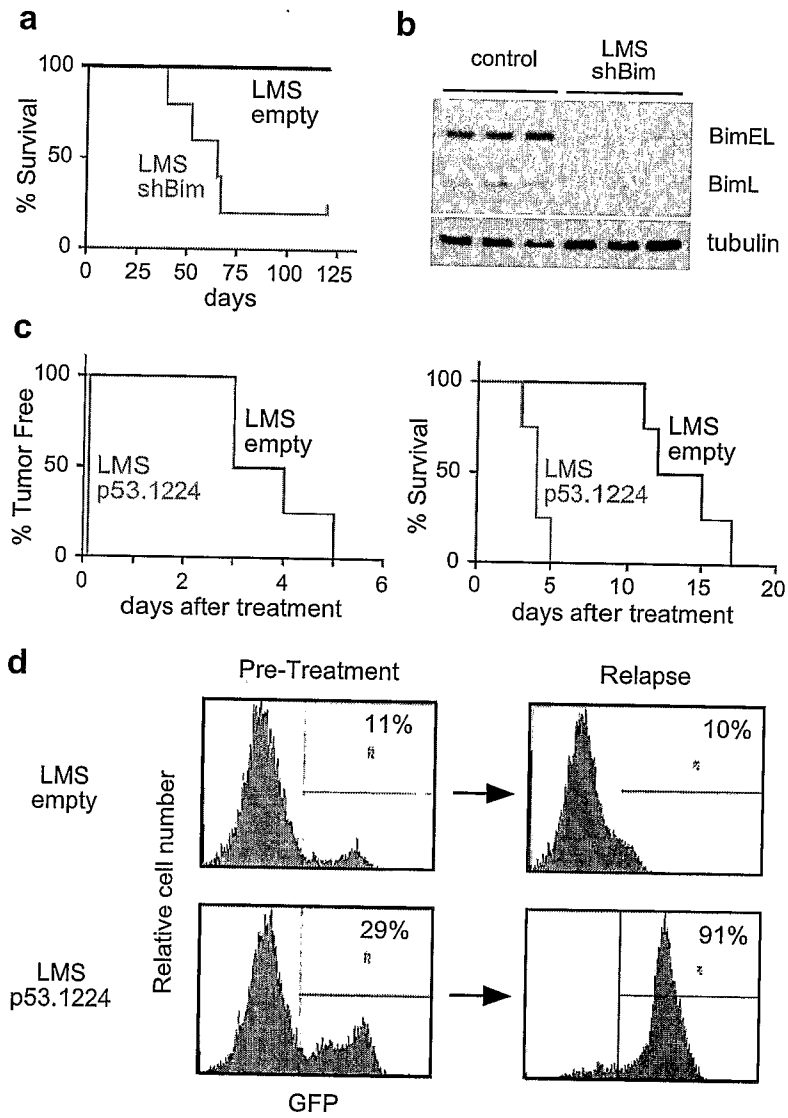


Figure 2

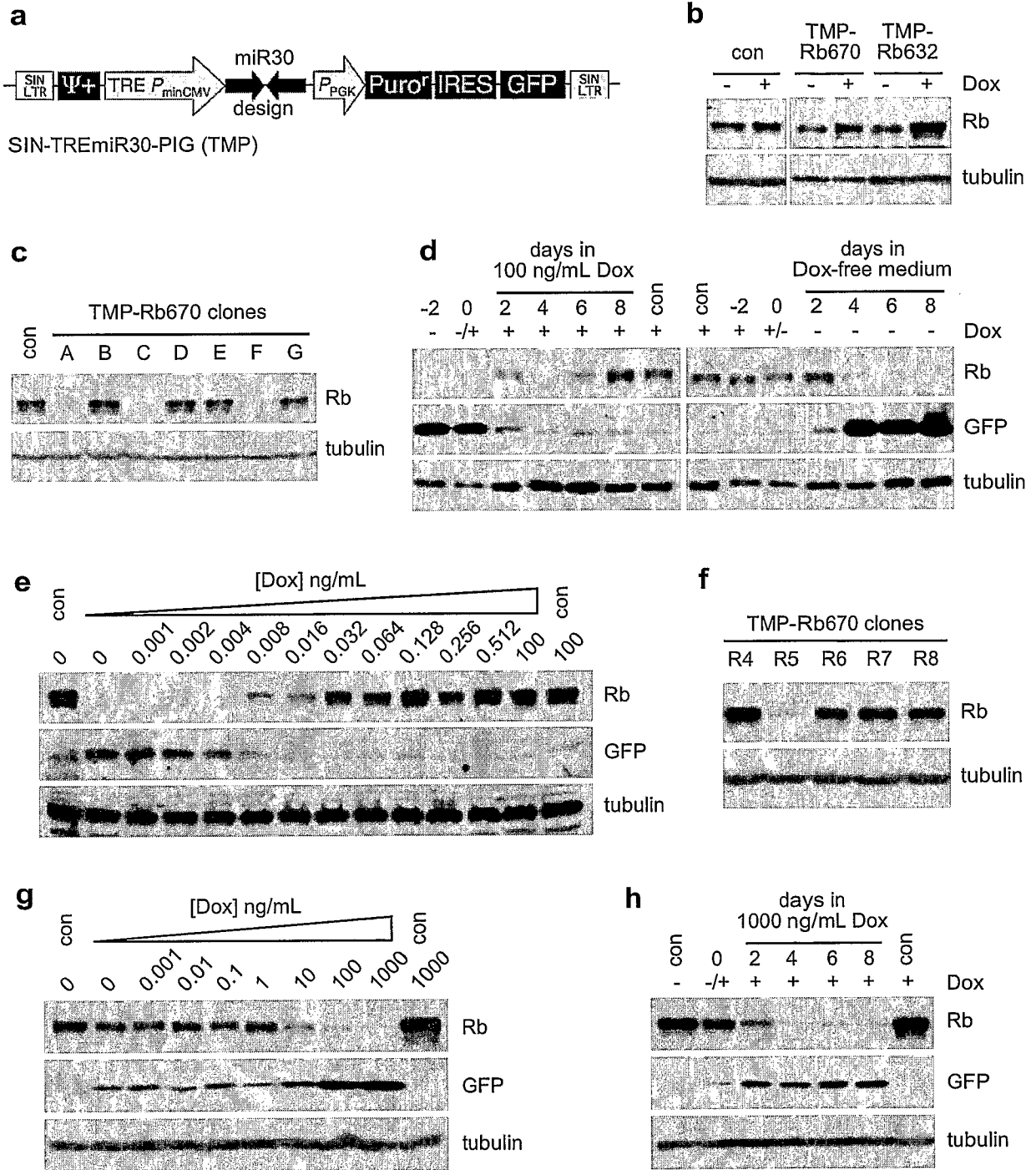


Figure 3

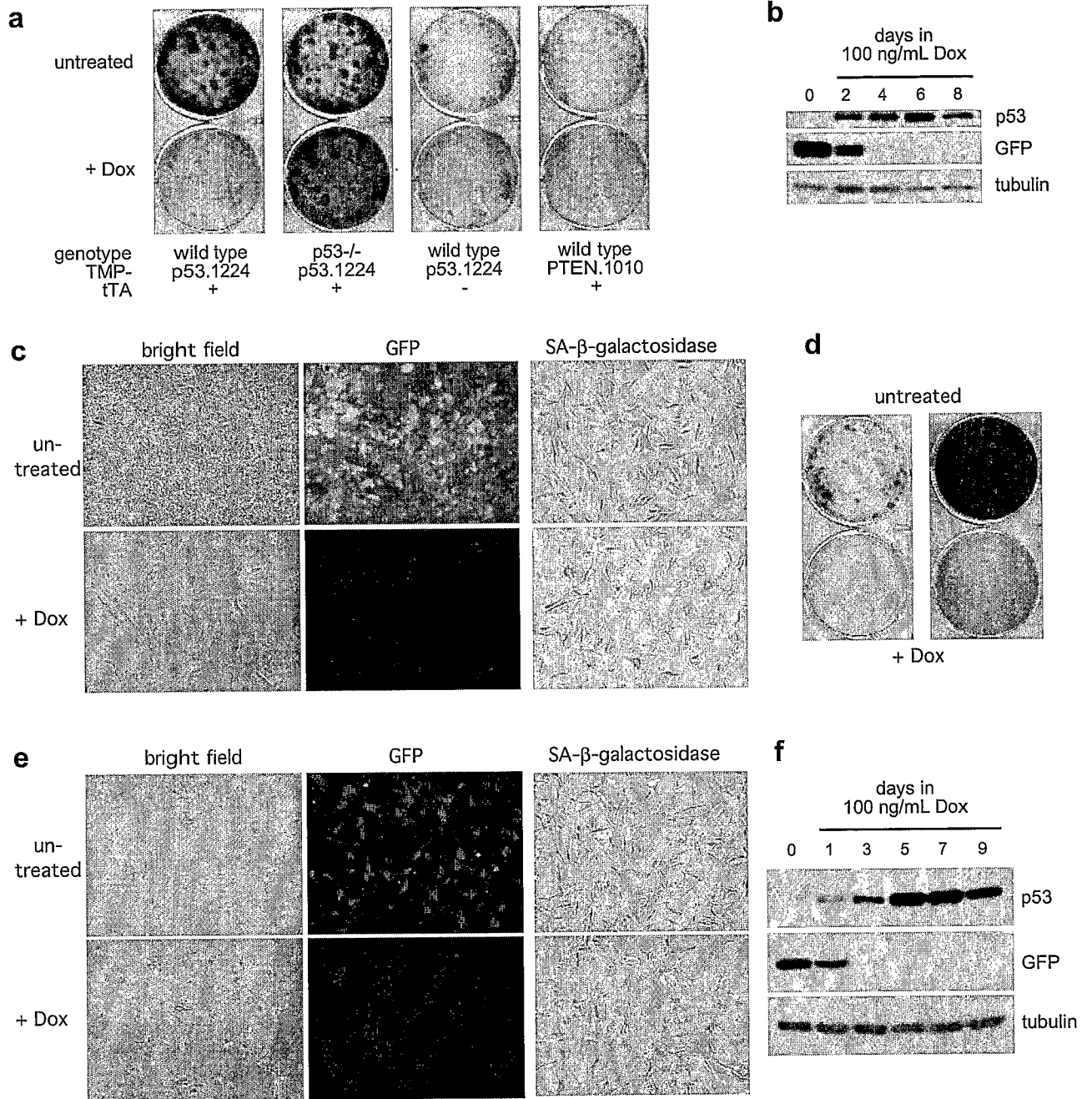


Figure 4

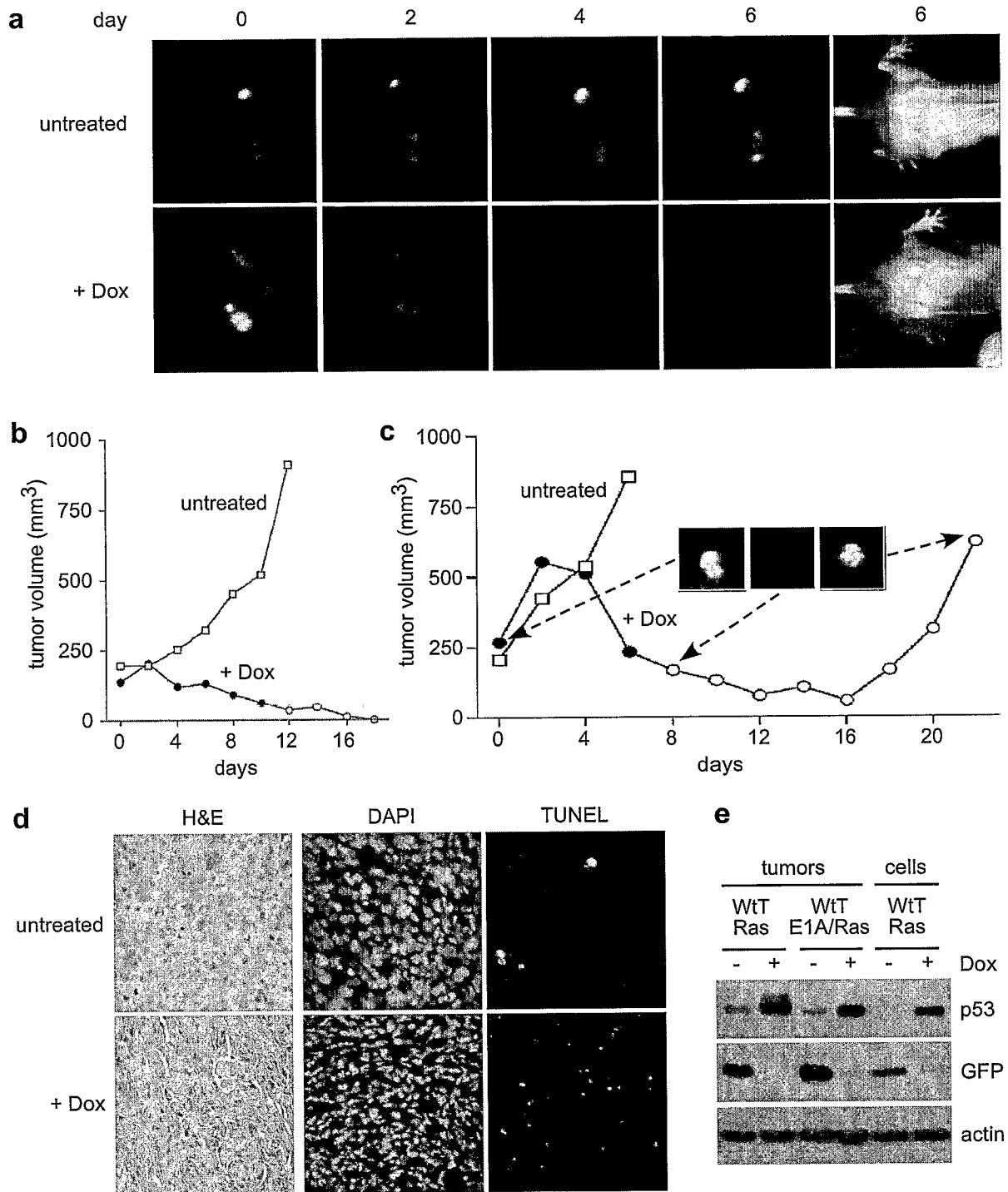


Figure 5

6/6

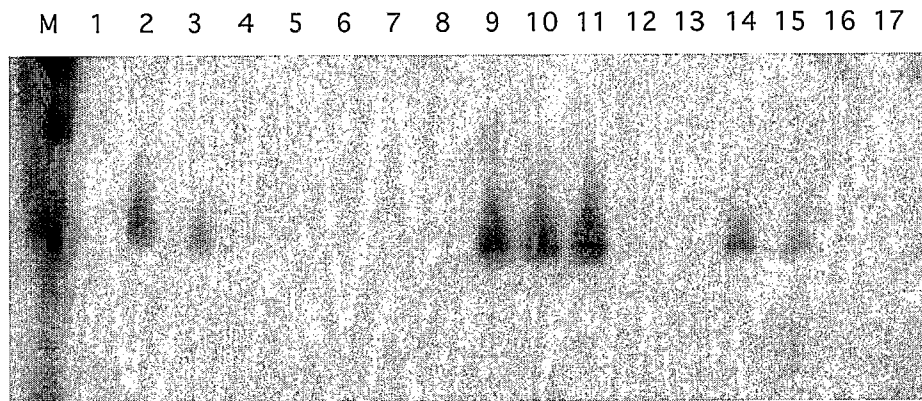


Figure 6