





(84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE,

SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Published:**

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

**Use of MicroRNA Precursors as Drugs for Inducing CD34-positive Adult Stem Cell Expansion**

**INVENTORS:** Shi-Lung Lin (Arcadia, CA), Donald Chang (Cerritos, CA) and David TS Wu (Arcadia, CA)

**PRIORITY:**

The present invention claims priority to the U.S. Provisional Application Serial No. 62/262,280 filed on December 2, 2015, which was entitled “miR-302 Attenuates A $\beta$ -induced Neurotoxicity through Activation of Akt Signaling”. Also, the present invention claims priority to the U.S. Patent Application Serial No. 15/048,964 filed on February 19, 2016, which was entitled “A Composition and Method of Using miR-302 Precursors as Drugs for Treating Alzheimer’s Diseases”. Furthermore, the present invention claims priority to the U.S. Patent Application Serial No. 15/167,226 filed on May 27, 2016, which was entitled “Use of MicroRNA Precursors as Drugs for Inducing CD34-positive Adult Stem Cell Expansion”.

**FIELD OF INVENTION:**

This invention generally relates to a composition and method for inducing CD34-cell expansion and/or regeneration with small hairpin RNA (shRNA) molecules, preferably microRNA precursor (pre-miRNA) and/or siRNA, which are useful for developing drugs/vaccines and/or therapies against a variety of ageing-related degenerative diseases in humans. Also, the present invention teaches the production and purification methods required for making high quality and high quantity of small hairpin-like RNA (shRNA) compositions such as microRNA precursors (pri-miRNA and/or pre-miRNA), and short interfering RNAs (siRNA), which are useful for treating human ageing-related diseases, such as, but not limited, Alzheimer’s diseases, Parkinson’s diseases, osteoporosis, diabetes, and cancers. The resulting shRNAs, preferably pre-miRNAs and siRNAs, are useful for developing therapeutic drugs against human degenerative diseases, particularly through a mechanism to induce CD34-positive stem cell expansion and/or regeneration. Furthermore, the present invention also reveals a novel pre-miRNA-based drug composition that is able to reprogram the malignant properties of high-grade liver cancers to a low-grade benign or even

relatively normal stage – a mechanism called “Cancer Reversion”. As cancer reversion is a totally new concept in drug designs, the present invention devises the first drug of its kinds using such a novel mechanism for cancer therapy.

**BACKGROUND:**

Stem cells are like a treasure box containing numerous effective ingredients useful for stimulating new cell growth/tissue regeneration, repairing and/or rejuvenating damaged/aged tissues, treating degenerative diseases, and preventing tumor/cancer formation/progression. Hence, it is conceivable that we can use these stem cells as a tool for novel drug screening, identification and production. As a result, the drugs so obtained will be useful for developing pharmaceutical and therapeutic applications, such as a biomedical utilization, device and/or apparatus for research, diagnosis, and/or therapy, and a combination thereof.

MicroRNA (miRNA) is one of the main effective ingredients in human embryonic stem cells (hESCs). Major hESC-specific miRNA species include, but not limited, members of the miR-302 family, miR-371~373 family, and miR-520 family. Among them, the miR-302 family has been found to play a functional role in tumor suppression (Lin et al., 2008 and 2010). MiR-302 contains eight (8) familial members, including four (4) sense miR-302 (a, b, c, and d) and four (4) antisense miR-302\* (a\*, b\*, c\*, and d\*). These sense and antisense members are partially matched and can form double-stranded duplex, respectively. Precursors of miR-302 are formed by miR-302a and a\* (pre-miR-302a), miR-302b and b\* (pre-miR-302b), miR-302c and c\* (miR-302c), and miR-302d and d\* (pre-miR-302d) with a link sequence in one end (stem loop), respectively. In order to activate miR-302 function, miR-302 precursors (pre-miR-302s) are first processed into mature miR-302s by cellular RNase III Dicers and further form RNA-induced silencing complexes (RISCs) with certain argonaute proteins, subsequently leading to either RNA interference (RNAi)-directed degradation or translational suppression of targeted gene transcripts (mRNAs), in particular oncogene mRNAs (Lin et al., 2008, 2010 and 2011).

MiR-302 is the most abundant ncRNA species found in hESCs and induced pluripotent stem cells (iPSCs). Our previous studies have shown that ectopic overexpression of miR-302 beyond the level found in hESCs is able to reprogram both human normal and cancerous cells to hESC-like iPSCs with a relatively slow cell cycle rate (20–24 hours/cycle) similar to that of a morula-stage early human zygote (Lin et al., 2008, 2010 and 2011; EP 2198025; U.S.

12/149,725; U.S. 12/318,806; U.S. 12/792,413). Relative quiescence is a defined characteristic of these miR-302-induced iPSCs, whereas hESCs and other previously reported four-factor-induced (either *Oct4-Sox2-Klf4-c-Myc* or *Oct4-Sox2-Nanog-Lin28*) iPSCs all showed a highly proliferative cell cycle rate (12–15 hours/cycle) similar to that of a tumor/cancer cell (Takahashi et al., 2006; Yu et al., 2007; Wernig et al., 2007; Wang et al., 2008). To disclose this tumor suppression effect of miR-302, we have identified the involvement of two miR-302-targeted G1-checkpoint regulators, including cyclin-dependent kinase 2 (CDK2) and cyclin D (Lin et al., 2010; U.S. 12/792,413; U.S. 13/964,705). It is known that cell cycle progression is driven by activities of cyclin-dependent kinases (CDKs), which forms functional complexes with positive regulatory subunits, cyclins, as well as by negative regulators, CDK inhibitors (CKIs, such as p14/p19Arf, p15Ink4b, p16Ink4a, p18Ink4c, p21Cip1/Waf1, and p27Kip1). In mammals, different cyclin-CDK complexes are involved in regulating different cell cycle transitions, such as cyclin-D-CDK4/6 for G1-phase progression, cyclin-E-CDK2 for G1-S transition, cyclin-A-CDK2 for S-phase progression, and cyclin-A/B-CDC2 (cyclin-A/B-CDK1) for entry into M-phase. Hence, our studies suggested that the tumor suppression function of miR-302 results from co-suppression of the cyclin-E-CDK2 and cyclin-D-CDK4/6 pathways during G1-S transition.

Although miR-302 is useful for designing and developing novel anti-cancer drugs/vaccines, its production is problematic because natural miR-302 can only be found in human pluripotent stem cells such as hESCs, of which the resource is very limited. Alternatively, synthetic small interfering RNAs (siRNA) may be used to mimic pre-miR-302; yet, since the structure of a pre-miR-302 is formed by two mis-matched strands of miR-302 and miR-302\*, those perfectly matched siRNA mimics can not replace the function of miR-302\*, of which the sequence is totally different from the antisense strand of siRNA. For example, the antisense strand of siRNA-302a mimic is 5'-UCACCAAAAC AUGGAAGCAC UUA-3' (SEQ.ID.NO.1), whereas native miR-302a\* is 5'-ACUAAAACGU GGAUGUACUU GCU-3' (SEQ.ID.NO.2). As miR-302 function results from both of its sense miR-302 and antisense miR-302\* strands, previous reports using those siRNA mimics have shown different results from native miR-302 function. On the other hand, our recent discovery of iPSCs may provide an alternative solution for pre-miR-302 production (EP 2198025; U.S. 12/149,725; U.S. 12/318,806). Nevertheless, the cost of growing these iPSCs is still too high to be used for industrial production.

Alternatively, the use of prokaryotic competent cells may be a possible approach for producing human microRNAs and their precursors. However, prokaryotic cells lack several essential enzymes required for eukaryotic microRNA expression and processing, such as Drosha and Dicer. Also, prokaryotic RNA polymerases do not efficiently transcribe small RNAs with high secondary structures, such as hairpin-like pre-miRNAs and shRNAs. In fact, there is no true microRNA encoded in bacterial genomes and bacteria do not naturally express microRNA. As a result, if we can force the expression of human microRNAs in prokaryotes, the resulting microRNAs will most likely remain in their precursor conformations similar to pri-miRNA (a large primary cluster of multiple pre-miRNAs) and/or pre-miRNA (one single hairpin RNA). Despite all of the above problems, the real challenge is how to force the expression of human microRNAs in prokaryotes. To overcome this major problem, our priority application U.S. No. 13/572,263 has established a preliminary method; yet, it is currently not sure whether these prokaryote-produced microRNAs (pro-miRNA) will possess the same structures and functions as their human counterparts. Also, the pro-miRNAs so obtained may be contaminated with bacterial endotoxin, which is not suitable for direct use in therapy.

As learning from current textbooks, every ordinary skill person in the art knows very well that prokaryotic and eukaryotic transcription machineries are different and hence not compatible to each other. For example, based on current understandings, eukaryotic RNA polymerases do not bind directly to a promoter sequence and require additional accessory proteins (cofactors) to initiate transcription, whereas prokaryotic RNA polymerases form a holoenzyme that binds directly to a promoter sequence to start transcription. It is also a common sense that eukaryotic messenger RNA (mRNA) is synthesized in the nucleus by type-II RNA polymerases (pol-2) and then processed and exported to the cytoplasm for protein synthesis, whereas prokaryotic RNA transcription and protein translation take place simultaneously off the same piece of DNA in the same place. This is because prokaryotes such as bacteria and archaea do not have any nucleus-like structure. Accordingly, these differences make a prokaryotic cell difficult or even impossible to produce eukaryotic RNAs using eukaryotic promoters.

Prior arts attempt at producing mammalian peptides and/or proteins in bacterial cells, such as U.S. Pat. No. 7,959,926 to Buechler and U.S. Pat. No. 7,968,311 to Mehta, used bacterial or bacteriophage promoters. For initiating expression, a desired gene was cloned

into a plasmid vector driven by a bacterial or bacteriophage promoter. The gene must not contain any non-coding intron because bacteria do not have any RNA splicing machinery to process the intron. Then, the vector so obtained was introduced into a competent strain of bacterial cells, such as *Escherichia coli* (*E. coli*), for expressing the transcripts (mRNAs) of the gene and subsequently translating the mRNAs into proteins. Nevertheless, the bacterial and bacteriophage promoters, such as Tac, Lac, T3, T7, and SP6 RNA promoters, are not pol-2 promoters and their transcription activities tend to be an error-prone process which causes mutations. In addition, Mehta further taught that glycerol/glycerin might be used to increase the efficiency of bacterial transformation; yet, no teaching was related to enhancement of RNA transcription, in particular pol-2 promoter-driven prokaryotic RNA transcription. Due to lack of possible compatibility between eukaryotic and prokaryotic transcription systems, these prior arts were still limited by the use of prokaryotic RNA promoters for gene expression in prokaryotes.

Due to the problems of system incompatibility and possible endotoxin contamination, there was previously no means for producing human pre-miRNA/shRNA-like drugs in prokaryotes. Also, a pre-miRNA/shRNA is sized about 70~85-nucleotides in length which is too large and costly to be made by a RNA synthesis machine. To overcome these problems, the present invention provides a novel breakthrough – By adding some defined chemicals mimicking certain transcriptional cofactors, we can create a novel adaptation environment for prokaryotic cells to use eukaryotic pol-2 and/or pol-2-like promoters for transcribing desired pre-miRNAs and shRNAs without going through error-prone prokaryotic promoters. The advantages are: first, cost-effective mass production due to the fast growth of bacteria; second, easy handling because of no need for growing dedicate hESCs or iPSCs; third, high fidelity productivity in terms of pol-2 promoter-driven RNA transcription; fourth, high purity of desired microRNAs due to lack of true microRNA in prokaryotes; and last, no endotoxin, which can be further removed by certain chemical treatments. Therefore, a method for producing human pre-miRNAs and/or shRNAs in prokaryotic cells without the problems of system incompatibility and endotoxin contamination is highly desirable. Furthermore, the drugs so obtained may present novel therapeutic effects other than the currently known function of a microRNA.

**SUMMARY OF THE INVENTION:**

The principle of the present invention is relied on the different and incompatible properties between prokaryotic and eukaryotic RNA transcription systems. Naturally, prokaryotic RNA polymerases do not recognize eukaryotic promoters and vice versa. However, the present invention has identified chemical agents that can serve as transcriptional inducers to trigger and/or enhance eukaryotic promoter-driven RNA transcription in prokaryotes. Hence, the knowledge taught in the present invention is a totally novel breakthrough beyond all current understandings regarding the differences between prokaryotic and eukaryotic transcription systems.

The present invention is related to an inducible gene expression composition using certain chemical inducers to stimulate and/or enhance eukaryotic promoter-driven RNA transcription in prokaryotes. These chemical inducers have not been used in a cell culture medium due to their bacteriostatic and/or bactericidal properties, including 3-morpholinopropane-1-sulfonic acid [or named 3-(N-morpholino)propanesulfonic acid; MOPS], glycerin and ethanol, as well as their functional analogs such as 2-(N-morpholino)ethanesulfonic acid (MES), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and mannitol. Conceivably, chemicals with a similar structure like these transcriptional inducers may share a similar function. For example, MOPS is frequently used as a buffering agent in bacterial cell lysis and hence is not suitable for growing bacteria. On the other hand, ethanol is a well-known sanitizer and glycerin is frequently used in bacterial transformation by destabilizing the bacterial cell walls, indicating that glycerin is bacteriostatic and ethanol is bactericidal, respectively. In view of these known functionalities of MOPS, ethanol and glycerin, an ordinary skill in the art would not anticipate the use of a trace amount (0.001% to 4% volume/volume concentration) of these chemicals for inducing eukaryotic promoter-driven gene expression in prokaryotic cells without first knowing the knowledge of the present invention.

Based on the above knowledge, the present invention is a design and method for utilizing prokaryotic cells to produce human microRNA precursors (pre-miRNAs) and/or shRNAs as therapeutic drugs and/or vaccines for cancer therapy. More specifically, the present invention is a design and method of utilizing prokaryotic cells to produce a special kind of pre-miRNA-like agents, named pro-miRNA, that are capable of reprogramming the malignant properties of high-grade human cancer cells into a low-grade benign or even

relatively normal-like state. Preferably, these pro-miRNAs are tumor suppressor microRNAs (TS-miRNA) similar to the precursors of miR-302a, b, c, d, e, and/or f (pre-miR-302s) and their natural familial cluster as well as their manually re-designed small hairpin RNA (shRNA) homologues/derivatives, and/or a combination thereof. The designs of pro-miRNA-like shRNA homologues/derivatives include imperfectly and perfectly matched hairpin compositions of the pro-miRNA and its homologous small interfering RNA (siRNA), which may be formed in a single unit or in a multiple unit cluster. These designs may improve the target specificity and/or reduce the copy number of pro-miR-302 required for effective delivery and therapy. The human cells suitable for such a drug treatment include normal, tumor, and cancerous cells in vitro, ex vivo and/or in vivo.

Preferably, the prokaryotic cells used for the present invention are bacterial competent cells in particular, *Escherichia coli* (E. coli), and the chemical inducer is MOPS, ethanol, or glycerin, or a mixture thereof. Also preferably, the eukaryotic RNA promoter used is either a eukaryotic pol-2 promoter (*i.e.* *EF1alpha* promoter) or a pol-2 compatible (pol-2-like) viral promoter (*i.e.* cytomegaloviral CMV promoter). The gene mediated by the eukaryotic RNA promoter may code for either a non-coding or protein-coding RNA, or both (such as an intron-containing gene transcript), selected from the group consisted of microRNA (miRNA), small hairpin RNA (shRNA), small interfering RNA (siRNA), messenger RNA (mRNA), their precursors and homologues, and a combination thereof. For inducing gene expression, the prokaryotic cells are transfected with the eukaryotic RNA promoter-mediated gene and then grown in a culture medium similar to Luria-Bertani (LB) broth at 37°C with addition of the chemical inducer(s) for >24 hours.

To demonstrate the inducibility of said chemical inducers for human microRNA production in prokaryotes, we modified a lentiviral vector *pSpRNAi-RGFP-miR302* from our priority U.S. Patent Applications No. 12/149,725 and No. 12/318,806 to a new plasmid vector *pLenti-EF1a-RGFP-miR302*, of which the *SpRNAi-RGFP* gene expression is driven by a eukaryotic promoter such as *EF1alpha* or *CMV* (FIG. 1A). After that, we transformed *E. coli* competent cells with it and then used the production of red fluorescent protein (RGFP) as a visible marker for measuring the transcription rate and process of microRNA miR-302, as shown in FIG. 1B. Because the miR-302 familial cluster was also modified to be encoded in the 5'-intron region [*e.g.* 5'-untranslated region (5'-UTR) or the first intron] of the *RGFP* gene, the transcription of each *RGFP* mRNA led to the production of one 4-hairpin miR-302

precursor cluster (pri-miR-302) and/or four 1-hairpin miR-302 precursors (pre-miR-302s), as shown in FIGs. 5 and 6. Due to lack of RNase III Dicer in prokaryotes, the pri-miR-302 transcripts would be eventually broke down (by certain single-strand RNases in *E. coli*) into 1-hairpin pre-miR-302s, all of which could be extracted and further used as therapeutic drugs in the present invention. Broadly speaking, 5'-UTR and 3'-UTR are considered as a part of intron in the present invention.

All miR-302 members share a totally identical sequence in their first 5'-seventeen (17) nucleotides 5'-UAAGUGCUUC CAUGUUU-3' (SEQ.ID.NO.3), and contain >82% homology in their full-length 23-nucleotides of a mature microRNA. Based on the results predicted by online computing programs TARGETSCAN (<http://www.targetscan.org/>) and PICTAR-VERT (<http://pictar.mdc-berlin.de/>), these miR-302s concurrently target against almost the same genes, including >600 human genes. In addition, miR-302 also shares many overlapping target genes with mir-92, mir-93, mir-200c, mir-367, mir-371, mir-372, mir-373, mir-374, and mir-520 familial members, all of which may possess similar functions. Most of these target genes are developmental signals and transcriptional factors involved in initiating and/or establishing certain lineage-specific cell differentiation during early embryogenesis (Lin et al., 2008). Many of these target genes are also well-known oncogenes; as a result, miR-302s likely functions as a tumor suppressor to prevent the deviation of normal hESC growth into tumor/cancer formation.

#### **Induction of Eukaryotic Promoter-driven Gene Expression in Prokaryotes.**

*Escherichia coli* (*E. coli*) competent cells were transformed by the *pLenti-EF1alpha-RGFP-miR302* plasmid (FIG. 1A) using a z-competent *E. coli* transformation kit (Zymo Research, Irvine, CA) and cultivated in Luria-Bertani (LB) broth supplemented with a mixture of 0.1% (v/v) MOPS and 0.05% (v/v) glycerin (inducers) at 37°C with frequent agitation at 170 rpm. After overnight incubation, the transformed *E. coli* competent cells expressed highly abundant red RGFP proteins that could be clearly seen in the color of the LB broth, whereas the blank control *E. coli* presented no RGFP, as shown in FIG. 2. The presence of functional RGFP indicated that both of its encoded RNA and protein are successfully produced and processed in the competent cells.

To further confirm the specificity of gene expression induced by the chemical inducers, two transformed *E. coli* strains were prepared: one carried a *pLVX-Grn-miR302+367* plasmid vector containing a *CMV* promoter-driven green fluorescent protein (GFP) gene and the other

carried the aforementioned *pLenti-EF1alpha-RGFP-miR302* vector. After overnight incubation with only 0.1% (v/v) MOPS, the *E. coli* transformed with *pLVX-Grn-miR302+367* were changed to green color while the other with *pLenti-EF1alpha-RGFP-miR302* still showed red color, as shown in FIG. 3. This result indicates that the chemical inducers like MOPS can stimulate a specific RNA transcription and its related protein production through either a eukaryotic pol-2 or a pol-2-like viral promoter. Particularly it was noted that the RGFP and GFP production is so abundant that even the *E. coli* cells are visually stained by the respective red and green colors.

Among all chemicals tested in the present invention, the top three most potent inducers are MOPS, glycerin and ethanol, as shown in FIG. 4. The quantitative result of the induced RGFP production was further confirmed by Western blot analysis, as shown in FIG. 5 and Example 3. Bacterial RuvB protein was served as a house-keeping standard to normalize RGFP expression. The inducibility of these identified inducers was also found to be dose-dependent in proportion to their concentrations. Without any treatment, negative control *E. coli* cells just showed their original color in absence of any fluorescent stain. Therefore, according to all these results, the present invention clearly provides a novel chemical-inducible composition and its application for modulating eukaryotic pol-2-driven or pol-2-like viral promoter-driven RNA production in prokaryotic cells. In view of the above demonstration, it is very obvious for an ordinary skill in the art to use other genes or the related cDNAs in place of the *RGFP* gene for producing functional RNAs and the related proteins in prokaryotes.

#### **Induction of Eukaryotic Promoter-driven microRNA Expression in Prokaryotes.**

Accompanying the experiments of RGFP induction shown above, we further measured the expression of pri-/pre-miR-302s and their mature miR-302s in the *pLenti-EF1alpha-RGFP-miR302*-transformed cells with or without chemical induction. As shown in FIG. 6 and Example 4, the quantitative results of induced pri-/pre-miR-302 production have been confirmed by Northern blot analysis. Similar to the results of the RGFP induction in FIGs. 4 and 5, the pri-/pre-miR-302 expression was strongly detected in transformed cells treated with MOPS, glycerin or ethanol, but not blank control, indicating that these chemical inducers indeed stimulated the expression of the encoded pri-/pre-miRNAs in prokaryotic cells through a eukaryotic pol-2 promoter (FIG. 6). Due to the structural similarity of pre-miRNAs and shRNAs, it is obvious for an ordinary skill in the art to use the present invention to produce

other kinds of pri-/pre-miRNA species, such as but not limited miR-34, miR-146, miR-371~373 and miR-520. For clarification, these prokaryote-produced pri-/pre-miRNAs are called pro-miRNAs.

Since *pLenti-EF1alpha-RGFP-miR302* contains a miR-302 familial cluster located in the 5'-UTR of the *RGFP* gene (FIGs. 1A and 1B), the induced *RGFP* gene expression will also generate the miR-302 cluster (pri-miR-302) and its derivative pre-miR-302a, b, c and d (pre-miR-302s) as demonstrated in FIG. 1B. Due to lack of RNase III Dicer in prokaryotes, the pri-miR-302 and pre-miR-302s so obtained were found to remain as hairpin-like microRNA precursors, which are useful for developing therapeutic drugs. In human cells, these pre-miR-302s and pri-miR-302 can be processed into mature miR-302 for eliciting its tumor suppression function. Similarly, the present invention can also be used to produce other kinds of TS-miRNA species and their precursors, such as the miR-34a, miR-146a, miR-373 and miR-520 family.

The resulting pro-miRNAs can be easily extracted from competent *E. coli* cells (Examples 5 and 6) and further purified by high-performance liquid chromatography (HPLC) (FIGs. 10A and 10B). Within the purified pro-miR-302s, we have identified all of the miR-302 familial members (miR-302a, a\*, b, b\*, c, c\*, d, and d\*) using analyses of microRNA microarrays (FIGs. 11B and 12) and RNA sequencing [FIGs. 13A (pri-miR-302) and 13B (pre-miR-302s)]. Particularly, the sequencing results showed that these pro-miR-302s all share exactly the same sequences as their natural pre-miR-302 counterparts (FIG. 13B). Furthermore, we have formulated these pro-miR-302s into a soluble drug for IV/in-vivo injection in order to test their therapeutic effects on human liver cancers in vivo (Example 11). As shown in FIG.14, after 3 injection treatments, the pro-miR-302 drug successfully reduced >90% volume of the engrafted human liver cancers in vivo, shirking the average cancer size to <10% compared to the untreated cancers. Moreover, histological examination with hematoxylin & eosin (H&E) staining further demonstrated that this significantly therapeutic effect was resulted from not only the reported tumor suppression function of miR-302 (Lin et al., 2010) but also another novel reprogramming function that has not yet been observed before. For instance, FIG.15 clearly showed that the pro-miR-302 drug can reprogram the malignant properties of high-grade human liver cancers in vivo to a much more benign stage almost similar to that of normal liver tissues! These treated cancers can even form normal liver-like structures, such as classical liver lobules, central veins (CV) and portal

triads (PT). Therefore, these evidences strongly indicated that pro-miR-302 is able to not only inhibit tumor/cancer cell growth but also reset the malignancy of human cancers to a relatively benign or normal state in vivo, leading to a totally novel therapeutic effect for cancer drug design.

In the present invention, both of the plasmid vector and its encoded non-coding RNAs (*i.e.* pre-miRNA/shRNA) can be simultaneously amplified in the prokaryotic cells, preferably *E. coli DH5alpha* competent cells (Examples 1, 5 and 6). The method for isolating the amplified *pLenti-EF1alpha-RGFP-miR302* plasmid DNA and the transcribed pri-/pre-miR-302s is described in Examples 5 and 6. The technology for delivering plasmid vectors (*i.e.* *pLenti-EF1alpha-RGFP-miR302*) into prokaryotic cells is called cell transformation, while the method for delivering the amplified non-coding RNAs (*i.e.* pro-/pri-/pre-miR-302s) into eukaryotic cells can be selected from the group of endocytosis, chemical/glycerol infusion, peptide/liposomal/chemical-mediated transfection, electroporation, gene gun penetration, micro-injection, transposon/retrotransposon insertion and/or adenoviral/retroviral/lentiviral infection.

#### **Pre-miR-302-induced Pluripotent Stem Cell Derivation.**

MiR-302 has been reported to reprogram mammalian somatic cells to human embryonic stem cell (hESC)-like induced pluripotent stem cells (iPSCs) as demonstrated in our priority U.S. Patent Applications No. 12/149,725 and No. 12/318,806. Numerous stem cell applications and therapies have been designed and developed using these iPSCs. Nevertheless, since cultivating these iPSCs and hESCs is very costly and laborious, it is difficult and inefficient to collect miR-302 and its precursors from these pluripotent stem cells. On the other hand, making synthetic shRNA mimics is another possible alternative for pre-miR-302 production; yet, the cost is still very expensive. Also, the similarity between synthetic shRNA and natural pre-miR-302 is very questionable. To solve these problems, the present invention provides a simple, cheap and efficient method for mass production of pre-miR-302 in prokaryotes. Moreover, the extraction and purification of these prokaryote-produced pre-miR-302s (pro-miR-302s) is relatively easy and cost-effective, as shown in FIG. 6 and Example 6 of the present invention.

We have used the *pLenti-EF1alpha-RGFP-miR302*-transformed *E. coli* cells to produce and isolate high quantity and quality of *pLenti-EF1alpha-RGFP-miR302* vector and pro-miR-302s, as shown in Examples 5 and 6. Both *pLenti-EF1alpha-RGFP-miR302* and pro-miR-

302s are useful for generating iPSCs. Following Example 2, when the pro-miR-302s produced by the present invention were transduced into human skin primary keratinocytes, the transfected keratinocytes were reprogrammed to hESC-like iPSCs that expressed strong hESC marker Oct4 (FIG. 7). In FIG. 8 and Example 8, we further performed bisulfite DNA sequencing assays to show that global DNA demethylation did occur in the promoters of both *Oct4* and *Sox2* genes, two of the key reprogramming factors as well as hESC markers. As global DNA demethylation and Oct4 expression are known to be the first step of somatic cell reprogramming to form hESC-like iPSCs (Simonsson and Gurdon, *Nat Cell Biol.* 6: 984-990, 2004), the pro-miR-302s isolated from the MOPS-induced *E. coli* cell extracts is proven to be as effective as natural pre-miR-302s, which are useful for iPSC derivation. Hence, pro-miR-302 and pre-miR-302 likely possess the same function in stem cell induction.

**Pre-miR-302-induced CD34-positive Adult Stem Cell Expansion and/or Regeneration.**

MicroRNA miR-302 has been found to reprogram mammalian somatic cells to embryonic stem cell (ESC)-like induced pluripotent stem cells (iPSC) (Lin, 2008, 2010, 2011; U.S. Patent Applications No. 12/149,725 and No. 12/318,806 to Lin). Using these iPSCs, many stem cell-associated applications and therapies have been developed for advancing modern regenerative medicine. Yet, miR-302 is only abundantly found in human ESCs rather than differentiated tissue cells. Also, isolation of miR-302 from human ESCs is highly debatable, costly and tedious. To solve these problems, the present invention provides a simple, cheap, fast and inducible composition and method for mass production of hairpin-like miR-302 molecules and/or their precursors/homologs in prokaryotes. Moreover, the isolation of miR-302 and/or its precursors from prokaryotic cells is relatively easy and cost-effective, as shown in FIG. 6 and Example 6 of the present invention.

We have used the *pLenti-EF1alpha-RGFP-miR302*-transformed *E. coli* cells to produce and isolate high quantity and quality of the *pLenti-EF1alpha-RGFP-miR302* vector and pre-miR-302s, as shown in Examples 5 and 6. The use of *pLenti-EF1alpha-RGFP-miR302* had been shown to produce human ESC-like iPSCs in view of our previous U.S. Patent Applications No. 12/149,725 and No. 12/318,806. Also, the iPSCs so obtained can be further differentiated into various somatic tissue cells as demonstrated in our previous studies (Lin et al, 2008, 2010 and 2011).

The applications of isolated miR-302 and/or pre-miR-302 molecules may further include the induction and expansion of CD34-positive adult stem cells. As shown in FIGS.

17A and 17B, our recent studies in wound healing therapy and cancer therapy using a novel miR-302-formulated drug revealed that treatments of relatively low concentrations (50~500 $\mu$ g/mL) of the isolated miR-302/pre-miR-302 molecules not only greatly enhance scar-less wound healing but also induce CD34-positive adult stem cell expansion around the wounded area in pig skins in vivo. Based on the miR-302-treated (miR-302s/pre-miR-302s+antibiotic ointment) result of FIG. 17B in comparison with that of control (only antibiotic ointment) result of FIG. 17A, it clearly showed a  $\geq 40$ -fold increase of CD34-positive adult stem cell populations (labeled by green fluorescent antibodies) in vivo after miR-302 treatments. The currently known CD34-positive adult stem cell types include, but not limited, skin, hair, muscle, blood (hematopoietic), mesenchymal, and neural stem cells. As a result, since miR-302 can be used to induce CD34-positive adult stem cell expansion and/or regeneration in vivo, this therapeutic effect may also help to re-grow and/or revive functional adult stem cells for treating degenerative diseases in humans, such as, but not limited, Alzheimer's diseases, Parkinson's diseases, osteoporosis, diabetes, and cancers.

#### **Utilization of pro-miR-302 for Tumor/Cancer Therapy in vivo.**

Our previous studies have demonstrated the feasibility of this approach in treating human hepatocellular carcinoma HepG2 cells in vitro but not in vivo (Lin et al., 2010). As shown in FIG. 9, the treated tumor/cancer cells were reprogrammed to iPSCs (labeled as mirPS-HepG2) and formed embryoid body-like cell colonies. Moreover, miR-302 was also found to induce >95% apoptosis in the treated cancer cell population. The top panels of FIG. 9 further showed that flow cytometry analysis of the DNA content in response to cell cycle stages revealed a significant reduction in the mitotic cell populations after miR-302 treatments (from 45.6% to 17.2%). These results indicated that miR-302 can effectively attenuate the fast cell cycle rate of human liver cancer cells and hence causes significant apoptosis in these cancer cells.

The process of cancer progression was thought to be irreversible due to accumulative gene mutations; yet, the present invention discloses a novel pre-miRNA (pro-miR-302) function that can reprogram high-grade malignant cancers back to a low-grade benign or even normal-like stage in vivo, of which the mechanism may be related to a very rare natural healing process called spontaneous cancer regression. Spontaneous cancer regression occurs rarely at a rate of less than 1 in 100,000 cancer patients. We found that pro-mir-302 treatment is able to increase this rare healing rate to >90% in human liver cancers. As

shown in FIG. 14, the therapeutic results of using pro-miR-302s as a drug to treat human liver cancer xenografts in SCID-beige nude mice ( $n = 6$ ) demonstrated that this pro-mir-302 drug successfully reduced cancer sizes from  $728 \pm 328 \text{ mm}^3$  (untreated blank control, C) to  $75 \pm 15 \text{ mm}^3$  (pro-mir-302-treated, T), indicating a ~90% reduction rate in the average cancer size, whereas treatments of other synthetic siRNA mimics (siRNA-302) did not provide any similar therapeutic effect.

Further histological examination (the most right panels of FIG. 14) showed that normal liver lobule-like structures (circled and pointed by a black arrow) were observed only in pro-miR-302-treated cancer grafts but not other treatments or controls, suggesting that a reprogramming mechanism has occurred to reset the malignant cancer cell property back to a relatively normal-like state (Cancer Reversion). This novel reprogramming mechanism is likely resulted from the gene silencing effect of miR-302 on human oncogenes in particular, those mutated oncogenes involved in cancer progression. By silencing those mutated oncogenes, pro-miR-302 is able to reset the cancerous gene expression patterns back to a normal-like state, consequently leading to the therapeutic result of cancer reversion. Nevertheless, this in-vivo reprogramming mechanism is clearly different from the previously reported somatic cell reprogramming (Lin et al., 2008 and 2011) because no Oct4-positive pluripotent stem cell has been identified.

More detailed histological examination (FIG. 15) further confirmed that the pro-miR-302 drug did reprogram high-grade (Grade IV) human liver cancer grafts to a more benign low-grade (less than Grade II) state. As shown in FIG. 15, the treated cancer grafts formed classical liver lobules containing central vein (CV)-like and portal triad (PT)-like structures (indicated by black arrows), highly similar to normal liver tissue structures (top). Histological comparison among untreated, siRNA-treated, pro-miR-302-treated human liver cancer grafts and normal liver tissues in vivo (FIG. 16) also showed that the engrafted human liver cancers without treatment (top) aggressively invaded into surrounding normal tissues, such as muscles and blood vessels, and formed massive cell-cell and cancer-tissue fusion structures, demonstrating its high malignancy and metastasis. Treatment of siRNA mimics (siRNA-302) did not significantly reduce the malignancy of the engrafted liver cancers (upper middle), probably due to the short half-life of siRNA in vivo. In contrast, treatment of pro-miR-302 not only reprogrammed the engrafted cancer cells to a normal liver cell-like morphology (no fusion) but also successfully inhibited any cancer invasion into the

surrounding tissues (lower middle). Compared to normal liver tissues (bottom), pro-miR-302-treated cancers clearly displayed similar lobule structures, normal gland cell-like arrangements, and very clear boundaries between cell-cell and cancer-tissue junctions (black arrows), suggesting that these treated cancers have been greatly downgraded to a very benign state. Further continuous treatments of the pro-miR-302 drug over six to ten times could completely eliminate the cancer xenografts in all six samples (n = 6).

#### **A. Definitions**

To facilitate understanding of the invention, a number of terms are defined below:

**Nucleotide:** a monomeric unit of DNA or RNA consisting of a sugar moiety (pentose), a phosphate, and a nitrogenous heterocyclic base. The base is linked to the sugar moiety via the glycosidic carbon (1' carbon of the pentose) and that combination of base and sugar is a nucleoside. A nucleoside containing at least one phosphate group bonded to the 3' or 5' position of the pentose is a nucleotide. DNA and RNA are consisted of different types of nucleotide units called deoxyribonucleotide and ribonucleotide, respectively.

**Oligonucleotide:** a molecule comprised of two or more DNAs and/or RNAs, preferably more than three, and usually more than ten. An oligonucleotide longer than 13 nucleotide monomers is also called polynucleotide. The exact size will depend on many factors, which in turn depends on the ultimate function or use of the oligonucleotide. The oligonucleotide may be generated in any manner, including chemical synthesis, DNA replication, RNA transcription, reverse transcription, or a combination thereof.

**Nucleotide Analog:** a purine or pyrimidine nucleotide that differs structurally from adenine (A), thymine (T), guanine (G), cytosine (C), or uracil (U), but is sufficiently similar to substitute for the normal nucleotide in a nucleic acid molecule.

**Nucleic Acid Composition:** a nucleic acid composition refers to an oligonucleotide or polynucleotide such as a DNA or RNA sequence, or a mixed DNA/RNA sequence, in either a single-stranded or a double-stranded molecular structure.

**Gene:** a nucleic acid composition whose oligonucleotide or polynucleotide sequence codes for an RNA and/or a polypeptide (protein). A gene can be either RNA or DNA. A gene may encode a non-coding RNA, such as small hairpin RNA (shRNA), microRNA (miRNA), rRNA, tRNA, snoRNA, snRNA, and their RNA precursors as well as derivatives. Alternatively, a gene may encode a protein-coding RNA essential for protein/peptide synthesis, such as messenger RNA (mRNA) and its RNA precursors as well as derivatives.

In some cases, a gene may encode a protein-coding RNA that also contains at least a microRNA or shRNA sequence.

Primary RNA Transcript: an RNA sequence that is directly transcribed from a gene without any RNA processing or modification, which may be selected from the group consisting of mRNA, hnRNA, rRNA, tRNA, snoRNA, snRNA, pre-microRNA, viral RNA and their RNA precursors as well as derivatives.

Precursor messenger RNA (pre-mRNA): primary RNA transcripts of a protein-coding gene, which are produced by eukaryotic type-II RNA polymerase (Pol-II) machineries in eukaryotes through an intracellular mechanism termed transcription. A pre-mRNA sequence contains a 5'-untranslated region (UTR), a 3'-UTR, exons and introns.

Intron: a part or parts of a gene transcript sequence encoding non-protein-reading frames, such as in-frame intron, 5'-UTR and 3'-UTR.

Exon: a part or parts of a gene transcript sequence encoding protein-reading frames (cDNA), such as cDNA for cellular genes, growth factors, insulin, antibodies and their analogs/homologs as well as derivatives.

Messenger RNA (mRNA): assembly of pre-mRNA exons, which is formed after intron removal by intracellular RNA splicing machineries (spliceosomes) and served as a protein-coding RNA for peptide/protein synthesis. The peptides/proteins encoded by mRNAs include, but not limited, enzymes, growth factors, insulin, antibodies and their analogs/homologs as well as derivatives.

Complementary DNA (cDNA): a single-stranded or double-stranded DNA that contains a sequence complementary to an mRNA sequence and does not contain any intronic sequence.

Sense: a nucleic acid molecule in the same sequence order and composition as the homologous mRNA. The sense conformation is indicated with a "+", "s" or "sense" symbol.

Antisense: a nucleic acid molecule complementary to the respective mRNA molecule. The antisense conformation is indicated as a "-" or "\*" symbol or with an "a" or "antisense" in front of the DNA or RNA, *e.g.*, "aDNA" or "aRNA".

Base Pair (bp): a partnership of adenine (A) with thymine (T), or of cytosine (C) with guanine (G) in a double stranded DNA molecule. In RNA, uracil (U) is substituted for thymine. Generally the partnership is achieved through hydrogen bonding.

Base Pair (bp): a partnership of adenine (A) with thymine (T), or of cytosine (C) with guanine (G) in a double stranded DNA molecule. In RNA, uracil (U) is substituted for thymine. Generally the partnership is achieved through hydrogen bonding. For example, a sense nucleotide sequence "5'-A-T-C-G-U-3'" can form complete base pairing with its antisense sequence "5'-A-C-G-A-T-3'".

5'-end: a terminus lacking a nucleotide at the 5' position of successive nucleotides in which the 5'-hydroxyl group of one nucleotide is joined to the 3'-hydroxyl group of the next nucleotide by a phosphodiester linkage. Other groups, such as one or more phosphates, may be present on the terminus.

3'-end: a terminus lacking a nucleotide at the 3' position of successive nucleotides in which the 5'-hydroxyl group of one nucleotide is joined to the 3'-hydroxyl group of the next nucleotide by a phosphodiester linkage. Other groups, most often a hydroxyl group, may be present on the terminus.

Template: a nucleic acid molecule being copied by a nucleic acid polymerase. A template can be single-stranded, double-stranded or partially double-stranded, depending on the polymerase. The synthesized copy is complementary to the template, or to at least one strand of a double-stranded or partially double-stranded template. Both RNA and DNA are synthesized in the 5' to 3' direction. The two strands of a nucleic acid duplex are always aligned so that the 5' ends of the two strands are at opposite ends of the duplex (and, by necessity, so then are the 3' ends).

Nucleic Acid Template: a double-stranded DNA molecule, double stranded RNA molecule, hybrid molecules such as DNA-RNA or RNA-DNA hybrid, or single-stranded DNA or RNA molecule.

Conserved: a nucleotide sequence is conserved with respect to a pre-selected (referenced) sequence if it non-randomly hybridizes to an exact complement of the pre-selected sequence.

Homologous or Homology: a term indicating the similarity between a polynucleotide and a gene or mRNA sequence. A nucleic acid sequence may be partially or completely homologous to a particular gene or mRNA sequence, for example. Homology may be expressed as a percentage determined by the number of similar nucleotides over the total number of nucleotides.

Complementary or Complementarity or Complementation: a term used in reference to matched base pairing between two polynucleotides (*i.e.* sequences of an mRNA and a cDNA) related by the aforementioned "base pair (bp)" rules. For example, the sequence "5'-A-G-T-3'" is complementary to the sequence "5'-A-C-T-3'", and also to "5'-A-C-U-3'". Complementation can be between two DNA strands, a DNA and an RNA strand, or between two RNA strands. Complementarity may be "partial" or "complete" or "total". Partial complementarity or complementation occurs when only some of the nucleic acid bases are matched according to the base pairing rules. Complete or total complementarity or complementation occurs when the bases are completely or perfectly matched between the nucleic acid strands. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, as well as in detection methods that depend on binding between nucleic acids. Percent complementarity or complementation refers to the number of mismatch bases over the total bases in one strand of the nucleic acid. Thus, a 50% complementation means that half of the bases were mismatched and half were matched. Two strands of nucleic acid can be complementary even though the two strands differ in the number of bases. In this situation, the complementation occurs between the portion of the longer strand corresponding to the bases on that strand that pair with the bases on the shorter strand.

Complementary Bases: nucleotides that normally pair up when DNA or RNA adopts a double stranded configuration.

Complementary Nucleotide Sequence: a sequence of nucleotides in a single-stranded molecule of DNA or RNA that is sufficiently complementary to that on another single strand to specifically hybridize between the two strands with consequent hydrogen bonding.

Hybridize and Hybridization: the formation of duplexes between nucleotide sequences which are sufficiently complementary to form complexes via base pairing. Where a primer (or splice template) "hybridizes" with target (template), such complexes (or hybrids) are sufficiently stable to serve the priming function required by a DNA polymerase to initiate DNA synthesis. There is a specific, *i.e.* non-random, interaction between two complementary polynucleotides that can be competitively inhibited.

Posttranscriptional Gene Silencing: a targeted gene knockout or knockdown effect at the level of mRNA degradation or translational suppression, which is usually triggered by either foreign/viral DNA or RNA transgenes or small inhibitory RNAs.

RNA Interference (RNAi): a posttranscriptional gene silencing mechanism in eukaryotes, which can be triggered by small inhibitory RNA molecules such as microRNA (miRNA), small hairpin RNA (shRNA) and small interfering RNA (siRNA). These small RNA molecules usually function as gene silencers, interfering with expression of intracellular genes containing either completely or partially complementarity to the small RNAs.

Gene Silencing Effect: a cell response after a gene function is suppressed, consisting but not limited of cell cycle attenuation, G0/G1-checkpoint arrest, tumor suppression, anti-tumorigenicity, cancer cell apoptosis, and a combination thereof.

Non-coding RNA: an RNA transcript that cannot be used to synthesize peptides or proteins through intracellular translation machineries. Non-coding RNA includes long and short regulatory RNA molecules such as microRNA (miRNA), small hairpin RNA (shRNA), small interfering RNA (siRNA) and double strand RNA (dsRNA). These regulatory RNA molecules usually function as gene silencers, interfering with expression of intracellular genes containing either completely or partially complementarity to the non-coding RNAs.

MicroRNA (miRNA): single-stranded RNA capable of binding to targeted gene transcripts (mRNAs) that have partial complementarity to the sequence of microRNA. Mature microRNA is usually sized about 17–27 oligonucleotides in length and is able to either directly degrade its intracellular mRNA target(s) or suppress the protein translation of its targeted mRNA(s), depending on the complementarity between the microRNA and its target mRNA(s). Native microRNAs are found in almost all eukaryotes, functioning as a defense against viral infections and allowing regulation of specific gene expression during development of plants and animals. In principle, one microRNA often target multiple target mRNAs to fulfill its full functionality while on the other hand multiple miRNAs may target the same gene transcripts to enhance the effect of gene silencing.

MicroRNA Precursor (Pre-miRNA): hairpin-like single-stranded RNA containing stem-arm and stem-loop regions for interacting with intracellular RNase III Dicer endoribonucleases to produce one or multiple mature microRNAs (miRNAs) capable of silencing a targeted gene or a specific group of targeted genes that contain full or partial complementarity to the mature microRNA sequence(s). The stem-arm of a pre-miRNA can

form either a perfectly (100%) or a partially (mis-matched) hybrid duplexes, while the stem-loop connects one end of the stem-arm duplex to form a circle or hairpin-loop conformation required for being assembled into an RNA-induced silencing complex (RISC) with some argonaute proteins (AGO).

Prokaryote-produced MicroRNA Precursor (Pro-miRNA): small hairpin-like RNA similar to natural microRNA precursor (pre-miRNA) but transcribed from an artificially recombinant microRNA-expressing plasmid driven by a eukaryotic promoter in prokaryotic competent cells. For example, pro-miR-302 is structurally as same as pre-miR-302 (FIGs. 13A and 13B) but transcribed from either a *pLVX-Grn-miR302+367* or *pLenti-EF1alpha-RGFP-miR302* vector in *E. coli DH5alpha* competent cells (Example 1). As prokaryotic cells normally do not express short RNAs with high secondary structures such as eukaryotic pre-miRNA, the production of pro-miRNA in prokaryotes usually requires the addition of chemical inducer(s) in order to stimulate the eukaryotic promoter-driven pre-miRNA transcription (FIGs. 2–4).

Small interfering RNA (siRNA): short double-stranded RNA sized about 18-27 perfectly base-paired ribonucleotide duplexes and capable of degrading target gene transcripts with almost perfect complementarity.

Small or short hairpin RNA (shRNA): single-stranded RNA that contains a pair of partially or completely matched stem-arm nucleotide sequences divided by an unmatched loop oligonucleotide to form a hairpin-like structure. Many natural miRNAs are derived from hairpin-like RNA precursors, namely precursor microRNA (pre-miRNA).

Vector: a recombinant nucleic acid composition such as recombinant DNA (rDNA) capable of movement and residence in different genetic environments. Generally, another nucleic acid is operatively linked therein. The vector can be capable of autonomous replication in a cell in which case the vector and the attached segment is replicated. One type of preferred vector is an episome, *i.e.*, a nucleic acid molecule capable of extrachromosomal replication. Preferred vectors are those capable of autonomous replication and expression of nucleic acids. Vectors capable of directing the expression of genes encoding for one or more polypeptides and/or non-coding RNAs are referred to herein as “expression vectors” or “expression-competent vectors”. Particularly important vectors allow cloning of cDNA from mRNAs produced using a reverse transcriptase. A vector may contain components consisting of a viral or a type-II RNA polymerase (Pol-II or pol-2)

promoter, or both, a Kozak consensus translation initiation site, polyadenylation signals, a plurality of restriction/cloning sites, a pUC origin of replication, a SV40 early promoter for expressing at least an antibiotic resistance gene in replication-competent prokaryotic cells, an optional SV40 origin for replication in mammalian cells, and/or a tetracycline responsive element. The structure of a vector can be a linear or circular form of single- or double-stranded DNA selected from the group consisting of plasmid, viral vector, transposon, retrotransposon, DNA transgene, jumping gene, and a combination thereof.

Promoter: a nucleic acid to which a polymerase molecule recognizes, or perhaps binds to, and initiates RNA transcription. For the purposes of the instant invention, a promoter can be a known polymerase or its cofactor binding site, an enhancer and the like, any sequence that can initiate synthesis of RNA transcripts by a desired polymerase.

Eukaryotic Promoter: a sequence of nucleic acid motifs which are required for RNA and/or gene transcription and can be recognized by eukaryotic type II RNA polymerases (pol-2), pol-2 equivalent, and/or pol-2 compatible (pol-2-like) viral polymerases for initiating the RNA/gene transcription.

Type-II RNA Polymerase (Pol-II or pol-2) Promoter: an RNA promoter that can be recognized by eukaryotic type-II RNA polymerases (Pol-II or pol-2) and hence is able to initiate the transcription of eukaryotic messenger RNAs (mRNAs) and/or microRNAs (miRNAs). For example, but not limited, a pol-2 promoter can be a mammalian RNA promoter or a cytomegaloviral (CMV) promoter.\_\_\_\_

Type-II RNA Polymerase (Pol-II or pol-2) Equivalent: a eukaryotic transcription machinery selected from the group consisting of mammalian type-II RNA polymerases (Pol-II or pol-2) and Pol-II compatible (pol-2-like) viral RNA polymerases.

Pol-II Compatible (pol-2-like) Viral Promoter: a viral RNA promoter capable of using the eukaryotic pol-2 or pol-2 equivalent transcription machineries for initiating gene and/or RNA expression. For example, but not limited, a pol-2-like viral promoter can be a cytomegaloviral (CMV) promoter or a retroviral long terminal repeat (LTR) promoter.

Cistron: a sequence of nucleotides in a DNA molecule coding for an amino acid residue sequence and including upstream and downstream DNA expression control elements.

Intron Excision: a cellular mechanism responsible for RNA processing, maturation and degradation, including RNA splicing, exosome digestion, nonsense-mediated decay (NMD) processing, and a combination thereof.

RNA Processing: a cellular mechanism responsible for RNA maturation, modification and degradation, including RNA splicing, intron excision, exosome digestion, nonsense-mediated decay (NMD), RNA editing, RNA processing, and a combination thereof.

Targeted Cell: a single or a plurality of human cells selected from the group consisting of a somatic cell, a tissue, a stem cell, a germ-line cell, a teratoma cell, a tumor cell, a cancer cell, and a combination thereof.

Cancerous Tissue: a neoplastic tissue derived from the group consisting of skin cancer, prostate cancer, breast cancer, liver cancer, lung cancer, brain tumor/cancer, lymphoma, leukemia and a combination thereof.

Expression-Competent Vector: a linear or circular form of single- or double-stranded DNA selected from the group consisting of plasmid, viral vector, transposon, retrotransposon, DNA transgene, jumping gene, and a combination thereof.

Antibiotic Resistance Gene: a gene capable of degrading antibiotics selected from the group consisted of penicillin G, streptomycin, ampicillin (Amp), neomycin, G418, kanamycin, erythromycin, paromycin, phophomycin, spectromycin, tetracycline (Tet), doxycycline (Dox), rifapicin, amphotericin B, gentamycin, chloramphenicol, cephalothin, tylosin, and a combination thereof.

Restriction/Cloning Site: a DNA motif for restriction enzyme cleavage including but not limited *AatII*, *AccI*, *AflIII/III*, *AgeI*, *ApaI/LI*, *AseI*, *Asp718I*, *BamHI*, *BbeI*, *BclI/II*, *BglII*, *BsmI*, *Bsp120I*, *BspHI/LU111/120I*, *BsrI/BI/GI*, *BssHIII/SI*, *BstBI/U1/XI*, *ClaI*, *Csp6I*, *DpnI*, *DraI/II*, *EagI*, *Ecl136II*, *EcoRI/RII/47III/RV*, *EheI*, *FspI*, *HaeIII*, *HhaI*, *HinPI*, *HindIII*, *HinfI*, *HpaI/II*, *KasI*, *KpnI*, *MaeII/III*, *MfeI*, *MluI*, *MscI*, *MseI*, *NaeI*, *NarI*, *NcoI*, *NdeI*, *NgoMI*, *NotI*, *NruI*, *NsiI*, *PmlI*, *Ppu10I*, *PstI*, *PvuI/II*, *RsaI*, *SacI/II*, *Sall*, *Sau3AI*, *SmaI*, *SnaBI*, *SphI*, *SspI*, *StuI*, *TaiI*, *TaqI*, *XbaI*, *XhoI*, *XmaI* cleavage site.

Gene Delivery: a genetic engineering method selected from the group consisting of polysomal transfection, liposomal transfection, chemical transfection, electroporation, viral infection, DNA recombination, transposon insertion, jumping gene insertion, microinjection, gene-gun penetration, and a combination thereof.

Genetic Engineering: a DNA recombination method selected from the group consisting of DNA restriction and ligation, homologous recombination, transgene incorporation, transposon insertion, jumping gene integration, retroviral infection, and a combination thereof.

Cell Cycle Regulator: a cellular gene involved in controlling cell division and proliferation rates, consisting but not limited of CDK2, CDK4, CDK6, cyclins, BMI-1, p14/p19Arf, p15Ink4b, p16Ink4a, p18Ink4c, p21Cip1/Waf1, and p27Kip1, and a combination thereof.

Tumor Suppression Effect: a cellular anti-tumor and/or anti-cancer mechanism and response consisting of, but not limited, cell cycle attenuation, cell cycle arrest, inhibition of tumor cell growth, inhibition of cell tumorigenicity, inhibition of tumor/cancer cell transformation, induction of tumor/cancer cell apoptosis, induction of normal cell recovery, reprogramming high-grade malignant cancer cells to a more benign low-grade state (tumor regression), and a combination thereof.

Cancer Therapy Effect: a cell response and/or cellular mechanism resulted from a drug treatment, including, but not limited, inhibition of oncogene expression, inhibition of cancer cell proliferation, inhibition of cancer cell invasion and/or migration, inhibition of cancer metastasis, induction of cancer cell death, prevention of tumor/cancer formation, prevention of cancer relapse, suppression of cancer progression, repairing damaged tissue cells, reprogramming high-grade malignant cancers to a more benign low-grade state (cancer regression/remission), and a combination thereof.

Gene Silencing Effect: a cell response after a gene function is suppressed, consisting of, but not limited, inhibition of oncogene expression, inhibition of cell proliferation, cell cycle arrest, tumor suppression, cancer regression, cancer prevention, cell apoptosis, cell repairing and/or rejuvenation, cell reprogramming, reprogramming diseased cells to a relatively normal state (spontaneous healing), and a combination thereof.

Cancer Reversion: a reprogramming mechanism that resets the malignant properties of high-grade cancers back to a relatively normal-like low-grade state in vitro, ex vivo or in vivo.

Targeted Cell: a single or a plurality of human cells selected from the group consisting of a somatic cell, a tissue, a stem cell, a germ-line cell, a teratoma cell, a tumor cell, a cancer cell, and a combination thereof.

Cancerous Tissue: a neoplastic tissue derived from the group consisting of skin cancer, prostate cancer, breast cancer, liver cancer, lung cancer, brain tumor/cancer, lymphoma, leukemia and a combination thereof.

Transcriptional Inducer: a chemical agent that can induce and/or enhance eukaryotic RNA and/or gene transcription from a pol-2 or pol-2-like promoter in prokaryotic cells. For example, a transcription inducer contains, but not limited, a chemical structure similar to MOPS, ethanol, glycerin, as well as their functional analogs such as 2-(N-morpholino)ethanesulfonic acid (MES), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and mannitol, or a mixture thereof.

Antibody: a peptide or protein molecule having a pre-selected conserved domain structure coding for a receptor capable of binding a pre-selected ligand.

Pharmaceutical and/or Therapeutic Application: a biomedical utilization, device and/or apparatus useful for diagnosis, stem cell generation, stem cell research and/or therapy development, tissue/organ repair and/or rejuvenation, wound healing treatment, tumor suppression, cancer therapy and/or prevention, disease treatment, drug production, and a combination thereof.

## **B. Compositions and Applications**

A composition and method for producing a new kind of prokaryote-produced microRNA precursors (pro-miRNAs) capable of reprogramming the malignant properties of human cancers into a low-grade benign or normal-like state in vitro, ex vivo and in vivo, comprising:: (a) at least a chemical inducer agent containing a structure similar to 3-morpholinopropane-1-sulfonic acid (MOPS), ethanol, or glycerin, or a mixture thereof; and (b) a plurality of prokaryotic cells that contain at least a pre-miRNA-encoding gene mediated by eukaryotic pol-2 and/or pol-2-like promoter-driven transcription; wherein said (a) and (b) are mixed together under a condition to induce the expression of said gene, so as to generate the encoded pre-miRNA in the prokaryotic cells. Notably, the chemical inducer is able to stimulate eukaryotic promoter-driven RNA transcription in prokaryotes!

In principle, the present invention provides a novel composition design and its applicable strategy for inducing a quick adaptation of prokaryotes to use eukaryotic pol-2 and pol-2-like promoters for directly expressing certain desired microRNA precursors (pre-miRNA) without the need of using error-prone prokaryotic promoters or growing laborious and costly hybridomas or mammalian cells.

Preferably, said prokaryote is a bacterial cell strain in particular, *Escherichia coli* (E. coli), and said chemical inducer is 3-morpholinopropane-1-sulfonic acid (MOPS), ethanol, or

glycerin, or a mixture thereof. Also preferably, said eukaryotic promoter is either a eukaryotic pol-2 promoter, such as EF1alpha, or a pol-2 compatible (pol-2-like) viral promoter, such as cytomegaloviral (CMV) promoter or retroviral long terminal repeat (LTR) promoter. The pre-miRNA-encoding gene mediated by said eukaryotic promoter is coded for either a non-coding or a protein-coding RNA transcript, or both (such as an intron-containing gene transcript), selected from the group consisted of microRNA (miRNA), small hairpin RNA (shRNA), small interfering RNA (siRNA), messenger RNA (mRNA) and their precursors as well as shRNA/siRNA homologues, and a combination thereof. The protein-coding RNA may be selected from, but not limited to, the group consisted of a gene encoding enzyme, growth factor, antibody, insulin, botulinum toxin (botox), a functional protein and/or its analogs, and a combination thereof. Preferably, said condition for inducing the expression of said pre-miRNA-encoding gene is a bacterial culturing condition such as Luria-Bertani (LB) broth at 37°C with the addition of said chemical inducer(s).

#### **BRIEF DESCRIPTION OF THE DRAWINGS:**

The patent or application file contains at least one color drawing. Copies of this patent or patent application publication with color drawing will be provided by the USPTO upon request and payment of the necessary fee.

Referring particularly to the drawings for the purpose of illustration only and not limitation, there is illustrated:

FIGs. 1A and 1B show a eukaryotic promoter-driven expression vector composition (1A) and its expression mechanism (1B) for RNA transcript and/or protein production in prokaryotes. For demonstrating the present invention, a new *pLenti-EF1alpha-RGFP-miR302* vector (FIG. 1A) is served as an example composition for transforming *E. coli DH5alpha* competent cells to produce RGFP proteins as well as miR-302s and their precursors (pre-miR-302s) under the stimulation of MOPS, glycerin and/or ethanol. *pLenti-EF1alpha-RGFP-miR302* is a lentiviral plasmid vector that is designed by the inventors to expresses various microRNAs/shRNAs, mRNAs and/or proteins/peptides in both prokaryotes and eukaryotes. According to the disclosed mechanism (1B), it is easy for an ordinary skill in the art to use any microRNA/shRNA in place miR-302 or any mRNA/protein in place of RGFP as described in the present invention. Black arrows indicate the pathways occurring

in both prokaryotic and eukaryotic cells, while blank arrows indicate the steps only occurring in the eukaryotic cells.

FIG. 2 depicts the results of bacterial culture broths treated with (left) or without (right) the mixture of 0.1% (v/v) MOPS and 0.05% (v/v) glycerin. The *E. coli* competent cells have been transformed by *pLenti-EF1alpha-RGFP-miR302* before the treatment of chemical inducers.

FIG. 3 shows the results of different bacterial pellets after treated with 0.1% (v/v) MOPS. The *E. coli* competent cells have been transformed by either *pLVX-Grn-miR302+367* (green) or *pLenti-EF1alpha-RGFP-miR302* (red) before the MOPS treatment.

FIG. 4 shows the inducibility of various chemical inducers for inducing pol-2 promoter-driven gene expression in *E. coli* competent cells. Among all chemicals tested, the top three most potent inducers are MOPS, glycerin and ethanol. The chemical concentration used can be ranged from about 0.001% to 4%, most preferably, from 0.01% to 1%.

FIG. 5 shows the Western blotting results of red RGFP protein expression induced by MOPS, glycerin and ethanol, respectively. Bacterial RuvB protein was used as a house-keeping standard to normalize the detected RGFP expression. Proteins extracted from blank *E. coli* cells, *i.e.* transformed with no vector, were used as a negative control.

FIG. 6 shows the Northern blotting results of the expression of the miR-302 familial cluster (~700nt) and its derivative precursors (pre-miR-302s with 1 to 4 hairpins) induced by MOPS, glycerin and ethanol, respectively. RNAs extracted from blank *E. coli* cells were used as a negative control.

FIG. 7 shows iPSC generation using miR-302 and/or pre-miR-302 isolated from bacterial competent cell extracts (BE), which is confirmed by Northern blot analysis as shown in FIG. 6. As reported, miR-302-reprogrammed iPSCs (or called mirPSCs) form sphere-like cell colonies and express strong Oct4 as a standard hESC marker.

FIG. 8 shows the global DNA demethylation of Oct4 and Sox2 gene promoters induced by the miR-302 and/or pre-miR-302 isolated from bacterial competent cell extracts (BE), which is confirmed by Northern blot analysis as shown in FIG. 6. As demonstrated by Simonsson and Gurdon (*Nat Cell Biol.* 6, 984-990, 2004), both signs of global DNA demethylation and Oct4 expression are required for somatic cell reprogramming to form iPSCs.

FIG. 9 shows the *in vitro* tumorigenicity assays of human liver cancer cell line HepG2 in response to miR-302 transfection. The cells obtained after miR-302 transfection are labeled as mirPS-HepG2, indicating the change of their cancer cell properties into an induced pluripotent stem cell (iPSC)-like state. Changes of morphology and cell cycle rate before and after miR-302 transfection were compared. Each cell DNA content respective to cell cycle stages was shown by a peak chart of flow cytometry analysis above the cell morphology ( $n = 3$ ,  $p < 0.01$ ).

FIGs. 10A and 10B show the results of HPLC purification and analysis using a synthetic standard uDNA (by Sigma-Genosys) and freshly extracted pro-miR-302s isolated from *pLenti-EF1alpha-RGFP-miR302*-transformed *E. coli* cells. The standard uDNA was designed to be equal to a natural pre-miR-302a as: 5'-CCACCACUUA AACGUGGAUG UACUUGCUIIU GAAACUAAAG AAGUAAGUGC UUGCAUGUIIU UGGUGAUGG-3' (SEQ.ID.NO.4).

FIGs. 11A and 11B show the results of microRNA (miRNA) microarray analyses using small RNAs extracted from either blank *E. coli* competent cells or *pLenti-EF1alpha-RGFP-miR302* (*RGFP-miR302*)-transfected cells. The extracted small RNAs were further purified by HPLC as shown in the green-labeled area of FIG. 10B. FIG. 11A shows that RNAs from blank *E.coli* cells present almost no microRNA (green dots mean non-statistically significant whereas red dots indicate positive results). This is because prokaryotes lack several essential enzymes required for microRNA expression and processing, such as Pol-2, Drosha and RNase III Dicer. Also, prokaryotic RNA polymerases do not efficiently transcribe small RNAs with high secondary structures, such as hairpin-like pre-miRNAs and shRNAs. As a result, only using the present invention we can stimulate the expression of specific microRNAs, such as miR-302a, a\*, b, b\*, c, c\*, d and d\* as shown in FIG. 11B, in prokaryotic cells. Since prokaryotic cells possess no Dicer, all microRNAs remain in their precursor conformations, such as pri-miRNA (4-hairpin cluster) and/or pre-miRNA (1 hairpin precursor). Taken together, the results of FIGs. 10B and 11B have established two facts as: (1) small RNAs extracted from the *RGFP-miR302*-transfected cells contain mostly pure miR-302 precursors, and (2) there is almost no other kind of microRNA contamination in the *E. coli* competent cells.

FIG. 12 shows the lists of expressed microRNAs extracted from either blank *E. coli* competent cells (Group 1 as shown in FIG. 11A) or *pLenti-EF1alpha-RGFP-miR302*-

transfected cells (Group 2 as shown in FIG. 11B). Signals less than 500 are not statistically significant (as shown in green in FIGs. 11A and 11B), which may be caused by either low copy number expression or high background.

FIGs. 13A and 13B show the sequencing results of the miR-302 familial cluster (13A) and the individual pro-miR-302a, pro-miR-302b, pro-miR-302c, and pro-miR-302d sequences (13B). The result of the miR-302 familial cluster (= pri-miR-302) is 5'-AAUUUUUUUC UUCUAAAGUU AUGCCAUUUU GUUUUCUUUC UCCUCAGCUC UAAAUACUCU GAAGUCCAAA GAAGUUGUAU GUUGGGUGGG CUCCCUUCAA CUUUAACAUG GAAGUGCUUU CUGUGACUUU AAAAGUAAGU GCUUCCAUGU UUUAGUAGGA GUGAAUCCAA UUUACUUCUC CAAAUAAGAA CACGCUAACC UCAUUUGAAG GGAUCCCCUU UGCUUUAACA UGGGGGUACC UGCUGUGUGA AACAAAAGUA AGUGCUUCCA UGUUUCAGUG GAGGUGUCUC CAAGCCAGCA CACCUUUUGU UACAAA AUU UUUUGUUAU GUGUUUUAAG GUUACUAAGC UUGUUAACAGG UUAAGGAU CUAACUUUUU CCAAGACUGG GCUCCCCACC ACUUAACGU GGAUGUACUU GCUUUGAAAC UAAAGAAGUA AGUGCUUCCA UGUUUUGGUG AUGGUAAGUC UUCUUUUUAC AUUUUUUAUA UUUUUUAGA AAAUAACUU AUUGUAUUGA CCGCAGCUC UAUUUUAAG CUUUAUUUUG UAUUUUACA UCUGUUAAGG GGCCCCCUCU ACUUUAACAU GGAGGCACUU GCUGUGACAU GACAAAAUA AGUGCUUCCA UGUUUGAGUG UGGUGGUUCC UACCUAUCA GCAAUUGAGU UAACGCCAC ACUGUGUGCA GUUCUUGGCU ACAGGCCAUU ACUGUUGCUA-3' (SEQ.ID.NO.5), while the individual sequences of pro-miR-302a, pro-miR-302b, pro-miR-302c, and pro-miR-302d are as follows: 5'-CCACCACUUA AACGUGGAUG UACUUGCUUU GAAACUAAAG AAGUAAGUGC UUCAUGUUU UGGUGAUGG-3' (SEQ.ID.NO.6), 5'-GCUCCCUUCA ACUUUAACAU GGAAGUGCUU UCUGUGACUU UAAAAGUAAG UGCUUCCAUG UUUUAGUAGG AGU-3' (SEQ.ID.NO.7), 5'-CCUUUGCUUU AACAUUGGGG UACCUGCUGU GUGAAACAAA AGUAAGUGCU UCCAUGUUUC AGUGGAGG-3' (SEQ.ID.NO.8), and 5'-CCUCUACUU AACAUUGGAGG CACUUGCUGU GACAUGACAA AAAUAAGUGC UUCAUGUUU GAGUGUGG-3' (SEQ.ID.NO.9), respectively.

FIG. 14 shows the in vivo therapeutic results of a pre-investigational new drug (pre-IND) trial using pro-miR-302 as an injection drug to treat human liver cancer xenografts in SCID-beige nude mice. Following three treatments (once per week), the pro-miR-302 drug

(= pre-miR-302) successfully reduced cancer sizes from  $728 \pm 328 \text{ mm}^3$  (untreated blank control, C) to  $75 \pm 15 \text{ mm}^3$  (pro-mir-302-treated, T), indicating a ~90% reduction rate in the average cancer size! No significant therapeutic effect was found in the treatments of synthetic siRNA mimics (siRNA-302). Further histological examination (most right) found that normal liver lobule-like structures (circles pointed by a black arrow) were formed only in pro-miR-302-treated cancers but not other treatments or controls, suggesting that a reprogramming mechanism may occur to reset the malignant cancer cell properties back to a relatively normal-like state, called “Cancer Reversion”.

FIG. 15 shows the histological similarity between normal liver tissues and pro-mir-302-treated human liver cancer xenografts in vivo. After three treatments (once per week), the pro-mir-302 drug successfully reprogrammed high-grade (grade IV) human liver cancer grafts to a more benign low-grade (less than grade II) state. Similar to normal liver tissues (top), the treated cancer grafts could form classical liver lobules, containing central vein (CV)-like and portal triad (PT)-like structures (indicated by black arrows). As cancer cells are generally more acidic than normal liver cells, the result of hematoxylin & eosin (H&E) staining shows more purple in cancer cells whereas more red in normal liver cells.

FIG. 16 shows the patho-histological comparison among untreated, siRNA-treated, pro-mir-302-treated human liver cancer grafts and normal liver tissues in SCID-beige nude mice. Without treatment (top), the engrafted human liver cancer aggressively invaded into normal tissues, such as muscles and blood vessels, and formed massive cell–cell and cancer–tissue fusion structures, indicating its malignancy and high metastasis. Treatment of siRNA mimics (siRNA-302) did not significantly reduce the malignancy of the engrafted cancer (upper middle), probably due to the short half-life of siRNA. In contrast, pro-miR-302 treatment not only reprogrammed the engrafted cancer to a relatively normal-like morphology (no fusion) but also greatly inhibited cancer invasion into the surrounding tissues (lower middle). Compared to normal liver tissues (bottom), pro-miR-302-treated cancers formed normal-like lobule structures, gland-like cell arrangements, and clear boundaries between cell–cell and cancer–tissue junctions (black arrows), indicating that these treated cancers have been downgraded to a very benign state.

FIGs. 17A and 17B show comparison of the healing results between untreated (17A) and miR-302-treated (17B) wounds in vivo. The isolated miR-302 molecules (20~400  $\mu\text{g/mL}$ ) were formulated with di-/tri-glycylglycerins, a delivery reagent, and antibiotic

ointment to form candidate drugs for testing topic treatments of large 2cm x 2cm open wounds on pig back skins in vivo (n = 6 for each group). After about two-week treatments (one treatment per day), the healed wounds were dissected and further made into tissue sections for histological examination under a microscope. The data showed that no or very little scar (scarless) could be seen in the miR-302-treated wounds (17B top, n = 6/6), whereas almost all untreated (treated with only antibiotic ointment) wounds contained large scars (17A top, n = 5/6). Also, a significantly large amount of CD34-positive adult stem cell clusters (labeled by green fluorescent antibodies) were found in the miR-302-treated wounds (17B bottom, n = 6/6), but not in untreated control wounds (17A bottom, n = 0/6). These results indicate that pre-miR-302 is able to induce CD34-positive adult stem cell expansion and/or regeneration, so as to enhance tissue repairing and regeneration, leading to a very beneficial therapeutic effect on lesions caused by human degenerative diseases, such as Alzheimer's diseases, Parkinson's diseases, osteoporosis, diabetes, and cancers. Such therapeutic effect may also help to reprogram high-grade malignant cancers into low-grade benign or even normal-like tissues, a novel mechanism called Cancer Reversion or Cancer Regression.

## EXAMPLES:

In the experimental disclosure which follows, the following abbreviations apply: M (molar); mM (millimolar);  $\mu$ M (micromolar); mol (moles); pmol (picomoles); gm (grams); mg (milligrams)  $\mu$ g (micrograms); ng (nanograms); L (liters); ml (milliliters);  $\mu$ l (microliters); °C (degrees Centigrade); RNA (ribonucleic acid); DNA (deoxyribonucleic acid); dNTP (deoxyribonucleotide triphosphate); PBS (phosphate buffered saline); NaCl (sodium chloride); HEPES (N-2-hydroxyethylpiperazine-N-2-ethanesulfonic acid); HBS (HEPES buffered saline); SDS (sodium dodecylsulfate); Tris-HCl (tris-hydroxymethylaminomethane-hydrochloride); ATCC (American Type Culture Collection, Rockville, MD); hESC (human embryonic stem cells); and iPSC (induced pluripotent stem cells).

### **1. Bacterial Cell Culture and Chemical Treatments**

*E. coli DH5alpha* competent cells were acquired as a part from the z-competent *E. coli* transformation kit (Zymo Research, Irvine, CA) and then transformed by mixing with 5  $\mu$ g of a pre-made plasmid vector such as *pLenti-EF1alpha-RGFP-miR302* or *pLVX-Grn-*

*miR302+367*. Non-transformed cells were normally grown in Luria-Bertani (LB) broth supplemented with 10 mM MgSO<sub>4</sub> and 0.2 mM glucose at 37°C with frequent agitation at 170 rpm, whereas the transformed cells are cultivated in the above LB broth further supplemented with additional 100 µg/ml ampicillin. For chemical induction, 0.5 to 2 ml of MOPS, glycerin and/or ethanol, respectively or in combination, was added into 1 liter LB broth supplemented with 10 mM MgSO<sub>4</sub> and 0.2 mM glucose in the presence of 100 µg/ml ampicillin. For negative control, the transformed cells were cultivated in the above ampicillin-supplemented LB broth but without adding any chemical inducer. The results are shown in FIGs. 2–4.

## **2. Human Cell Culture and MicroRNA Transfection**

Human liver cancer cell line HepG2 was obtained from ATCC and maintained according to manufacturer's suggestions. For transfection, 15 µg of pre-miR-302 was dissolved in 1 ml of fresh RPMI medium and mixed with 50 µl of X-tremeGENE HP DNA transfection reagent (Roche, Indianapolis, IN). After 10 min incubation, the mixture is added into a 100-mm cell culture dish containing 50%~60% confluency of HepG2. The medium was refreshed 12 to 18 hours later. After these transfected cells formed sphere-like iPSC colonies, the medium was changed to a knockout DMEM/F-12 medium (Invitrogen) supplemented with 20% knockout serum, 1% MEM nonessential amino acids, 100 µM β-mercaptoethanol, 1 mM GlutaMax, 1 mM sodium pyruvate, 10 ng/ml bFGF, 10 ng/ml FGF-4, 5 ng/ml LIF, 100 IU/ml penicillin/100 µg/ml streptomycin, 0.1 µM A83-01, and 0.1 µM valproic acid (Stemgent, San Diego, CA), and the cells were cultivated at 37°C under 5% CO<sub>2</sub>. The result is shown in FIG. 9.

## **3. Protein Extraction and Western Blot Analysis**

Cells (10<sup>6</sup>) were lysed with a CellLytic-M lysis/extraction reagent (Sigma) supplemented with protease inhibitors, Leupeptin, TLCK, TAME and PMSF, following the manufacturer's suggestion. Lysates were centrifuged at 12,000 rpm for 20 min at 4°C and the supernatant was recovered. Protein concentrations were measured using an improved SOFTmax protein assay package on an E-max microplate reader (Molecular Devices, CA). Each 30 µg of cell lysate was added to SDS-PAGE sample buffer under reducing (+50 mM DTT) and non-reducing (no DTT) conditions, and boiled for 3 min before loading onto a 6~8% polyacrylamide gel. Proteins were resolved by SDS-polyacrylamide gel electrophoresis (PAGE), electroblotted onto a nitrocellulose membrane and incubated in

Odyssey blocking reagent (Li-Cor Biosciences, Lincoln, NB) for 2 hours at room temperature. Then, a primary antibody was applied to the reagent and incubated the mixture at 4°C. Primary antibodies included Oct3/4 (Santa Cruz Biotechnology, Santa Cruz, CA), RuvB (Santa Cruz) and RGFP (Clontech). After overnight, the membrane was rinsed three times with TBS-T and then exposed to goat anti-mouse IgG conjugated secondary antibody to Alexa Fluor 680 reactive dye (1:2,000; Invitrogen–Molecular Probes), for 1 hour at the room temperature. After three additional TBS-T rinses, fluorescent scanning of the immunoblot and image analysis was conducted using Li-Cor Odyssey Infrared Imager and Odyssey Software v.10 (Li-Cor). The results are shown in FIG. 5.

#### **4. RNA Extraction and Northern Blot Analysis**

Total RNAs (10 µg) were isolated with a *mirVana*<sup>TM</sup> miRNA isolation kit (Ambion, Austin, TX), fractionated by either 15% TBE-urea polyacrylamide gel or 3.5% low melting point agarose gel electrophoresis, and electroblotted onto a nylon membrane. Detection of miR-302s and the related pre-miR-302s was performed with a [LNA]-DNA probe (5'-[TCACTGAAAC] ATGGAAGCAC TTA-3') (SEQ.ID.NO.10) probe. The probe has been purified by high-performance liquid chromatography (HPLC) and tail-labeled with terminal transferase (20 units) for 20 min in the presence of [<sup>32</sup>P]-dATP (> 3000 Ci/mM, Amersham International, Arlington Heights, IL). The results are shown in FIG. 6.

#### **5. Plasmid Amplification and Plasmid DNA/Total RNA Extraction**

*E. coli DH5alpha* competent cells after transformation (from Example 1) were cultivated in LB broth supplemented with 10 mM MgSO<sub>4</sub> and 0.2 mM glucose at 37°C with frequent agitation at 170 rpm. For inducing eukaryotic promoter-driven RNA transcription, 0.5 to 2 ml of MOPS, glycerin, and/or ethanol was added into every 1 liter of LB broth for propagating the transformed cells overnight. The amplified plasmid DNAs and expressed mRNAs/microRNAs in the transformed cells were isolated using a HiSpeed plasmid purification kit (Qiagen, Valencia, CA), following the manufacturer's protocol but with a minor modification that RNase A was not added into the P1 buffer. After that, the final extracted products containing both plasmids and mRNAs/microRNAs were dissolved in DEPC-treated ddH<sub>2</sub>O and stored at -80°C before use. For purifying only the amplified plasmid vectors, RNase A was added into the P1 buffer and the extraction procedure was performed following the manufacturer's protocol.

## **6. MicroRNA and Pre-miRNA Isolation/Purification**

For purifying microRNAs and pre-miRNAs, the total RNAs isolated from Example 5 were further extracted using a *mirVana*<sup>™</sup> miRNA isolation kit (Ambion, Austin, TX), following the manufacturer's protocol. The final products so obtained were dissolved in DEPC-treated ddH<sub>2</sub>O and stored at -80°C before use. Because bacterial RNAs are naturally degraded very fast (within a few hours) whereas eukaryotic hairpin-like microRNA precursors (pre-miRNAs and pri-miRNAs) remain largely stable at 4°C (half-life up to 3-4 days), we can use this half-life difference to acquire relatively pure pri-/pre-miRNAs for other applications. For example, the pre-miR-302s so obtained can be used to reprogram somatic cells to hESC-like iPSCs, as shown in FIG. 9.

## **7. Immunostaining Assay**

Embedding, sectioning and immunostaining tissue samples were performed as previously reported (Lin et al., 2008). Primary antibodies include Oct4 (Santa Cruz) and RGFP (Clontech, Palo Alto, CA). Fluorescent dye-labeled goat anti-rabbit or horse anti-mouse antibody was used as the secondary antibody (Invitrogen—Molecular Probes, Carlsbad, CA). Positive results were examined and analyzed at 100x or 200x magnification under a fluorescent 80i microscopic quantitation system with a Metamorph imaging program (Nikon). The result is shown in FIG. 7.

## **8. Bisulfite DNA Sequencing**

Genomic DNAs were isolated from ~2,000,000 cells using a DNA isolation kit (Roche) and 1 µg of the isolated DNAs was further treated with bisulfite (CpGenome DNA modification kit, Chemicon, Temecula, CA), following the manufacturers' suggestion. The bisulfite treatment converted all unmethylated cytosine to uracil, while methylated cytosine remained as cytosine. For bisulfite DNA sequencing, we amplified the promoter region of the *Oct4* gene with PCR primers: 5'-GAGGCTGGAG CAGAAGGATT GCTTTGG-3' (SEQ.ID.NO.11) and 5'-CCCTCCTGAC CCATCACCTC CACCACC-3' (SEQ.ID.NO.12). For PCR, the bisulfite-modified DNAs (50 ng) were mixed with the primers (total 100 pmol) in 1x PCR buffer, heated to 94°C for 2 min, and immediately cooled on ice. Next, 25 cycles of PCR were performed as follows: 94°C for 1 min and 70°C for 3 min, using an Expand High Fidelity PCR kit (Roche). The PCR product with a correct size was further fractionized by 3% agarose gel electrophoresis, purified by a gel extraction filter (Qiagen), and then used in DNA sequencing. After that, a detailed profile of DNA methylation sites

was generated by comparing the unchanged cytosine in the converted DNA sequence to the unconverted one, as shown in FIG. 8.

#### **9. DNA-density Flow Cytometry**

Cells were trypsinized, pelleted and fixed by re-suspension in 1 ml of pre-chilled 70% methanol in PBS for 1 hour at -20°C. The cells were pelleted and washed once with 1 ml of PBS and then pelleted again and resuspended in 1 ml of 1 mg/ml propidium iodide, 0.5 µg/ml RNase in PBS for 30 min at 37°C. After that, about 15,000 cells were analyzed on a BD FACSCalibur (San Jose, CA). Cell doublets were excluded by plotting pulse width versus pulse area and gating on the single cells. The collected data were analyzed using the software package Flowjo using the “Watson Pragmatic” algorithm. The result was shown in the top panels of FIG. 9.

#### **10. MicroRNA (miRNA) Microarray Analysis**

At about 70% confluency, small RNAs from each cell culture were isolated, using the *mirVana*<sup>TM</sup> miRNA isolation kit (Ambion). The purity and quantity of the isolated small RNAs were assessed, using 1% formaldehyde-agarose gel electrophoresis and spectrophotometer measurement (Bio-Rad), and then immediately frozen in dry ice and submitted to LC Sciences (San Diego, CA) for miRNA microarray analyses. Each microarray chip was hybridized a single sample labeled with either Cy3 or Cy5 or a pair of samples labeled with Cy3 and Cy5, respectively. Background subtraction and normalization were performed as manufacturer's suggestions. For a dual sample assay, a *p*-value calculation was performed and a list of differentially expressed transcripts more than 3-fold (yellow–red signals) was produced. The final microarray results were shown in FIGs. 11A and 11B, and the list of differentially expressed microRNAs was shown in FIG. 12, which compared the small RNAs extracted from blank *E. coli* cell lysates (Group 1) to those extracted from *pLenti-EF1alpha-RGFP-miR302*-transformed cell lysates (Group 2).

#### **11. In Vivo In Vivo Liver Cancer Therapy Trials**

Xenografting human liver cancers into immunocompromised SCID-beige mice is a valid animal model for studying liver cancer metastasis and therapy. To establish this model, we mixed 5 million human hepatocarcinoma (HepG2) cells with 100 µL of matrix gel and subcutaneously engrafted the mixture into each flank of the mouse hind limbs, respectively. As a result, both sides of the mouse hind limbs were subjected to approximately the same amount of cancer cell engraftment. Cancers were observed about

two weeks post-engraftment and sized about  $15.6 \pm 8 \text{ mm}^3$  in average (starting cancer size before treatment). For each mouse, we selected the side with a larger cancer as the treatment group and the other smaller one as the control group. Since the same mouse was treated with a blank formulation reagent (negative control) in one side and the formulated drug (pro-mir-302) in the other side, the results so obtained can minimize any possible variation due to individual differences.

To deliver pro-mir-302 into the targeted cancer regions in vivo, we contracted a professional formulation company, Latitude (San Diego, CA), to liposomally encapsulate pro-miR-302s into 160~200 nm-diameter nanoparticles. These pro-miR-302-containing nanoparticles have been tested to be almost 100% stable at room temperature for over two weeks and at 4°C for over one month, whereas other synthetic siRNA mimics (siRNA-302) were all quickly degraded over 50% within 3 to 5 days under the same conditions, indicating that pro-miRNA rather than siRNA is stable enough to be used as a drug for therapy. For toxicity assay, we have further injected maximally 300  $\mu\text{L}$  of the formulated pro-miR-302 (1 mg/mL) into the mouse tail vein ( $n = 8$ ), respectively, and observed no detectable side effect in all tested mice over six months. In general, non-modified ribonucleic acids are relatively not immunogenic and can be easily metabolized by tissue cells, rendering a safe tool for in vivo therapy.

For testing drug potency, we subcutaneously injected 200  $\mu\text{L}$  of the formulated pro-mir-302 in one side and 200  $\mu\text{L}$  of the blank formulation reagent in the other side of the mice, respectively, and continued the same injection pattern for three times (one injection per week). The drug and reagent were applied to the surrounding region of the cancer site and absorbed by the cancer and its surrounding tissues within 18 hours. Samples were collected one week after the third injection. Hearts, livers, kidneys and the engrafted cancers were removed for further histological examination. Tumor formation was monitored by palpation and tumor volume was calculated using the formula  $(\text{length} \times \text{width}^2)/2$ . Tumor lesions were counted, dissected, weighed, and subjected to histological examination using H&E and immunostaining assays. Histological examination showed no detectable tissue lesions in heart, liver, and kidney. The results were shown in FIGs. 14, 15 and 16.

## **12. Statistic Analysis**

Any change over 75% of signal intensity in the analyses of immunostaining, western blotting and northern blotting was considered as a positive result, which in turn is analyzed

and presented as mean  $\pm$  SE. Statistical analysis of data was performed by one-way ANOVA. When main effects were significant, the Dunnett's post-hoc test was used to identify the groups that differed significantly from the controls. For pairwise comparison between two treatment groups, the two-tailed student *t* test was used. For experiments involving more than two treatment groups, ANOVA was performed followed by a post-hoc multiple range test. Probability values of  $p < 0.05$  was considered significant. All  $p$  values were determined from two-tailed tests.

## REFERENCES:

1. Lin SL and Ying SY. (2006) Gene silencing *in vitro* and *in vivo* using intronic microRNAs. Ying SY. (Ed.) *MicroRNA protocols*. Humana press, Totowa, New Jersey, pp295-312.
2. Lin SL, Chang D and Ying SY. (2006) Transgene-like animal models using intronic microRNAs. Ying SY. (Ed.) *MicroRNA protocols*. Humana press, Totowa, New Jersey, pp 321-334.
3. Lin SL, Chang D, Chang-Lin S, Lin CH, Wu DTS, Chen DT, and Ying SY. (2008) Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state. *RNA* 14, 2115-2124.
4. Lin SL and Ying SY. (2008) Role of mir-302 microRNA family in stem cell pluripotency and renewal. Ying SY. (Ed.) *Current Perspectives in MicroRNAs*. Springer Publishers press, New York, pp 167-185.
5. Lin SL, Chang D, Ying SY, Leu D and Wu DTS. (2010) MicroRNA miR-302 inhibits the tumorigenicity of human pluripotent stem cells by coordinate suppression of CDK2 and CDK4/6 cell cycle pathways. *Cancer Res.* 70, 9473-9482.
6. Lin SL, Chang D, Lin CH, Ying SY, Leu D and Wu DTS. (2011) Regulation of somatic cell reprogramming through inducible mir-302 expression. *Nucleic Acids Res.* 39, 1054-1065.
7. Simonsson S and Gurdon J. (2004) DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. *Nat Cell Biol.* 6, 984-990.
8. Takahashi et al. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663-676.

9. Wang et al. (2008). Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. *Nat. Genet.* 40, 1478–1483.
10. Wernig et al. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 448, 318–324.
11. Yu et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917–1920.
12. U.S. Patent No. 7,959,926 to Buechler.
13. U.S. Patent No. 7,968,311 to Mehta.
14. European Patent No. EP 2198025 to Lin.
15. U.S. Patent Application No. 12/149,725 to Lin.
16. U.S. Patent Application No. 12/318,806 to Lin.
17. U.S. Patent Application No. 12/792,413 to Lin.
18. U.S. Patent Application No. 13/572,263 to Lin.
19. U.S. Patent Application No. 13/964,705 to Lin.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 12

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA

(A) DESCRIPTION: /desc = "synthetic"

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: YES

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

UCACCAAAAC AUGGAAGCAC UUA

23

## (3) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA

(A) DESCRIPTION: /desc = "natural"

39

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:  
ACUUAAACGU GGAUGUACUU GCU

23

(4) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA

(A) DESCRIPTION: /desc = "natural"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:  
UAAGUGCUUC CAUGUUU

17

(5) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 69 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: hairpin

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic"

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCACCACUUA AACGUGGAUG UACUUGCUUU GAAACUAAAG AAGUAAGUGC  
UCCAUGUUU UGGUGAUGG 69

(6) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 720 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: multiple hairpins

(ii) MOLECULE TYPE: RNA

(A) DESCRIPTION: /desc = "recombinant"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AAUUUUUUUC UUCUAAAGUU AUGCCAUUUU GUUUUCUUUC UCCUCAGCUC  
UAAAUACUCU GAAGUCCAAA GAAGUUGUUA GUUGGGUGGG CUCCCUUCAA  
CUUUAACAUG GAAGUGCUUU CUGUGACUUU AAAAGUAAGU GCUUCCAUGU  
UUUAGUAGGA GUGAAUCCAA UUUACUUCUC CAAAUAAGAA CACGCUAACC  
UCAUUUGAAG GGAUCCCCUU UGCUUUAACA UGGGGGUACC UGCUGUGUGA  
AACAAAAGUA AGUGCUUCCA UGUUUCAGUG GAGGUGUCUC CAAGCCAGCA  
CACCUUUUGU UACAAAUUU UUUUGUUAU GUGUUUUAAG GUUACUAAGC

UUGUUACAGG UAAAAGGAUU CUAACUUUUU CCAAGACUGG GCUCCCCACC  
ACUAAAACGU GGAUGUACUU GCUUUGAAAC UAAAGAAGUA AGUGCUUCCA  
UGUUUUGGUG AUGGUAAGUC UUCUUUUUAC AUUUUUUAUA UUUUUUAGA  
AAAUAACUUU AUUGUAUUGA CCGCAGCUCA UAUUUUAAG CUUUAUUUUG  
UAUUUUUACA UCUGUUAAGG GGCCCCCUCU ACUUUAACAU GGAGGCACUU  
GCUGUGACAU GACAAAAUA AGUGCUUCCA UGUUUGAGUG UGGUGGUUCC  
UACCUAUUA GCAAUUGAGU UAACGCCCAC ACUGUGUGCA GUUCUUGGCU  
ACAGGCCAUU ACUGUUGCUA 720

(7) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 69 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: hairpin

(ii) MOLECULE TYPE: RNA

- (A) DESCRIPTION: /desc = "recombinant"

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCACCACUUA AACGUGGAUG UACUUGCUUU GAAACUAAAG AAGUAAGUGC  
UCCAUGUUU UGGUGAUGG 69

(8) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 73 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: hairpin

(ii) MOLECULE TYPE: RNA

(A) DESCRIPTION: /desc = "recombinant"

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCUCCCUUCA ACUUUAACAU GGAAGUGCUU UCUGUGACUU UAAAAGUAAG  
UGCUUCCAUG UUUUAGUAGG AGU 73

(9) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 68 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: hairpin

(ii) MOLECULE TYPE: RNA

(A) DESCRIPTION: /desc = "recombinant"

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CCUUUGCUUU AACAUAGGGGG UACCUGCUGU GUGAAACAAA AGUAAGUGCU  
UCCAUGUUUC AGUGGAGG 68

(10) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 68 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: hairpin

## (ii) MOLECULE TYPE: RNA

- (A) DESCRIPTION: /desc = "recombinant"

## (iii) HYPOTHETICAL: YES

## (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCUCUACUUU AACAUUGGAGG CACUUGCUGU GACAUGACAA AAAUAAGUGC  
UCCAUGUUU GAGUGUGG 68

## (11) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: LNA-DNA hybrid

- (A) DESCRIPTION: /desc = "synthetic"

## (iii) HYPOTHETICAL: YES

## (iv) ANTI-SENSE: YES

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

44

TCACTGAAAC ATGGAAGCAC TTA

23

## (12) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acids

(A) DESCRIPTION: /desc = "synthetic"

## (iii) HYPOTHETICAL: YES

## (iv) ANTI-SENSE: YES

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GAGGCTGGAG CAGAAGGATT GCTTTGG

27

## (13) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acids

(A) DESCRIPTION: /desc = "synthetic"

## (iii) HYPOTHETICAL: YES

## (iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:  
CCCTCCTGAC CCATCACCTC CACCACC

**CLAIMS:**

1. A method for inducing CD34-positive adult stem cell expansion or regeneration using microRNA precursor (pre-miRNA), comprising:
  - (a) providing at least a microRNA precursor containing SEQ.ID.NO.3;
  - (b) providing a cell substrate containing at least a CD34-positive cell, and;
  - (c) contacting the microRNA precursor of (a) with the cell substrate of (b), so as to induce the expansion or regeneration of the CD34-positive cell population.
2. The method as defined in Claim 1, wherein said microRNA precursor is produced by eukaryotic promoter-driven RNA transcription in prokaryotic cells.
3. The method as defined in Claim 2, wherein said eukaryotic promoter-driven RNA transcription is induced by contacting a chemical agent containing a molecular structure of 3-morpholinopropane-1-sulfonic acid (MOPS) with at least a transformed prokaryotic cell carrying at least an expression vector encoding a sequence of SEQ.ID.NO.8, SEQ.ID.NO.7, SEQ.ID.NO.8, or SEQ.ID.NO.9.
4. The method as defined in Claim 3, wherein said expression vector is a recombinant plasmid encoding the sequence of SEQ.ID.NO.5.
5. The method as defined in Claim 3, wherein said expression vector is pLenti-EF1alpha-RGFP-miR302 containing either a cytomegalovirus (CMV) or mammalian EF1alpha promoter, or both.
6. The method as defined in Claim 2, wherein said prokaryotic cells are E. coli competent cells.
7. The method as defined in Claim 1, wherein said microRNA precursor contains at least a hairpin-like sequence of SEQ.ID.NO.6, SEQ.ID.NO.7, SEQ.ID.NO.8, or SEQ.ID.NO.9.
8. The method as defined in Claim 1, wherein said microRNA precursor is a kind of prokaryote-produced miR-302 precursor (pre-miR-302).

9. The method as defined in Claim 8, wherein said pre-miR-302 contains at least a sequence of miR-302a, miR-302b, miR-302c, or miR-302d.
10. The method as defined in Claim 8, wherein said pre-miR-302 can be used as a part of drug ingredients for pharmaceutical and therapeutic applications.
11. The method as defined in Claim 1, wherein said microRNA precursor can be used as a part of drug ingredients for pharmaceutical and therapeutic applications.
12. The method as defined in Claim 1, wherein said CD34-positive adult stem cell expansion is useful for pharmaceutical or therapeutic applications.
13. The method as defined in Claim 12, wherein said CD34-positive adult stem cell expansion is useful for reprogramming the malignant properties of human cancer cells into a low-grade benign or normal-like state in vivo.
14. The method as defined in Claim 12, wherein said CD34-positive adult stem cell expansion is useful for enhancing scarless wound healing in vivo.
15. The method as defined in the Claim 1, wherein said CD34-positive cells include skin, hair, muscle, blood (hematopoietic), mesenchymal, and neural stem cells, or a combination thereof.

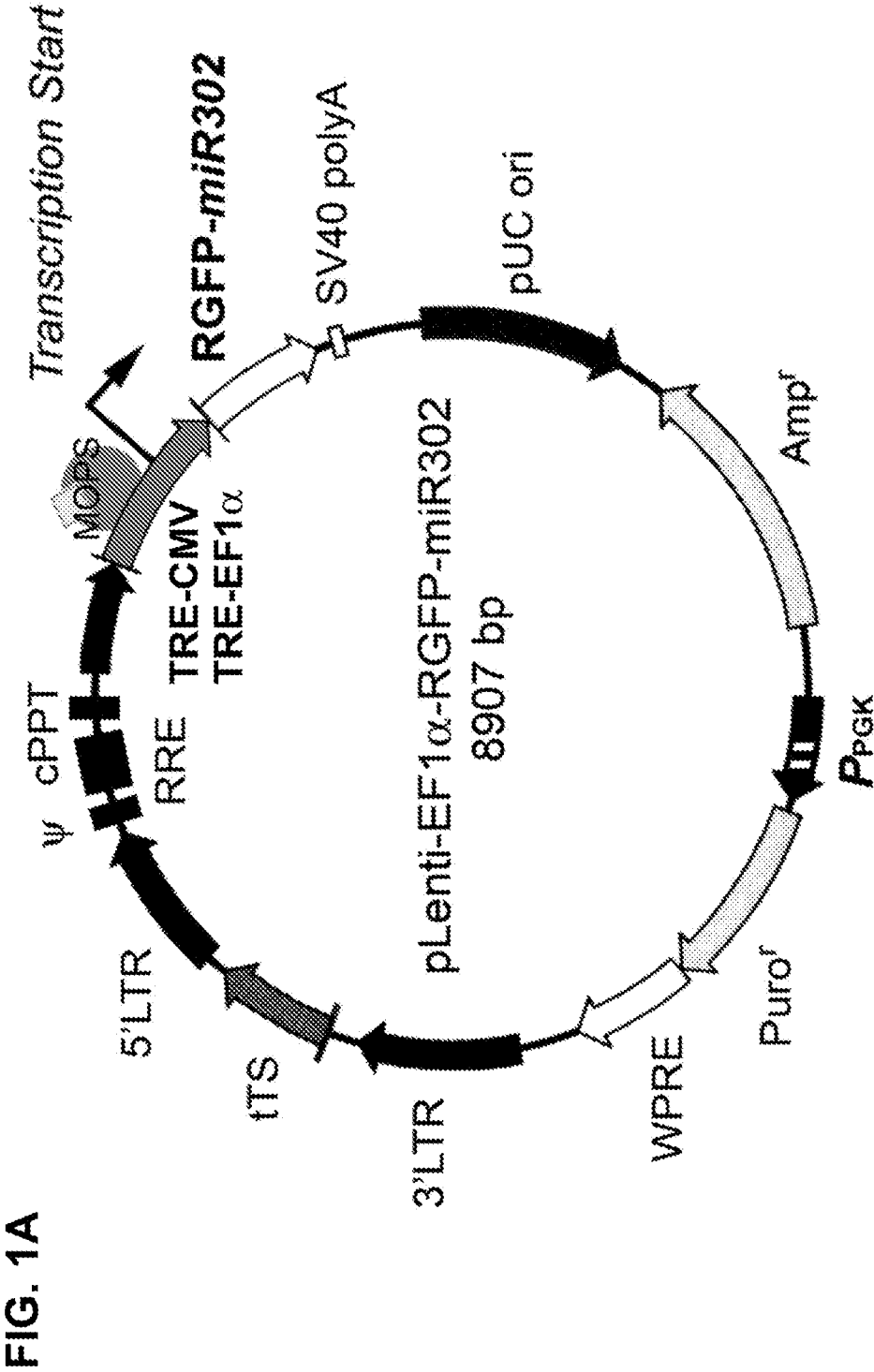
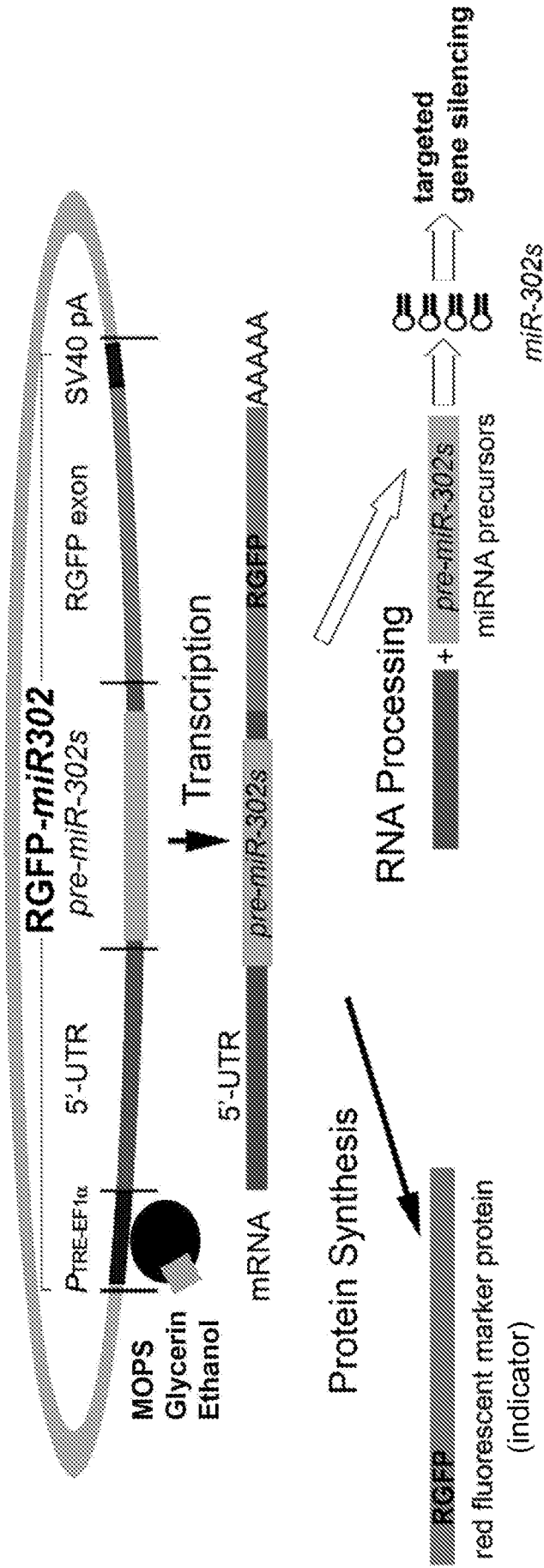
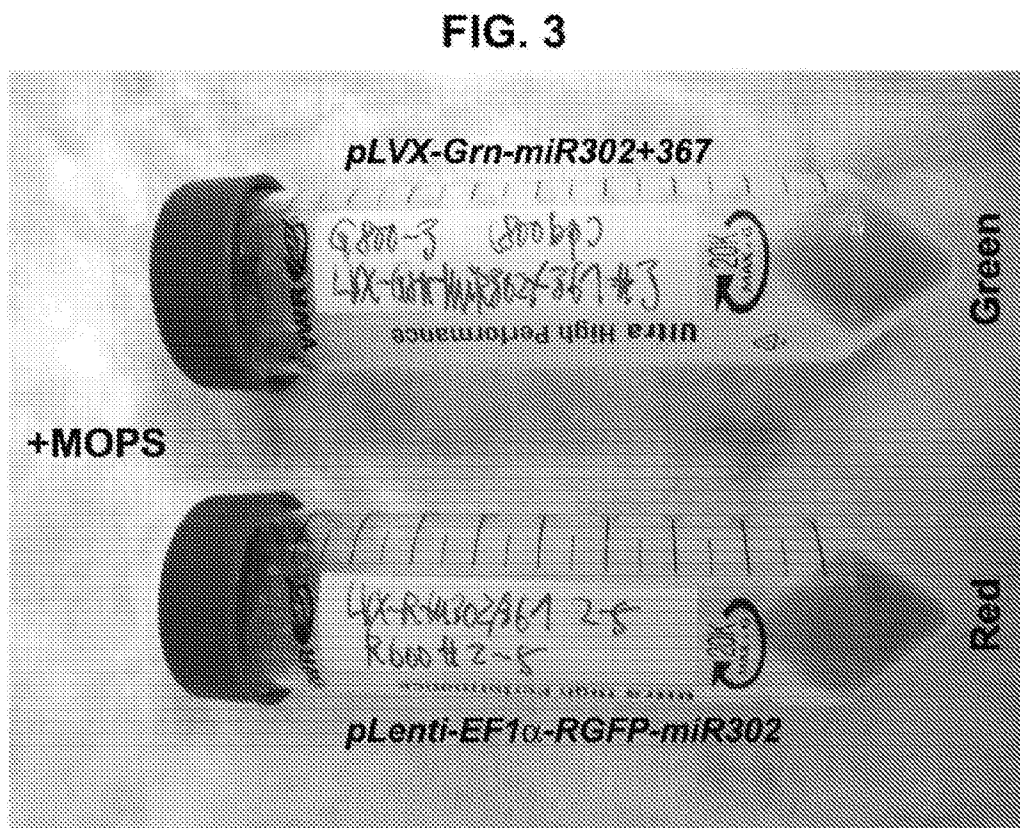
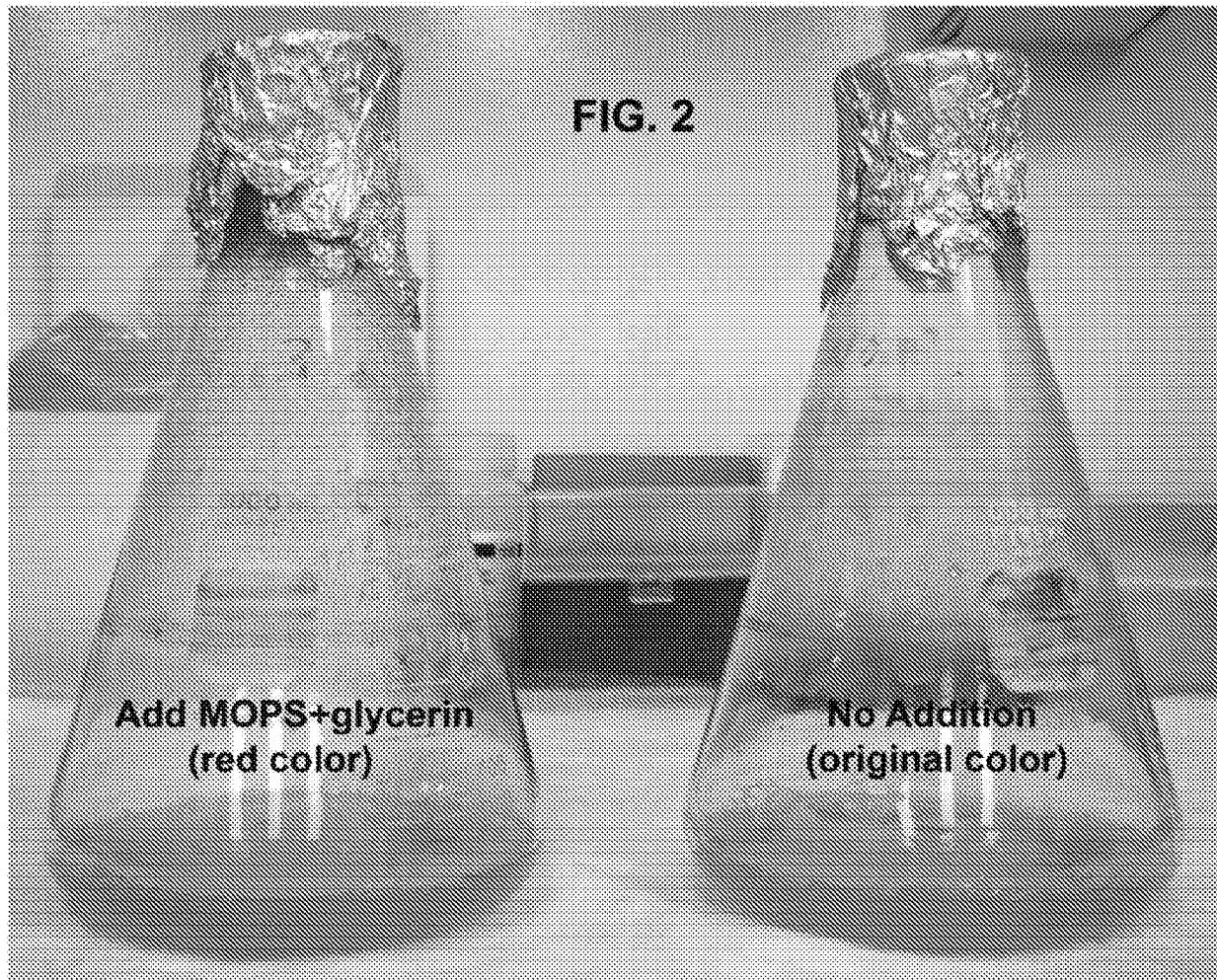


FIG. 1B





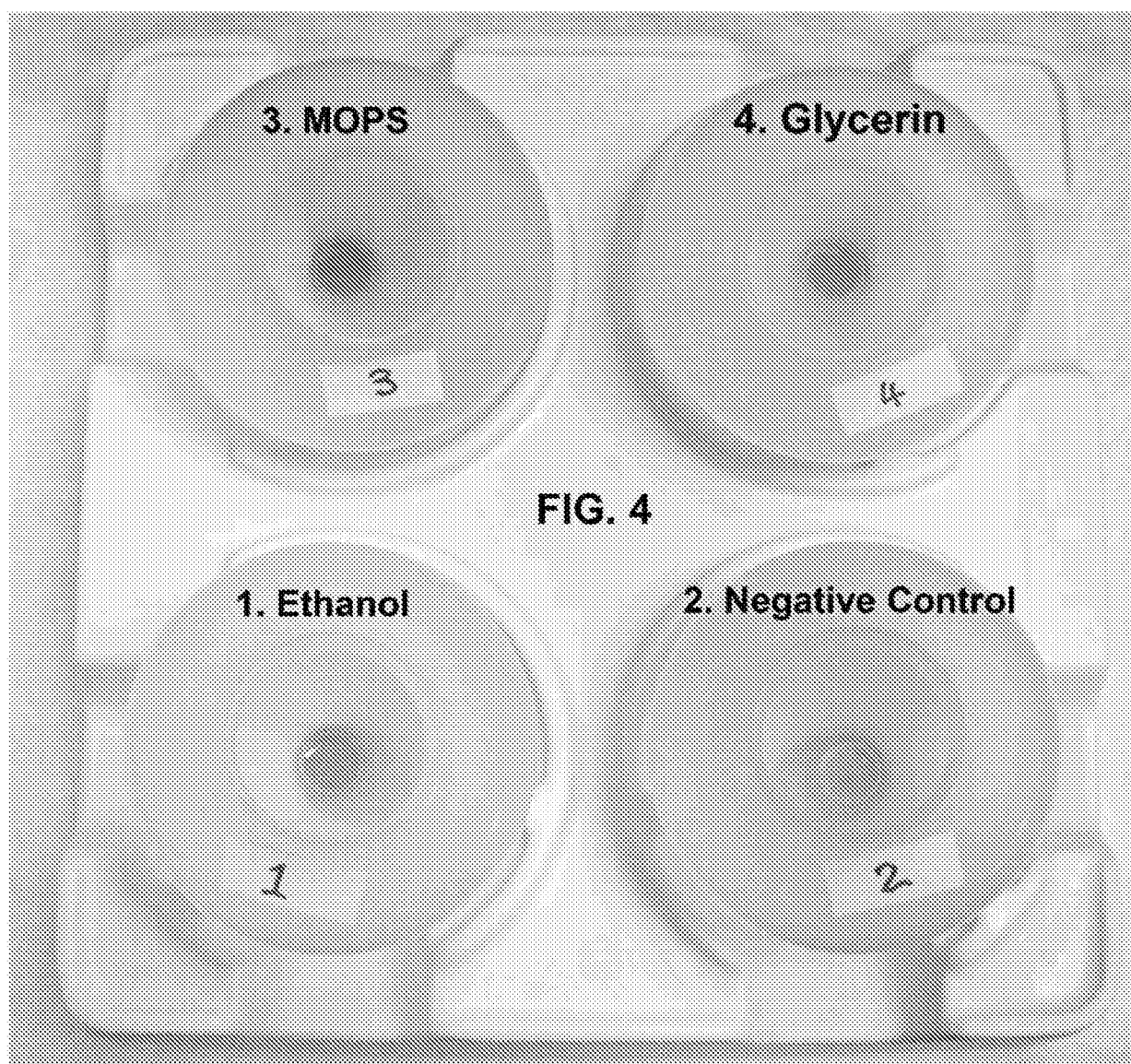


FIG. 5

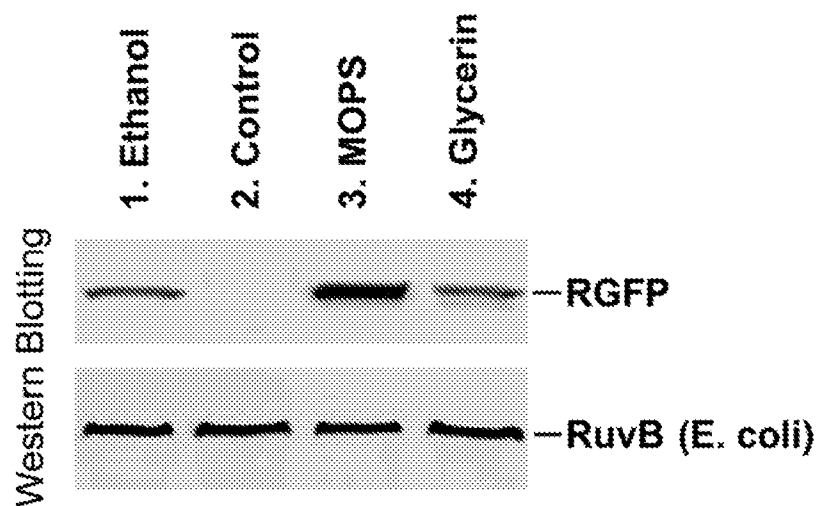


FIG. 6

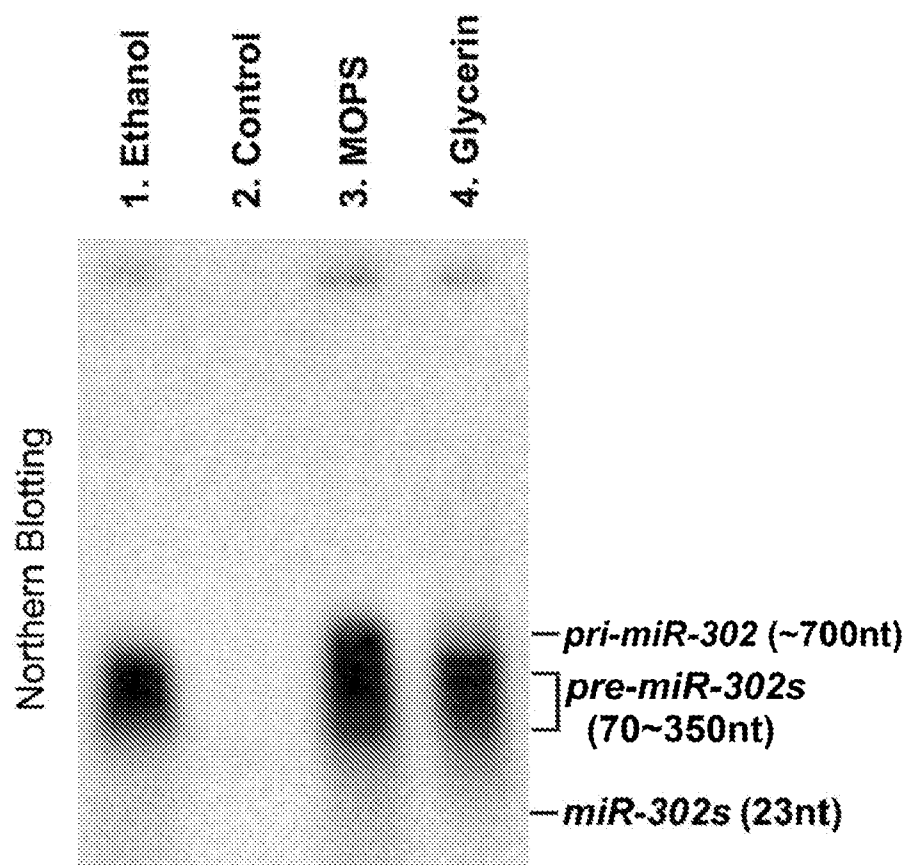


FIG. 7

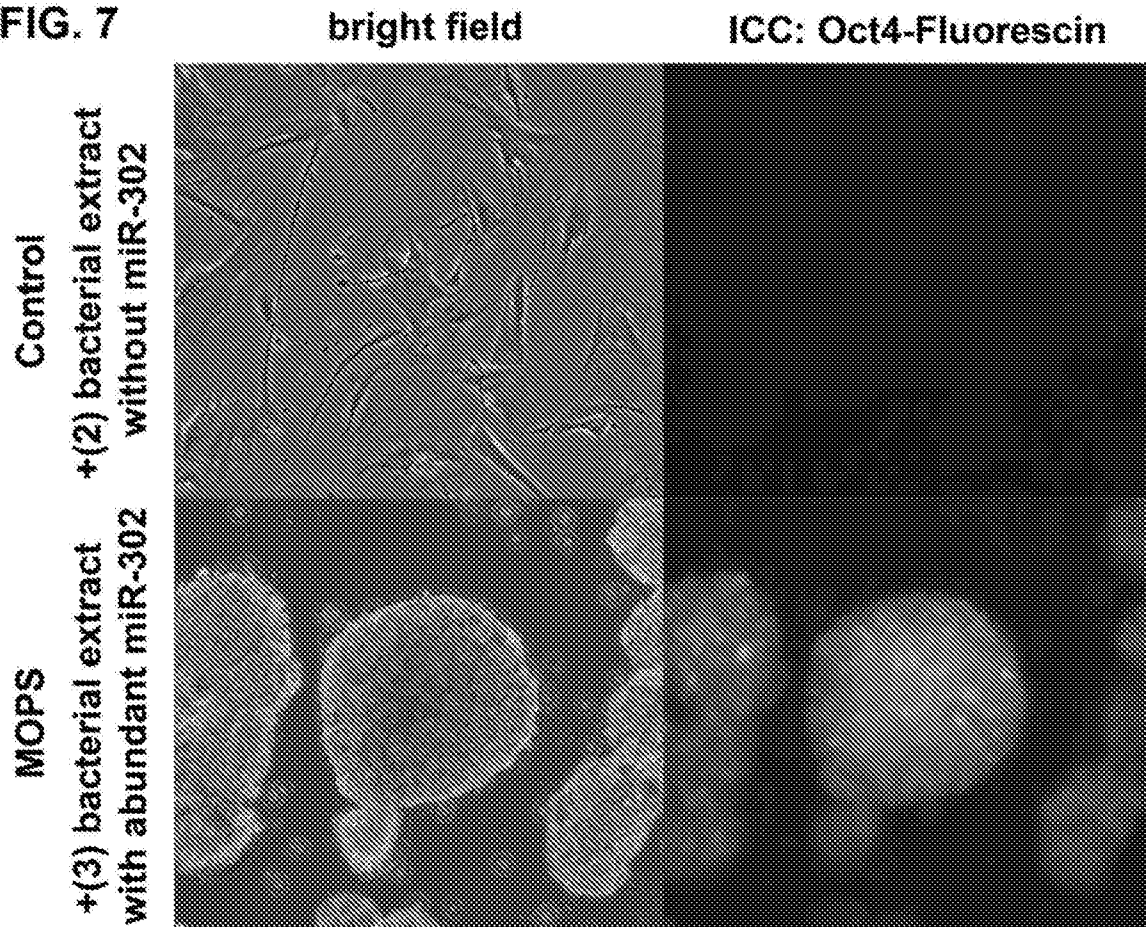
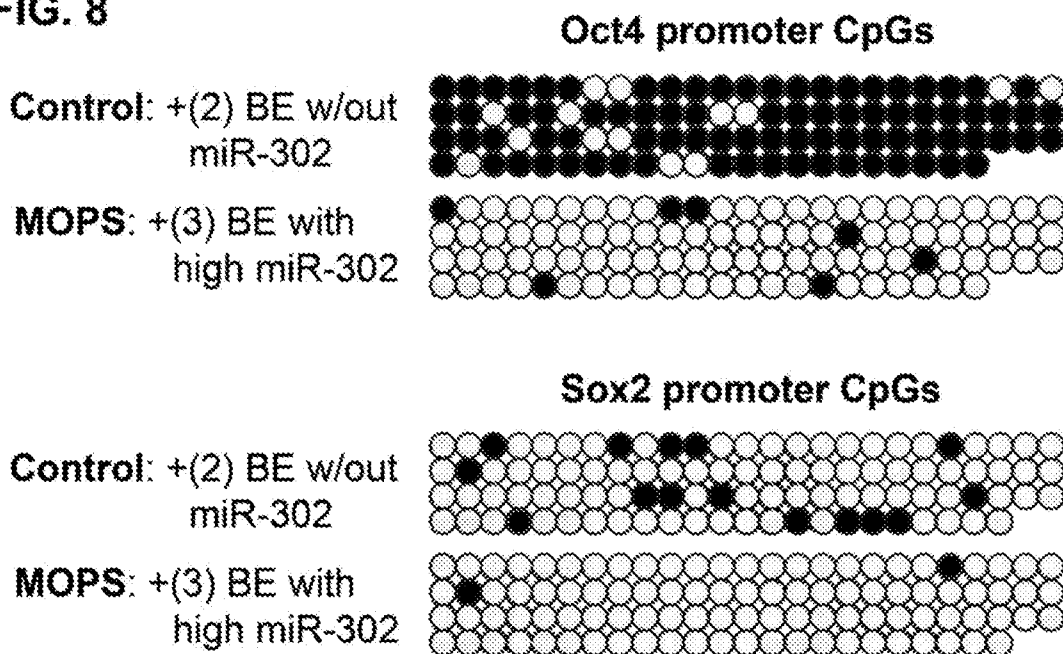


FIG. 8



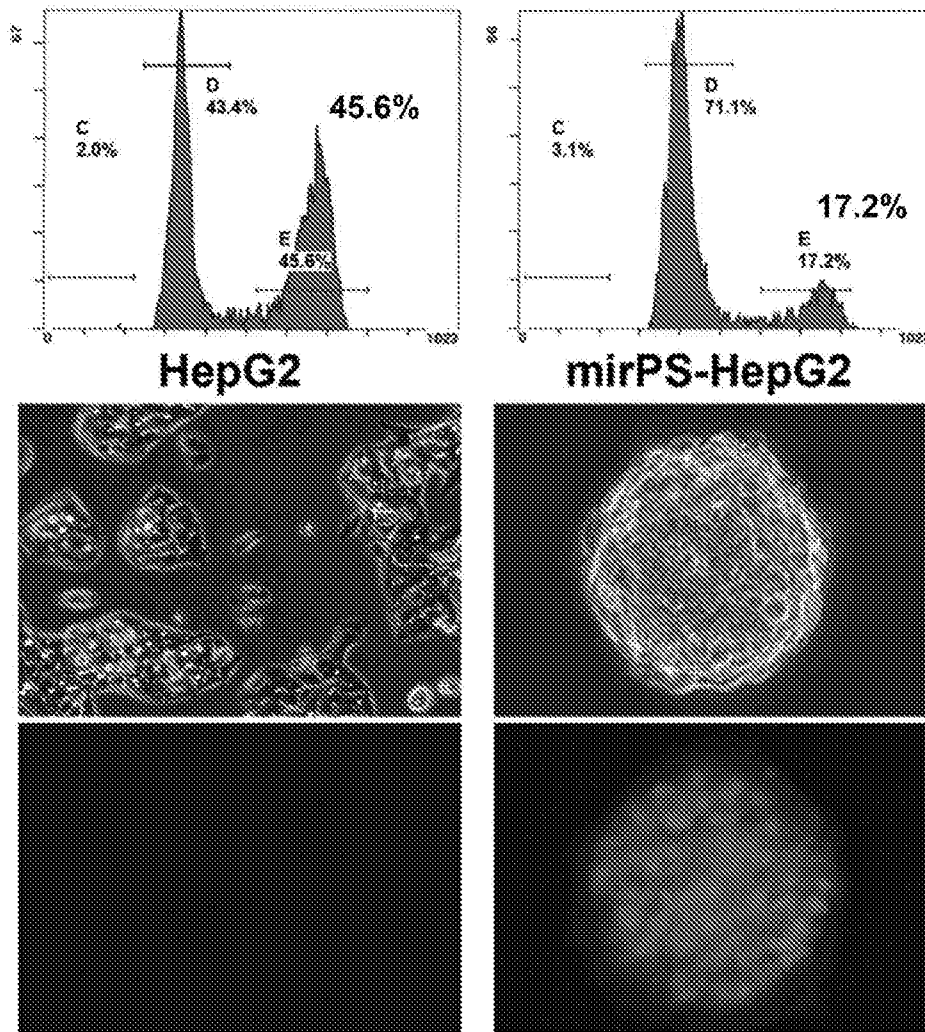
**FIG. 9**

FIG. 10A

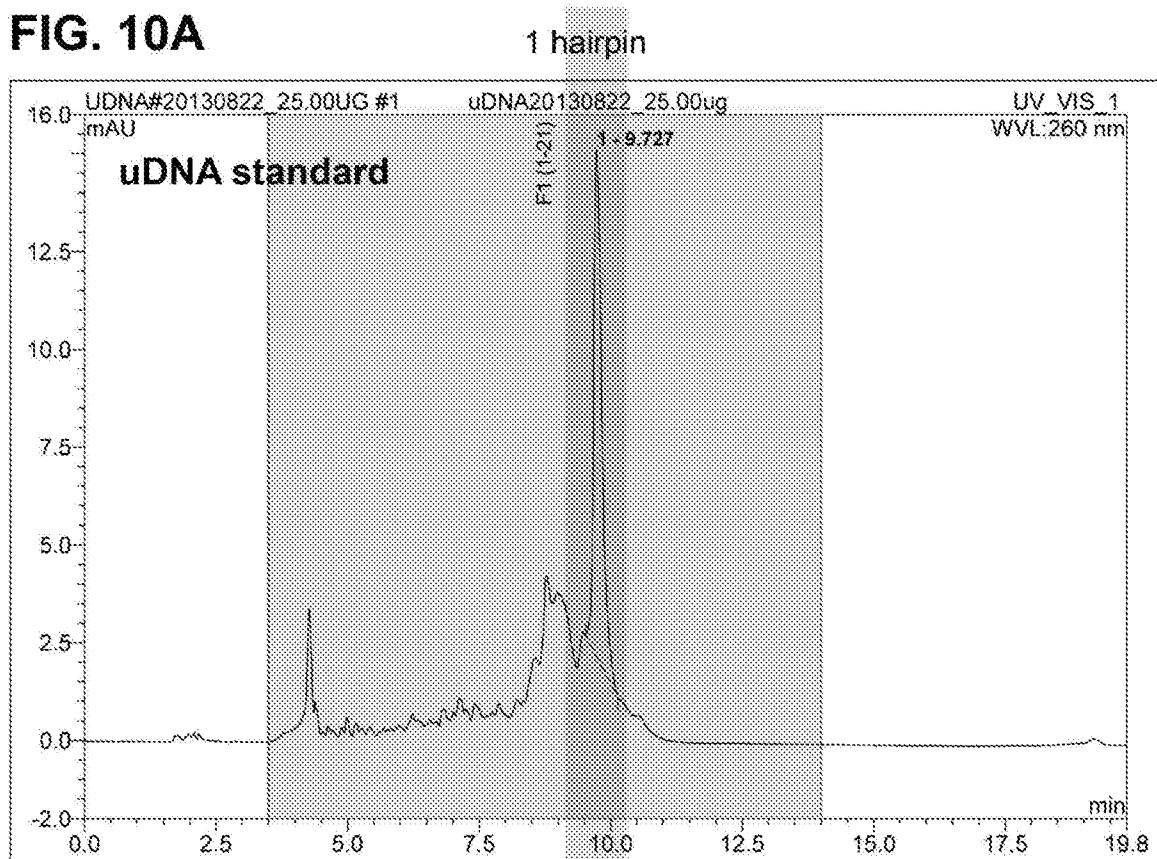
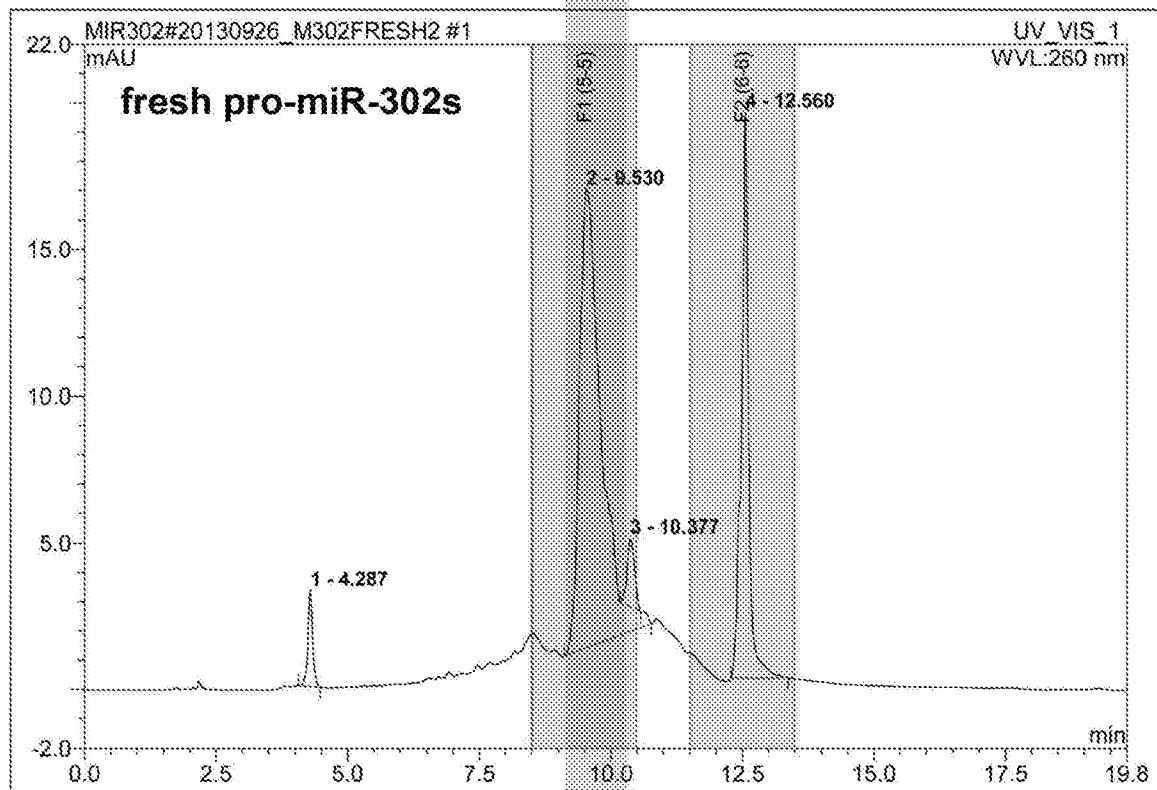
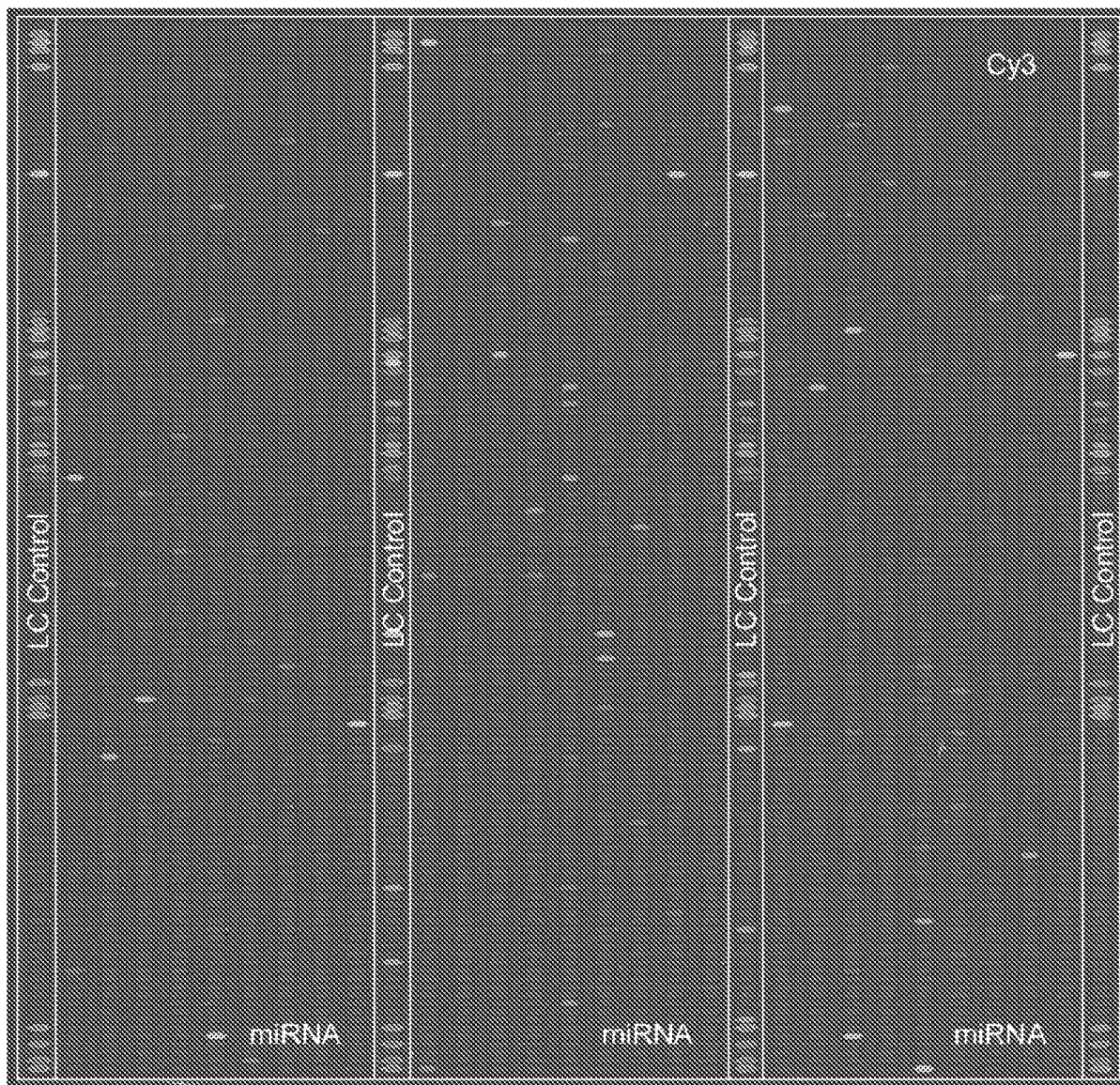


FIG. 10B



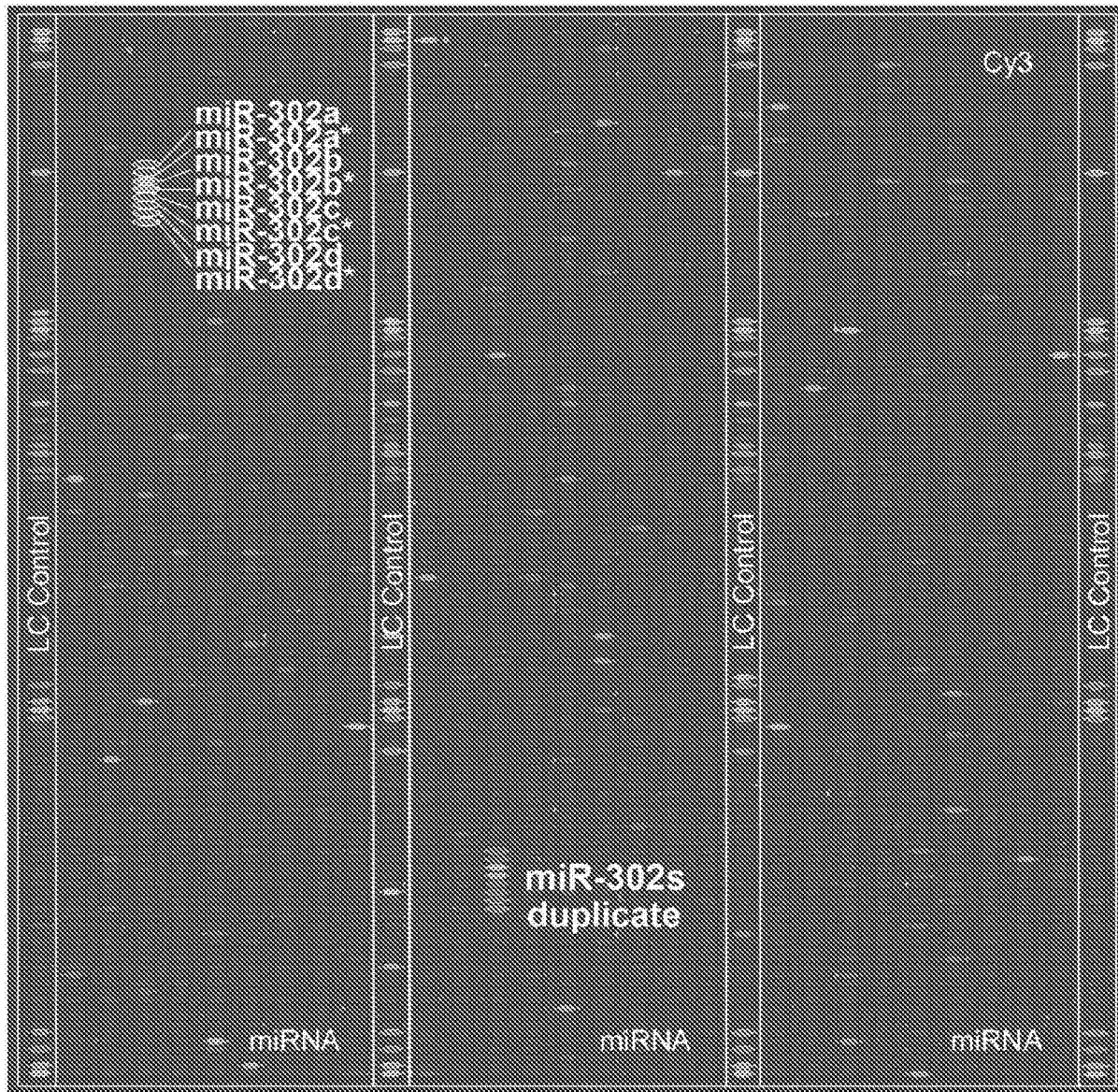
**FIG. 11A**

small RNAs extracted from blank *E. coli* competent cells (control):



**FIG. 11B**

small RNAs extracted from RGFP-miR302 transformed *E. coli* cells:



**FIG. 12**

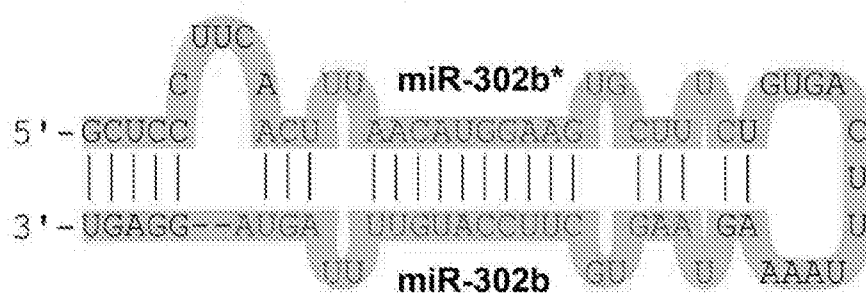
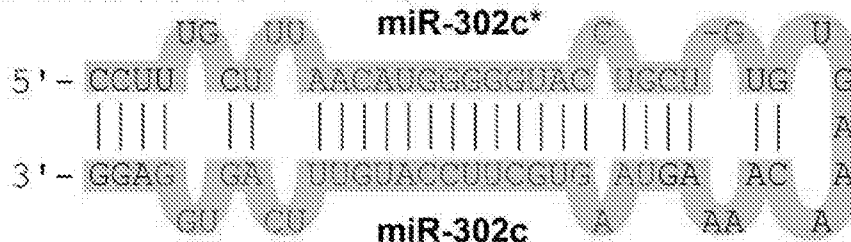
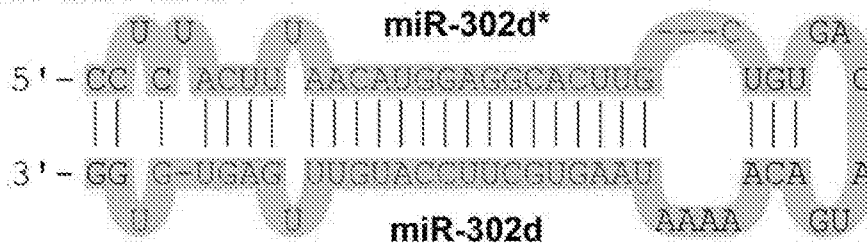
FIG.11A      FIG.11B

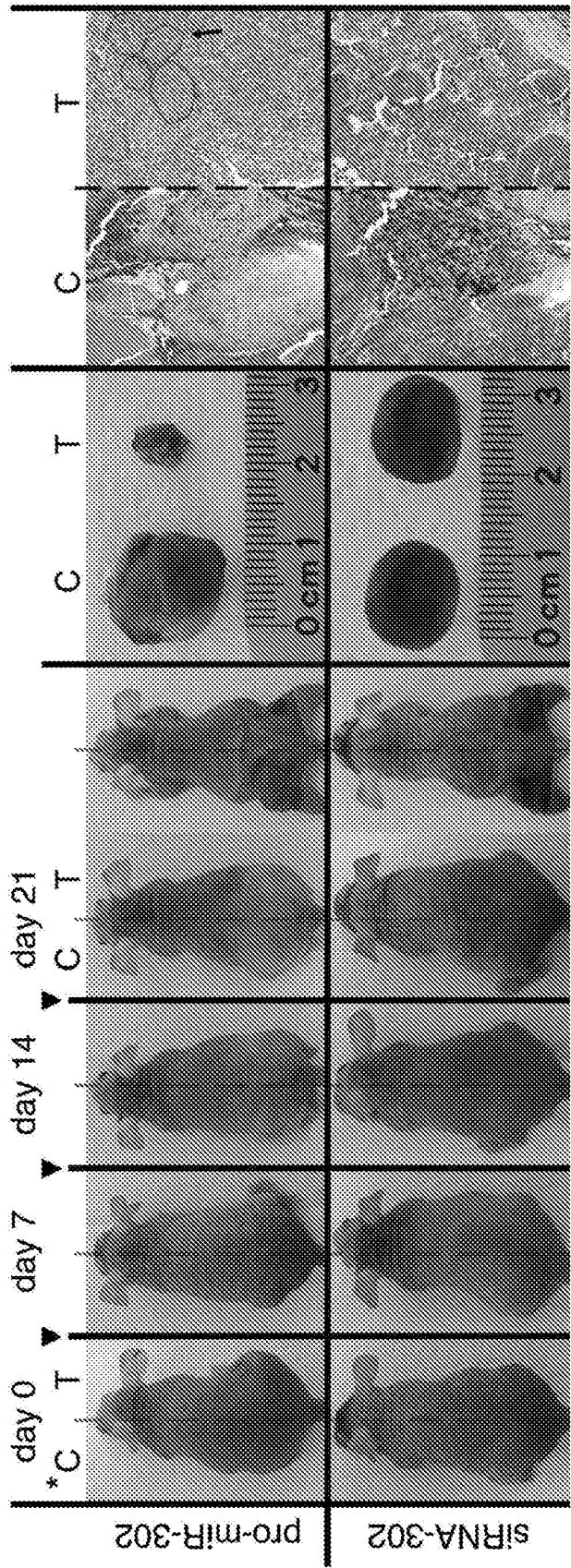
			Group 1	Group 2	
			#1	#2	
Rpter Index	Reporter Name	p-value	Mean	Mean	Log2 (G2/G1)
445	hsa-miR-302c-5p	1.28E-03	0	52,804	22.65
443	hsa-miR-302a-5p	1.38E-03	0	26,579	21.85
441	hsa-miR-302a-3p	1.52E-03	0	66,718	22.92
444	hsa-miR-302c-3p	3.51E-03	0	15,973	20.71
447	hsa-miR-302d-5p	2.23E-02	3	22,124	12.87
313	hsa-miR-302b-5p	3.09E-02	1	940	3.75
442	hsa-miR-302d-3p	4.86E-02	0	9,579	18.53
591	hsa-miR-320b-3p	4.88E-02	6	1,963	2.23
131	hsa-miR-1275	8.78E-02	749	364	-1.22
Following transcripts are not statistically significant due to low signals (signal < 500)					
554	hsa-miR-3177-3p	3.06E-02	402	76	-2.40
986	hsa-miR-4459	3.59E-02	66	135	1.04
967	hsa-miR-4442	4.52E-02	32	65	1.02
1298	hsa-miR-4758-5p	4.63E-02	7	74	3.45
993	hsa-miR-4466	5.70E-02	58	137	1.23
1120	hsa-miR-4649-5p	5.76E-02	187	385	1.04
1500	hsa-miR-5195-3p	5.81E-02	49	316	2.68
851	hsa-miR-4253	5.86E-02	296	136	-1.12
299	hsa-miR-146a-5p	5.93E-02	7	62	3.21
1156	hsa-miR-4668-5p	7.08E-02	315	70	-2.17
974	hsa-miR-4447	7.41E-02	13	89	2.76
1251	hsa-miR-4728-3p	7.96E-02	42	16	-1.42
750	hsa-miR-146a-3p	8.97E-02	23	88	1.92
870	hsa-miR-4270	9.39E-02	29	113	1.95
1123	hsa-miR-4651	9.46E-02	25	116	2.19

**FIG. 13A**

Sequence of the miR-302 familial cluster (SEQ.ID.NO.13)

1	AATTTTTC TTCTAAAGTT ATGCCATTTT GTTTCTTTC TCCTCAGCTC TAAATACTCT	sequence of
61	GAAGTCCAAA GAAGTTGTAT GTTGGGTGGG CTCCCTTCAA CTTTAACATG GAAGTGCTTT	pro-miR-302b
121	CTGTGACTTT AAAAGTAAGT GCTTCCATGT TTTAGTAGGA GTGAATCCAA TTTACTTCTC	
181	CAAAATAGAA CACGCTAACC TCATTGAAG GGATCCCCTT TGCTTTAACA TGGGGGTACC	pro-miR-302c
241	TGCTGTGTGA AACAAAAGTA AGTGCTTCCA TGTTTCAGTG GAGGTGTCTC CAAGCCAGCA	
301	CACCTTTTGT TACAAAATTT TTTTGTATT GTGTTTAAAG GTTACTAAGC TTGTACAGG	
361	TTAAAGGATT CTAACTTTTT CCAAGACTGG GCTCCCACC ACTTAAACGT GGATGTACTT	pro-miR-302a
421	GCTTTGAAAC TAAAGAAGTA AGTGCTTCCA TGTTTGGTG ATGGTAAGTC TTCTTTTAC	
481	ATTTTATTA TTTTTTAGA AAATAACTTT ATTGTATGA CCGCAGCTCA TATATTTAAG	
541	CTTTATTTTG TATTTTACA TCTGTAAAGG GGCCCCCTCT ACTTTAACAT GGAGGCACTT	pro-miR-302d
601	GCTGTGACAT GACAAAAATA AGTGCTTCCA TGTTGAGTG TGGTGGTCC TACCTAATCA	
661	GCAATTGAGT TAACGCCAC ACTGTGTGCA GTTCTGGCT ACAGGCCATT ACTGTTGCTA	

**FIG. 13B****pro-miR-302a (SEQ.ID.NO.6)****pro-miR-302b (SEQ.ID.NO.7)****pro-miR-302c (SEQ.ID.NO.8)****pro-miR-302d (SEQ.ID.NO.9)**



\*C = cancer without treatment; T = cancer with drug treatments

FIG. 14

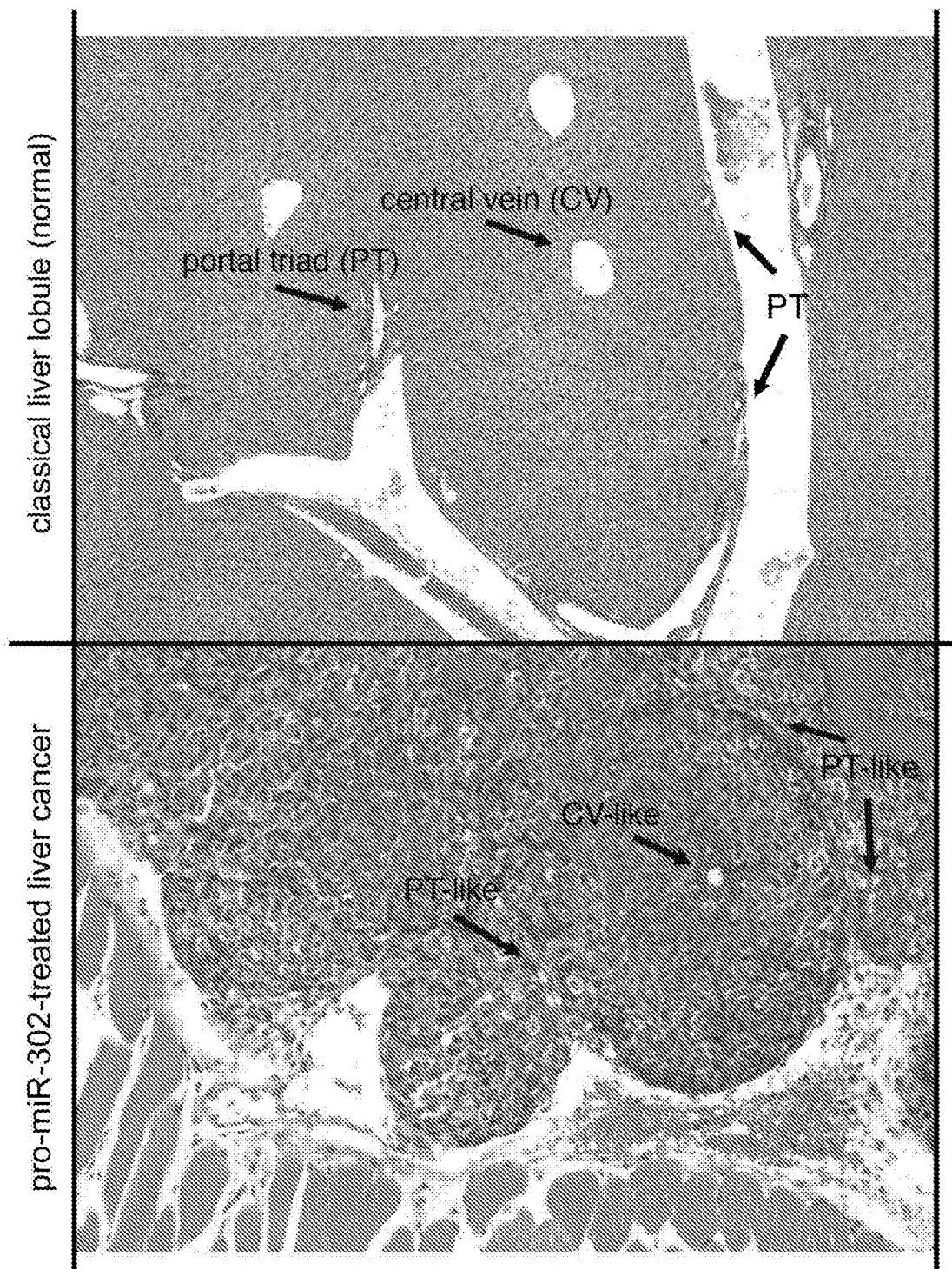


FIG. 15

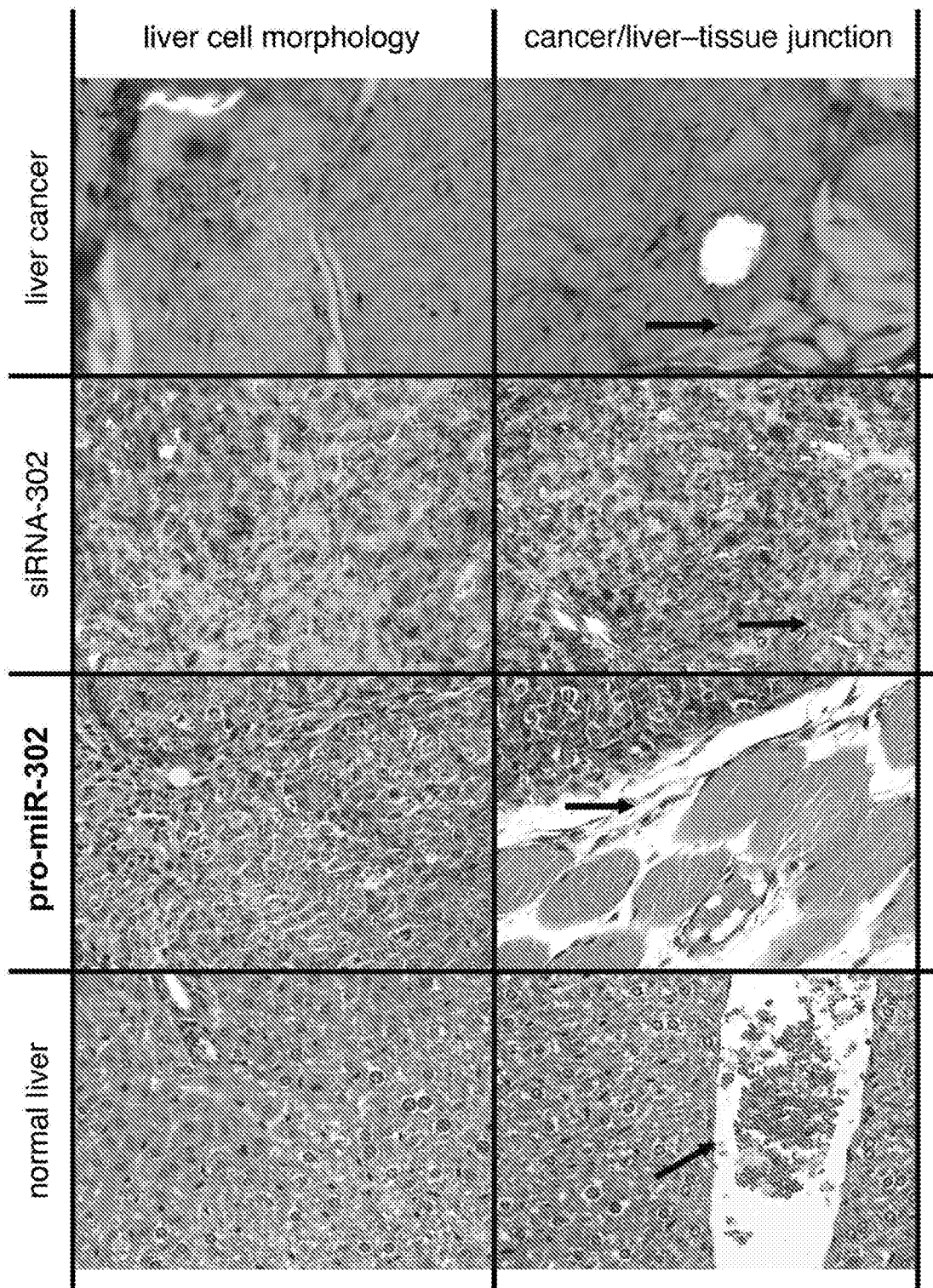
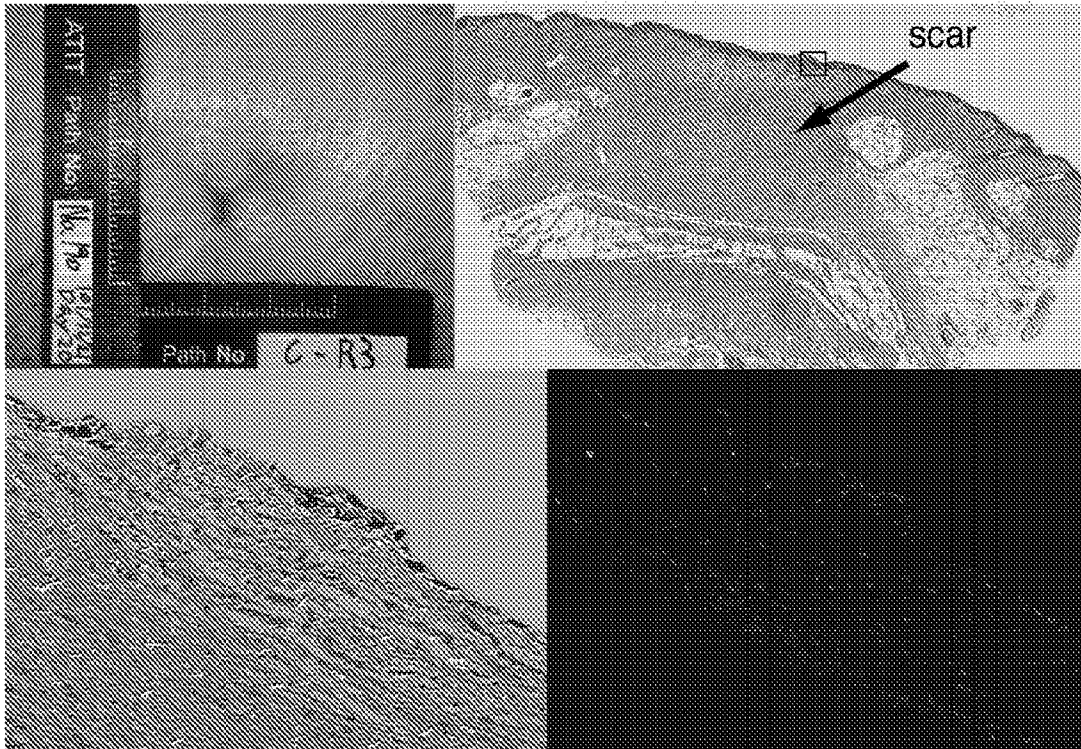
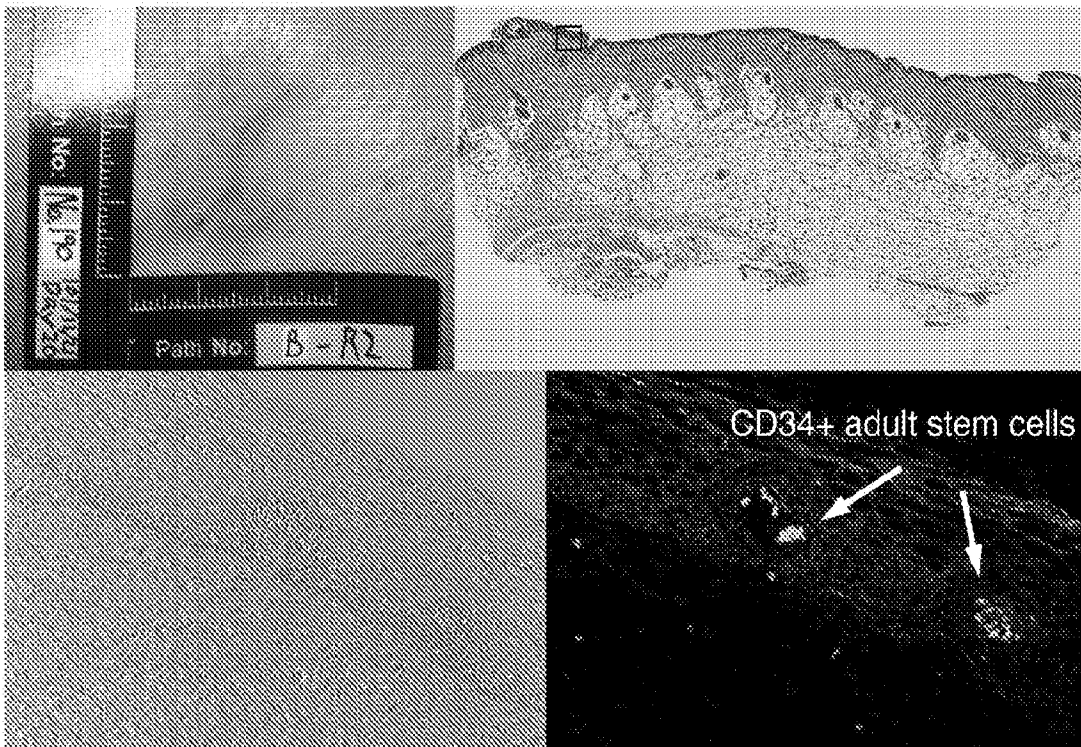


FIG. 16

**FIG. 17A** Example of control section shows large area of scar tissue underneath (190-CR3)



**FIG. 17B** Example of treated section shows no scar tissue underneath (190-BR2)



## INTERNATIONAL SEARCH REPORT

International application No

PCT/US2016/049583

## A. CLASSIFICATION OF SUBJECT MATTER

INV. C12N15/113 A61K31/713 C12N5/10 C12N5/074  
 ADD. A61P25/28 A61P39/00

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

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

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

EPO-Internal, BIOSIS, EMBASE, WPI Data

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	MING YAN ET AL: "Modulation of Gene Expression by Polymer Nanocapsule Delivery of DNA Cassettes Encoding Small RNAs", PLOS ONE, vol. 10, no. 6, 2 June 2015 (2015-06-02), page e0127986, XP055319413, DOI: 10.1371/journal.pone.0127986 the whole document	1-15
Y	WO 2013/025248 A1 (MELLO BIOTECHNOLOGY INC [US]; LIN SHI-LUNG [US]; CHANG DONALD C [US]) 21 February 2013 (2013-02-21) cited in the application the whole document	1-15
	----- -/--	



Further documents are listed in the continuation of Box C.



See patent family annex.

\* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search

23 November 2016

Date of mailing of the international search report

02/12/2016

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel. (+31-70) 340-2040,  
 Fax: (+31-70) 340-3016

Authorized officer

Andres, Serge

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2016/049583

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

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
  - a. ☒ forming part of the international application as filed:
    - ☒ in the form of an Annex C/ST.25 text file.
    - ☐ on paper or in the form of an image file.
  - b. ☐ furnished together with the international application under PCT Rule 13~~ter~~.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
  - c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
    - ☐ in the form of an Annex C/ST.25 text file (Rule 13~~ter~~.1(a)).
    - ☐ on paper or in the form of an image file (Rule 13~~ter~~.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

## INTERNATIONAL SEARCH REPORT

International application No

PCT/US2016/049583

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2014/106011 A1 (LIN SHI-LUNG [US]; CHANG DONALD [US]; WU DAVID TS) 3 July 2014 (2014-07-03) the whole document -----	1-15
A	US 2015/132805 A1 (WU DAVID TS [TW] ET AL) 14 May 2015 (2015-05-14) the whole document -----	1-15
A	BAOJIAN LIAO ET AL: "MicroRNA Cluster 302-367 Enhances Somatic Cell Reprogramming by Accelerating a Mesenchymal-to-Epithelial Transition", JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 286, no. 19, 22 March 2011 (2011-03-22), pages 17359-17364, XP055320364, US ISSN: 0021-9258, DOI: 10.1074/jbc.C111.235960 the whole document -----	1-15
E	US 2016/264974 A1 (LIN SHI-LUNG [US] ET AL) 15 September 2016 (2016-09-15) the whole document -----	1-15

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2016/049583

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2013025248 A1	21-02-2013	EP 2742127 A1	18-06-2014
		JP 2014530596 A	20-11-2014
		TW 201321507 A	01-06-2013
		US 2013210120 A1	15-08-2013
		US 2015064771 A1	05-03-2015
		US 2015118734 A1	30-04-2015
		WO 2013025248 A1	21-02-2013
-----			
WO 2014106011 A1	03-07-2014	US 2013324590 A1	05-12-2013
		WO 2014106011 A1	03-07-2014
-----			
US 2015132805 A1	14-05-2015	NONE	
-----			
US 2016264974 A1	15-09-2016	NONE	
-----			