

METHODS, SYSTEMS AND COMPOSITIONS RELATING
TO CELL CONVERSION VIA PROTEIN-INDUCED IN-VIVO CELL REPROGRAMMING

GRANT REFERENCE

[0001] This invention was made with government support under Grant No. 1 R01 CA172480-01A1, awarded by the National Institutes of Health/National Cancer Institute. The Government has certain rights in the invention.

FIELD OF THE INVENTION

[0002] The present invention relates generally to methods, systems and compositions for treatment of a pathological condition in a subject in need thereof. In specific aspects, the present invention relates to methods, systems and compositions for protein-induction of cell conversion in vivo for treatment of a pathological condition in a subject in need thereof.

BACKGROUND OF THE INVENTION

[0003] Despite recent medical progress, there is a continuing need for methods and compositions for treatment of disease and injury.

SUMMARY

[0004] Methods of treating a subject in need thereof are provided according to embodiments of the present disclosure which include administering a pharmaceutical composition comprising a protein transduction reagent-modified reprogramming protein to the subject, wherein the protein transduction reagent is non-covalently bound to the reprogramming protein and wherein the protein transduction reagent comprises a cation reagent and a lipid.

[0004a] In one aspect, there is provided a pharmaceutical composition comprising a protein transduction reagent-modified reprogramming protein for use in therapy, wherein the protein transduction reagent is non-covalently bound to the reprogramming protein and wherein the protein transduction reagent comprises a cation reagent and a lipid, wherein the cation reagent is a polyethylenimine (PEI); and wherein the lipid is selected from the group consisting of DOTMA, (N-1-(2,3-dioleoyloxy)propyl-N,N,N-trimethyl-ammonium chloride; DOGS, (dioctadecylamido-glycylspermine); DOTAP, 1,2-dioleoyl-3-trimethylammonium-propane; DOPE, 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine; POPC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; DMPE 1,2-dimyristoyl-sn-glycero-3-phosphocholine and combinations thereof.

[0004b] In another aspect, there is provided use of a protein transduction reagent-modified reprogramming protein in the manufacture of a medicament for the treatment of a condition characterized by damaged, and/or defective cells in a subject, wherein the protein transduction reagent is non-covalently bound to the reprogramming protein and wherein the protein transduction reagent comprises a cation reagent and a lipid; wherein the cation reagent is a polyethylenimine (PEI); and wherein the lipid is selected from the group consisting of DOTMA, (N-1(-(2,3-dioleoyloxy)propyl-N,N,N-trimethyl-ammonium chloride; DOGS, (dioctadecylamido-glycylspermine); DOTAP, 1,2-dioleoyl-3-trimethylammonium-propane; DOPE, 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine; POPC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; DMPE 1,2-dimyristoyl-sn-glycero-3-phosphocholine and combinations thereof; and wherein the condition is selected from the group consisting of cancer, heart disease, liver disease, atherosclerosis, neurological disease, neurodegenerative disease, blood disease, diabetes, pancreatic disease, obesity, muscle disease, breast disease and vascular disease.

[0005] Methods of treating a subject in need thereof are provided according to embodiments of the present disclosure which include administering a pharmaceutical composition comprising a protein transduction reagent-modified reprogramming protein to the subject, wherein the protein transduction reagent is non-covalently bound to the reprogramming protein and wherein the protein transduction reagent comprises a cation reagent and a lipid, wherein the subject has a disease selected from the group consisting of: cancer; pancreatic disease or injury; heart disease or heart injury such as acute myocardial infarction, chronic myocardial infarction and heart failure; liver injury or liver disease such as familial hypercholesterolaemia (FH), Crigler-Najjar syndrome and hereditary tryosinemia I; atherosclerosis; neurological disease or injury such as spinal cord

injury, traumatic brain injury, amyotrophic lateral sclerosis, spinal muscular atrophy and Parkinson's disease; arthritis; joint disease or injury; blood disease; diabetes; obesity; muscle disease or injury; cartilage disease or injury; breast disease or injury; and vascular disease or injury.

[0006] Methods of treating a subject having a condition characterized by damaged and/or defective cells are provided according to embodiments of the present disclosure including: administering a pharmaceutical composition including a protein transduction reagent-modified reprogramming protein to the subject, wherein the protein transduction reagent is non-covalently bound to the reprogramming protein and wherein the protein transduction reagent comprises a cation reagent and a lipid.

[0007] Methods of treating a subject having cancer including: administering a pharmaceutical composition including a protein transduction reagent-modified reprogramming protein to the subject, wherein the protein transduction reagent is non-covalently bound to the reprogramming protein and wherein the protein transduction reagent comprises a cation reagent and a lipid.

[0008] Methods of treating a subject having cancer are provided according to embodiments of the present disclosure which include administering a pharmaceutical composition including a protein transduction reagent-modified reprogramming protein selected from the group consisting of: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog.

[0009] Methods of treating a subject having cancer are provided according to embodiments of the present disclosure which include administering a pharmaceutical composition including protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog.

[0010] Methods of treating a subject having a brain tumor or breast cancer and are provided according to embodiments of the present disclosure which include administering a pharmaceutical composition including a protein transduction reagent-modified reprogramming protein selected from the group consisting of: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog.

[0011] Methods of treating a subject having a brain tumor or breast cancer and are provided according to embodiments of the present disclosure which include administering a pharmaceutical composition including protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog.

[0012] Methods of treating a subject having a condition characterized by damaged and/or defective cells are provided according to embodiments of the present disclosure including: administering a pharmaceutical composition including a protein transduction reagent-modified reprogramming protein to the subject, wherein the protein transduction reagent is non-covalently bound to the reprogramming protein and wherein the protein transduction reagent comprises a cation reagent and a lipid, wherein the condition is a heart disease or heart damage and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Gata4, protein transduction reagent-modified Hand2, protein transduction reagent-modified MEF2c and protein transduction reagent-modified Tbx5; wherein the condition is a liver disease or liver damage and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Hnf1a, protein transduction reagent-modified Foxa1, protein transduction reagent-modified Foxa2 and protein transduction reagent-modified Foxa3; wherein the condition is a liver disease or liver damage and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Gata4, protein transduction reagent-modified Hnf1a and protein transduction reagent-modified Foxa3; wherein the condition is atherosclerosis and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified CEBP α , protein transduction reagent-modified CEBP β and protein transduction reagent-modified PU.1; wherein the condition is a neurodegenerative disease or neuronal tissue damage and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Sox2 and protein transduction reagent-modified Foxg1; one or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Lmx1a, protein transduction reagent-modified Nurr1, protein transduction reagent-modified Brn2 and protein transduction reagent-modified Myt1l; one or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Myt1l and protein transduction reagent-modified NeuroD1; protein transduction reagent-modified Ngn2; one or both of: protein transduction reagent-modified Sox2 and protein transduction reagent-modified NeuroD1; one or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Asc1, protein transduction reagent-modified Myt1l, protein transduction reagent-modified Lhx3 and protein transduction reagent-modified Hb9; one or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Asc1, protein transduction reagent-modified Myt1l, protein transduction reagent-modified Lhx3, protein transduction reagent-modified Hb9, protein transduction reagent-modified Lsl1 and protein transduction reagent-modified Ngn2; protein transduction reagent-modified Dlx2; protein

transduction reagent-modified Dlx2 and protein transduction reagent-modified Ascl1; one or more of: protein transduction reagent-modified Ascl1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Myt1l, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Foxa2; or one or more of: protein transduction reagent-modified Ascl1, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Nurr1; wherein the condition is a disease or disorder of the blood and the pharmaceutical composition comprises: protein transduction reagent-modified Oct4; wherein the condition is diabetes, a pancreatic disease or pancreatic tissue damage and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Ngn3, protein transduction reagent-modified Pdx1 and protein transduction reagent-modified Pax4 or one or more of: protein transduction reagent-modified Ngn3, protein transduction reagent-modified Pdx1 and protein transduction reagent-modified MafA; wherein the condition is obesity and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Prdm16 and protein transduction reagent-modified C/EBPb; wherein the condition is a muscle disease or muscle damage and the pharmaceutical composition comprises protein transduction reagent-modified MyoD; wherein the condition is arthritis or joint disease or injury and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Runx2, protein transduction reagent-modified Sox5 and protein transduction reagent-modified Sox6; wherein the condition is a breast disease or breast tissue damage and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Slug, protein transduction reagent-modified Stat5a, protein transduction reagent-modified Gata3 and protein transduction reagent-modified Brca-11; wherein the condition is a vascular disease or blood vessel damage and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Erg1, protein transduction reagent-modified Er71, protein transduction reagent-modified Fli1 and protein transduction reagent-modified Gata2 or one or more of: protein transduction reagent-modified Hey1, protein transduction reagent-modified Hey2, protein transduction reagent-modified FoxC1 and protein transduction reagent-modified FoxC2 or one or both of: protein transduction reagent-modified Sox7 and protein transduction reagent-modified Sox18.

[0012a] A pharmaceutical composition is provided in the present disclosure, comprising a protein transduction reagent-modified reprogramming protein.

[0013] Pharmaceutical compositions are provided according to embodiments of the present disclosure which include one or more protein transduction reagent-reprogramming proteins selected from the group consisting of: protein transduction reagent-modified Gata4, protein transduction reagent-modified Hand2, protein transduction reagent-modified MEF2c, protein transduction

reagent-modified Tbox, protein transduction reagent-modified Hnf1a, protein transduction reagent-modified Foxa1, protein transduction reagent-modified Foxa2 and protein transduction reagent-modified Foxa3, protein transduction reagent-modified Hnf1a, protein transduction reagent-modified CEBP α , protein transduction reagent-modified CEBP β , protein transduction reagent-modified PU.1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Sox2 and protein transduction reagent-modified Foxg1, protein transduction reagent-modified Asc1, protein transduction reagent-modified Lmx1a, protein transduction reagent-modified Nurr1, protein transduction reagent-modified Myt1l, protein transduction reagent-modified NeuroD1, protein transduction reagent-modified Ngn2, protein transduction reagent-modified Lhx3, protein transduction reagent-modified Hb9, protein transduction reagent-modified Lsl1, protein transduction reagent-modified Dlx2, protein transduction reagent-modified Asc1, protein transduction reagent-modified Oct4, protein transduction reagent-modified Ngn3, protein transduction reagent-modified Pdx1, protein transduction reagent-modified Pax4, protein transduction reagent-modified MafA, protein transduction reagent-modified Prdm16, protein transduction reagent-modified MyoD, protein transduction reagent-modified Sox9, protein transduction reagent-modified Runx2, protein transduction reagent-modified Sox5 and protein transduction reagent-modified Sox6, protein transduction reagent-modified Slug, protein transduction reagent-modified Stat5a, protein transduction reagent-modified Gata3, protein transduction reagent-modified Brca-1l, protein transduction reagent-modified Erg1, protein transduction reagent-modified Er71, protein transduction reagent-modified Fli1, protein transduction reagent-modified Gata2, protein transduction reagent-modified Hey1, protein transduction reagent-modified Hey2, protein transduction reagent-modified FoxC1 and protein transduction reagent-modified FoxC2, protein transduction reagent-modified Sox7 and protein transduction reagent-modified Sox18.

[0014] Pharmaceutical compositions are provided according to embodiments of the present disclosure which include two or more protein transduction reagent-modified reprogramming proteins selected from the group consisting of: protein transduction reagent-modified Gata4, protein transduction reagent-modified Hand2, protein transduction reagent-modified MEF2c and protein transduction reagent-modified Tbox5; two or more of: protein transduction reagent-modified Hnf1a, protein transduction reagent-modified Foxa1, protein transduction reagent-modified Foxa2 and protein transduction reagent-modified Foxa3; two or more of: protein transduction reagent-modified Gata4, protein transduction reagent-modified Hnf1a and protein transduction reagent-modified Foxa3; protein transduction reagent-modified CEBP α , protein

transduction reagent-modified CEBP β , and protein transduction reagent-modified PU.1; two or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Sox2 and protein transduction reagent-modified Foxg1; two or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Lmx1a, protein transduction reagent-modified Nurr1, protein transduction reagent-modified Brn2 and protein transduction reagent-modified Myt1l; two or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Myt1l and protein transduction reagent-modified NeuroD1; protein transduction reagent-modified Ngn2; protein transduction reagent-modified Sox2 and protein transduction reagent-modified NeuroD1; two or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Asc1, protein transduction reagent-modified Myt1l, protein transduction reagent-modified Lhx3 and protein transduction reagent-modified Hb9; two or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Asc1, protein transduction reagent-modified Myt1l, protein transduction reagent-modified Lhx3, protein transduction reagent-modified Hb9, protein transduction reagent-modified Lsl1 and protein transduction reagent-modified Ngn2; protein transduction reagent-modified Dlx2; protein transduction reagent-modified Dlx2 and protein transduction reagent-modified Asc1; two or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Myt1l, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Foxa2; two or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Nurr1; protein transduction reagent-modified Ngn3, protein transduction reagent-modified Pdx1 and protein transduction reagent-modified Pax4; two or more of: protein transduction reagent-modified Ngn3, protein transduction reagent-modified Pdx1 and protein transduction reagent-modified MafA; protein transduction reagent-modified Prdm16 and protein transduction reagent-modified C/EBP β ; two or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Runx2, protein transduction reagent-modified Sox5 and protein transduction reagent-modified Sox6; two or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Slug, protein transduction reagent-modified Stat5a, protein transduction reagent-modified Gata3 and protein transduction reagent-modified Brca-11; two or more of: protein transduction reagent-modified Erg1, protein transduction reagent-modified Er71, protein transduction reagent-modified Fli1 and protein transduction reagent-modified Gata2; two or more of: protein transduction reagent-modified Hey1, protein transduction reagent-modified Hey2, protein transduction reagent-

modified FoxC1 and protein transduction reagent-modified FoxC2; or protein transduction reagent-modified Sox7 and protein transduction reagent-modified Sox18.

[0015] Methods according to the present disclosure allow of delivery of protein-transduction reagent-modified reprogramming proteins to cancer cells, such as tumor cells, as well as diseased cells of diseased tissues.

[0016] Methods according to embodiments of the present disclosure provide in vivo conversion of diseased cells into normal cells via protein-induced in situ cell reprogramming. Methods according to embodiments of the present disclosure provide in vivo conversion of cancer cells into normal cells via protein-induced in situ cell reprogramming.

[0017] Use of protein-transduction reagent-modified reprogramming proteins to reprogram cancer cells or diseased cells in situ into stem cells or transient protein-induced multipotent stem cells that are then induced to differentiate into normal cells in the tissue where the cancer cells or diseased cells were located is provided according to embodiments of the present disclosure.

[0018] Methods of treating cancer are provided according to embodiments of the present disclosure which treat the cancer in vivo not by killing the cancer cells but instead by converting the cancer cells into non-cancerous cells by administration, such as systemic or local administration, of one or more protein transduction reagent-modified reprogramming proteins.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019] Figure 1A is a representative MRI image of a 9L-intracranial tumor bearing rat brain showing enhanced permeability and retention effect in a 9L brain tumor for QQ-ferritin showing that the intravenous (i.v.) injected QQ-ferritin was delivered into 9L brain tumor via enhanced permeability and retention effect, causing an enhanced negative MRI image of brain tumor in the 9L-brain tumor bearing rat;

[0020] Figure 1B is a representative MRI image of a 9L-intracranial tumor bearing rat brain showing no enhanced permeability and retention effect in a 9L brain tumor for ferritin without QQ modification showing that the i.v. injected ferritin did not reach into 9L brain tumor and no enhanced negative MRI image of brain tumor was observed in the 9L-brain tumor bearing rat;

[0021] Figure 2A is a representative MRI image of a mouse bearing a 4T1 breast tumor prior to a tail-vein injection of QQ-ferritin;

[0022] Figure 2B is a representative MRI image of a mouse bearing a 4T1 breast tumor 0.5 hours after a tail-vein injection of QQ-ferritin;

[0023] Figure 2C is a representative MRI image of a mouse bearing a 4T1 breast tumor 2

hours after a tail-vein injection of QQ-ferritin;

[0024] Figure 2D is a representative MRI image of a mouse bearing a 4T1 breast tumor 3.5 hours after a tail-vein injection of QQ-ferritin;

[0025] Figure 2E is a representative MRI image of a mouse bearing a 4T1 breast tumor 8 hours after a tail-vein injection of QQ-ferritin;

[0026] Figure 2F is a representative MRI image of a mouse bearing a 4T1 breast tumor 18 hours after a tail-vein injection of QQ-ferritin;

[0027] Figure 3A is a graph showing proliferation of U251 cells and U251-piPSCs monitored by Ki67 assay;

[0028] Figure 3B is a graph showing dose-dependent chemotherapeutic drug sensitivity of patient-derived primary human GBM (12-14) cells (squares) and GBM (12-14)-piPSCs (circles) against the alkylating agent temozolomide;

[0029] Figure 3C is a graph showing dose-dependent chemotherapeutic drug sensitivity of 9L-cells (squares) and 9L-piPSCs (circles) against carboplatin;

[0030] Figure 4A is an image showing Western blots of several epithelial and mesenchymal markers of U251 cells and U251-piPSCs, including E-cadherin (E-cad), β -catenin (β -cat), vimentin (VMT), indicating a mesenchymal-to-epithelial transition (MET) during cell reprogramming. In addition, cell reprogramming of 9L-cells into 9L-piPSCs also caused an enhanced expression of glial fibrillary acidic protein (GFAP), a marker for astrocytes;

[0031] Figure 4B is a graph showing that the tumorigenicity reduction of U251 cells after cell reprogramming according to aspects of the present invention is confirmed by in vitro assays for cell migration of U251-cells and U251-piPSCs, indicating significantly reduced migration of U251-piPSCs ($p < 0.01$) in vitro;

[0032] Figure 4C is a graph showing that the tumorigenicity reduction of U251 cells after cell reprogramming according to aspects of the present invention is confirmed by in vitro assays for invasion of U251-cells and U251-piPSCs, indicating significantly reduced invasion by U251-piPSCs ($p < 0.01$) in vitro;

[0033] Figure 4D is a graph showing reduced proliferation of 4T1-piPSCs and 67NR-piPSCs as compared to those properties of the parental 4T1 cells and 67NR cells, indicating a significantly reduced tumorigenicity of 4T1 cells and 67NR cells after cell reprogramming into 4T1-piPSCs and 67NR-piPSCs in vitro;

[0034] Figure 4E is a graph showing reduced mammary sphere formation of 4T1-piPSCs as compared to those properties of the parental 4T1 cells, indicating a significantly reduced

tumorigenicity of 4T1 cells after cell reprogramming into 4T1-piPSCs in vitro;

[0035] Figure 4F is a graph showing reduced migration of 4T1-piPSCs as compared to those properties of the parental 4T1 cells, indicating a significantly reduced tumorigenicity of 4T1 cells after cell reprogramming into 4T1-piPSCs in vitro;

5 [0036] Figure 4G is a graph showing reduced invasion of 4T1-piPSCs as compared to those properties of the parental 4T1 cells, indicating a significantly reduced tumorigenicity of 4T1 cells after cell reprogramming into 4T1-piPSCs in vitro;

[0037] Figure 5A is an image showing cell morphology of U251 cells;

[0038] Figure 5B is an image showing changes in cell morphology of U251 cells indirectly
10 co-cultured with U251-piPSCs for 40-64 hours, indicative of mesenchymal-to-epithelial transition and demonstrating a bystander effect of the U251-piPSCs on U251 cells;

[0039] Figure 5C is a graph showing significantly reduced proliferation of U251-piPSCs and U251 cells indirectly co-cultured with U251-piPSCs at a 1:1 ratio of the U251-piPSCs and U251 cells, indicating a significantly reduced tumorigenicity of the co-cultured U251 cells and
15 supporting the observation of a bystander effect;

[0040] Figure 5D is a graph showing significantly reduced migration of U251-piPSCs and U251 cells indirectly co-cultured with U251-piPSCs at 1:1 and 8:1 ratios (U251:U251-piPSCs), ($p < 0.005$), indicating a significantly reduced tumorigenicity of the co-cultured U251 cells and supporting the observation of a bystander effect;

20 [0041] Figure 5E is a graph showing significantly reduced invasion of U251-piPSCs and U251 cells indirectly co-cultured with U251-piPSCs at 1:1 and 8:1 ratios (U251:U251-piPSCs), ($p < 0.005$), indicating a significantly reduced tumorigenicity of the co-cultured U251 cells and supporting the observation of a bystander effect;

[0042] Figure 5F is a graph showing significantly reduced proliferation of 4T1 cells co-
25 cultured with 4T1-piPSCs for 40 hours, indicating a significantly reduced tumorigenicity of the co-cultured 4T1 cells and supporting the observation of a bystander effect;

[0043] Figure 5G is a graph showing significantly reduced migration of 4T1 cells co-cultured with 4T1-piPSCs for 40 hours, indicating a significantly reduced tumorigenicity of the co-cultured 4T1 cells and supporting the observation of a bystander effect;

30 [0044] Figure 5H is a graph showing significantly reduced invasion of 4T1 cells co-cultured with 4T1-piPSCs for 40 hours, indicating a significantly reduced tumorigenicity of the co-cultured 4T1 cells and supporting the observation of a bystander effect;

[0045] Figure 6A is a graph showing tumor volume (mm³) in rats intracranially-implanted

with either 9L cells or 9L-piPSCs (n = 10) at day 14 and showing significantly reduced volume of the intracranial 9L-piPSC tumors as compared with those of 9L-tumors, indicating a significantly reduced tumorigenicity of 9L-cells after cell reprogramming into 9L-piPSCs in vivo;

[0046] Figure 6B is a graph showing tumor weight (grams) in rats subcutaneously-implanted with 9L cells and 9L-piPSCs at day 25 after implantation (n = 6) and showing significantly reduced weight of the subcutaneous 9L-piPSC tumors as compared with those of 9L-tumors, indicating a significantly reduced in vivo tumorigenicity of 9L-cells after cell reprogramming into 9L-piPSCs in vivo;

[0047] Figure 6C is a graph showing 4T1- and 4T1-piPSC tumor growth curve as measured using tumor volume over 25 days after 4T1- and 4T1-piPSC implantation, indicating a significant in vivo tumor stasis of 4T1 cells after cell reprogrammed into 4T1-piPSCs in vivo;

[0048] Figure 6D is a graph showing 4T1- and 4T1-piPSC tumor growth curve as measured using tumor weight at day 25 after 4T1- and 4T1-piPSC implantation, indicating a significant in vivo tumor stasis of 4T1 cells after cell reprogrammed into 4T1-piPSCs in vivo;

[0049] Figure 6E is a graph showing lung metastases in mice implanted with 4T1 cells or 4T1-piPSCs at day 25 after cell implantation, indicating a major in vivo metastasis inhibition caused by cell reprogramming of 4T1 cells into 4T1-piPSCs in vivo; the lung was divided into 200 fields and number of metastases was counted and reported in the y-axis;

[0050] Figure 6F is a Kaplan-Meier survival curve of mice implanted with 4T1 cells and 4T1-piPSCs in the #4 fat pads, demonstrating a significantly prolonged survival of the 4T1-piPSCs during this 250-day survival experiment, indicating that cell reprogramming of malignant cancer cells into piPSCs significantly reduces tumorigenicity and metastatic properties of the parental cancer cells;

[0051] Figure 7A is a graph showing dose-dependent 9L tumor growth curves measured by tumor volume in rats treated with QQ-reagent in PBS buffer (n = 10), QQ-SON proteins 1 µg/day (n = 5), QQ-SON proteins 5 µg/day (n = 5), and QQ-SON proteins 10 µg/day (n = 10) for 18 daily treatments where the treatment started at day 5 after 9L cell implantation and the rats were sacrificed at day 23;

[0052] Figure 7B is a graph showing 9L tumor weight in rats treated with QQ-reagent in PBS buffer (n = 10), QQ-SON proteins 1 µg/day (n = 5), QQ-SON proteins 5 µg/day (n = 5), and QQ-SON proteins 10 µg/day (n = 10) for 18 daily treatments where the treatment started at day 5 after 9L cell implantation and the rats were sacrificed at day 23;

[0053] Figure 7C is a Kaplan-Meier survival curve (130-days) of rats having subcutaneous

9L tumors treated with QQ-reagents in PBS (solid line, n = 6; median survival = 21 days) or QQ-SON proteins (dashed line, 10 μ g/day, n = 8; median survival = 127 days) for 30 daily treatments, the endpoint is when tumor volume reaches 12 cm³;

[0054] Figure 7D is a graph showing effects of various QQ-SON protein dosages on 4T1-tumor growth monitored by tumor volume in the breast over a 25-day time course;

[0055] Figure 7E is a graph showing effects of various QQ-SON protein dosages on 4T1-tumor weight in the breast at day 25;

[0056] Figure 7F is a graph showing changes of tumor volumes at different time points, determined by MRI imaging, of 4T1-tumor bearing mice treated either with QQ-SON proteins (n = 8) or QQ-PBS as a control (n = 8), showing major tumor stasis without primary tumor removal;

[0057] Figure 7G is a graph showing tumor weight of 4T1-tumor bearing mice treated either with QQ-SON proteins (n = 8) or QQ-PBS as the control (n = 8) at day 35, showing major tumor stasis without primary tumor removal;

[0058] Figure 7H is a Kaplan-Meier survival curve of 4T1-breast cancer bearing mice treated with QQ-PBS (Control, dotted line) and QQ-SON proteins (Treatment, solid line) during a 60-day survival experiment without primary 4T1 breast cancer removed as shown in Figure 7H. The treatment started at day 5 after 4T1-cell implantation. Mice that met the endpoint, including tumor size burden, labored breath, uncontrollable pain, etc. were sacrificed;

[0059] Figure 7I is a graph showing the percentage of QQ-SON or QQ-PBS treated 4T1 tumor-bearing mice having metastatic lesions in the lung, without primary tumor removal, as observed by MRI at indicated days, indicating a major metastasis inhibition in the 4T1-bearing mice caused by QQ-SON protein treatment without primary tumor removal;

[0060] Figure 7J is a graph showing the percentage of QQ-SON or QQ-PBS treated 4T1 tumor-bearing mice having metastatic lesions in lymph nodes, without primary tumor removal, as observed by MRI at indicated days, indicating a major metastasis inhibition in the 4T1-bearing mice caused by QQ-SON protein treatment without primary tumor removal;

[0061] Figure 7K is a graph showing the percentage of QQ-SON or QQ-PBS treated 4T1 tumor-bearing mice having metastatic lesions in the liver, without primary tumor removal, as observed by MRI at indicated days, indicating a major metastasis inhibition in the 4T1-bearing mice caused by QQ-SON protein treatment without primary tumor removal;

[0062] Figure 7L is a graph showing the percentage of QQ-SON or QQ-PBS treated 4T1 tumor-bearing mice having metastatic lesions in the spleen, without primary tumor removal, as observed by MRI at indicated days, indicating a major metastasis inhibition in the 4T1-bearing

mice caused by QQ-SON protein treatment without primary tumor removal;

[0063] Figure 7M is a graph showing the average number per mouse of QQ-SON or QQ-PBS treated 4T1 tumor -bearing mice having metastatic lesions in the lung, without primary tumor removal, as observed by MRI at indicated days, indicating a major metastasis inhibition in the 4T1-bearing mice caused by QQ-SON protein treatment without primary tumor removal;

[0064] Figure 7N is a graph showing the average number per mouse of QQ-SON or QQ-PBS treated 4T1 tumor-bearing mice having metastatic lesions in the lymph nodes, without primary tumor removal, as observed by MRI at indicated days, indicating a major metastasis inhibition in the 4T1-bearing mice caused by QQ-SON protein treatment without primary tumor removal;

[0065] Figure 7O is a graph showing the average number per mouse of QQ-SON or QQ-PBS treated 4T1 tumor-bearing mice having metastatic lesions in the liver, without primary tumor removal, as observed by MRI at indicated days, indicating a major metastasis inhibition in the 4T1-bearing mice caused by QQ-SON protein treatment without primary tumor removal;

[0066] Figure 7P is a graph showing the average number per mouse of QQ-SON or QQ-PBS treated 4T1 tumor-bearing mice having metastatic lesions in the spleen, without primary tumor removal, as observed by MRI at indicated days, indicating a major metastasis inhibition in the 4T1-bearing mice caused by QQ-SON protein treatment without primary tumor removal; and

[0067] Figure 8 is a Kaplan-Meier survival curve (250-day survival) of 4T1-breast cancer bearing mice after surgical removal of the primary 4T1-breast cancer at day 18 and QQ-SON protein treatment started at day 6.

DETAILED DESCRIPTION

[0068] Scientific and technical terms used herein are intended to have the meanings commonly understood by those of ordinary skill in the art. Such terms are found defined and used in context in various standard references illustratively including J. Sambrook and D.W. Russell, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press; 3rd Ed., 2001; F.M. Ausubel, Ed., Short Protocols in Molecular Biology, Current Protocols; 5th Ed., 2002; B. Alberts et al., Molecular Biology of the Cell, 4th Ed., Garland, 2002; D.L. Nelson and M.M. Cox, Lehninger Principles of Biochemistry, 4th Ed., W.H. Freeman & Company, 2004; Engelke, D.R., RNA Interference (RNAi): Nuts and Bolts of RNAi Technology, DNA Press LLC, Eagleville, PA, 2003; Herdewijn, P. (Ed.), Oligonucleotide Synthesis: Methods and Applications, Methods in Molecular Biology, Humana Press, 2004; A. Nagy, M. Gertsenstein, K. Vintersten, R. Behringer, Manipulating the Mouse Embryo: A Laboratory Manual, 3rd edition, Cold Spring Harbor Laboratory Press; December 15, 2002, ISBN-10: 0879695919; Kursad Turksen (Ed.), Embryonic

stem cells: methods and protocols in Methods Mol Biol. 2002;185, Humana Press; Current Protocols in Stem Cell Biology, ISBN: 9780470151808.

[0069] The singular terms "a," "an," and "the" are not intended to be limiting and include plural referents unless explicitly stated otherwise or the context clearly indicates otherwise.

5 **[0070]** Methods, systems and compositions according to aspects of the present invention provide protein-induced cell reprogramming in vivo to treat a subject in need thereof.

[0071] Methods, systems and compositions according to aspects of the present invention provide for conversion of diseased or injured cells into normal cells by introducing one or more QQ-modified reprogramming proteins into the diseased or injured cells in vivo without
10 introduction of nucleic acids encoding the one or more reprogramming proteins.

[0072] Methods, systems and compositions according to aspects of the present invention provide for conversion of cancer cells into non-cancerous cells by introducing one or more QQ-modified reprogramming proteins into the cancer cells in vivo without introduction of nucleic acids encoding the one or more reprogramming proteins.

15 **[0073]** Methods of treating a subject in need thereof are provided according to aspects of the present invention which include systemically and/or locally administering a pharmaceutical composition comprising a protein transduction reagent-modified reprogramming protein to the subject, wherein the protein transduction reagent is non-covalently bound to the reprogramming protein and wherein the protein transduction reagent comprises a cation reagent and a lipid.

20 **[0074]** Pharmaceutical compositions which include a protein transduction reagent-modified reprogramming protein are provided according to aspects of the present invention.

[0075] A "protein transduction reagent-modified reprogramming protein" is a reprogramming protein that has been treated with the protein transduction reagent, also termed a "QQ reagent" herein. The term "protein transduction reagent" refers to a composition effective to enable a
25 protein non-covalently bound to the protein transduction reagent to be delivered into mammalian cells and once present in mammalian cells, to dissociate from the protein to allow proper delivery of the protein to its proper subcellular location. The protein transduction reagent, also termed a "QQ reagent" herein, includes at least one cation reagent, at least one lipid, and optionally an enhancer. The term "QQ modified reprogramming protein" and grammatical variants thereof as
30 used herein is equivalent to "protein transduction reagent-modified reprogramming protein" and grammatical variants thereof as used herein. Similarly, one or more proteins termed "QQ" protein signifies that the proteins is modified by treatment with a protein transduction reagent and is a "protein transduction reagent-modified reprogramming protein. For example, the term "QQ-SON"

refers to a mixture of Sox2, Oct4 and Nanog proteins modified by treatment with a protein transduction reagent as described herein to produce protein transduction reagent-modified reprogramming Sox2, Oct4 and Nanog proteins.

[0076] One example of an appropriate cation reagent is polyethylenimine (PEI), such as, but not limited to, PEI Mw 1,200 (PEI 1.2K), PEI Mw 2000 (PEI 2K), PEI Mw 4000 (PEI 4K) and PEI Mw 8000 (PEI 8K). The lipid can be any lipid known to those of skill in the art to have the same general properties as those listed herein. Examples of such lipids include, but are not limited to, DOTMA (N-1-(2,3-dioleoyloxy)propyl-N,N,N-trimethyl-ammonium chloride; DOGS (dioctadecylamido-glycylspermine); DOTAP, 1,2-dioleoyl-3-trimethylammonium-propane; DOPE, 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine; POPC, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine; and DMPE 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine.

[0077] Optionally, the protein transduction reagent includes polyethylenimine as a cation reagent and the lipid is DOTAP or DOTMA and DOPE or DOGS; POPC and DMPE; or DOTAP or DOTMA, DOPE or DOGS, POPC and DMPE. Optionally, the protein transduction reagent is QQ1a, QQ2a, QQ3a, QQ4a, QQ5a, QQ6a, QQ7a, QQ8a, QQ9a as described in Table 1.

[0078] The optional enhancer can be any enhancer that significantly enhances cell loading of cationized proteins. Examples of such enhancers in cell cultures include, but are not limited to MG132, protease inhibitor, CaCl₂, DMSO and growth factors. Other enhancers can also be used, including, but not limited to, cell membrane surfactants. The reagent can also include stabilizers and other inert carriers that do not affect the function of the reagent. As shown in Table 1, the concentrations and specific compounds utilized can vary.

[0079] The term “reprogramming protein” as used herein refers to a DNA binding transcription factor protein, or effective portion thereof, which affects transcription of a gene and which induces a change from a first differentiated cell type to a second, different, differentiated cell type. The change from a first differentiated cell type to a second, different, differentiated cell type typically proceeds through an intermediate, less differentiated, cell type, such as a transient stem cell including protein-induced pluripotent stem cells. Reprogramming proteins and nucleic acids that encode them have been isolated from humans and other species. Reprogramming proteins include, but are not limited to, Asc1, Ascl1, Brca-11, Brn2, C/EBP α , CEBP β , c-MYC, Dlx2, EKLF, Erg1, Er71, Fli1, Foxa1, Foxa2, Foxa3, FoxC1, FoxC2, FOXG1, FOXP3, Gata1, Gata2, Gata3, Gata4, Gata6, GF11, Hand2, Hb9, Hey1, Hey2, HNF4A, Hnf1a, Klf4, Lhx3, LIN28A, Lmx1a, Lsl1, MafA, MEF2c, Myt11, MYF5, MyoD, NAB2, Nanog, NeuroD1, NEUROG2, NEUROG3, Nurr1, Oct4, Pdx1, Pax4, PAX5, Pax6, Prdm16, PU.1, ROR gamma, Runx2,

SLC7A10, Slug, Sox2, Sox5, Sox6, Sox7, Sox9, Sox18, Stat5a, T-bet and Tbox5. Amino acid sequences for such reprogramming proteins are known, as exemplified by the sequences shown herein as SEQ ID NOs:1-63, along with nucleic acids encoding them shown herein as SEQ ID NOs: 65-127.

5 **[0080]** A reprogramming protein to be QQ-modified is obtained by methods such as isolation, synthesis, or recombinant expression of a nucleic acid encoding the reprogramming protein. Such proteins may also be obtained commercially.

10 **[0081]** The term “nucleic acid” refers to RNA or DNA molecules having more than one nucleotide in any form including single-stranded, double-stranded, oligonucleotide or polynucleotide. The term “nucleotide sequence” refers to the ordering of nucleotides in an oligonucleotide or polynucleotide in a single-stranded form of nucleic acid.

15 **[0082]** Recombinant expression of a reprogramming protein to be QQ-modified includes expression of a nucleic acid encoding the protein wherein the nucleic acid is included in an expression construct.

20 **[0083]** A host cell may be transfected with the expression construct encoding the desired reprogramming protein such that the reprogramming protein is expressed in the cell.

25 **[0084]** The terms “expression construct” and “expression cassette” are used herein to refer to a double-stranded recombinant DNA molecule containing a desired nucleic acid coding sequence for a reprogramming factor to be expressed and containing one or more regulatory elements necessary or desirable for the expression of the operably linked coding sequence.

30 **[0085]** Expression constructs operable to express a desired protein include, for example, in operable linkage: a promoter, a DNA sequence encoding a desired protein and a transcription termination site.

35 **[0086]** The term “regulatory element” as used herein refers to a nucleotide sequence which controls some aspect of the expression of nucleic acid sequences. Exemplary regulatory elements illustratively include an enhancer, an internal ribosome entry site (IRES), an intron; an origin of replication, a polyadenylation signal (polyA), a promoter, a transcription termination sequence, and an upstream regulatory domain, which contribute to the replication, transcription, post-transcriptional processing of a nucleic acid sequence. Those of ordinary skill in the art are capable of selecting and using these and other regulatory elements in an expression construct with no more than routine experimentation. Expression constructs can be generated recombinantly or synthetically using well-known methodology.

[0087] The term “operably linked” as used herein refers to a nucleic acid in functional relationship with a second nucleic acid.

[0088] A regulatory element included in an expression cassette is a promoter in particular aspects.

5 [0089] The term “promoter” is well-known in the art and refers to one or more DNA sequences operably linked to a nucleic acid sequence to be transcribed and which bind an RNA polymerase and allow for initiation of transcription. A promoter is typically positioned upstream (5') of a nucleic acid encoding a peptide or protein to be expressed.

10 [0090] An included promoter can be a constitutive promoter or can provide inducible expression. One of skill in the art is familiar with various well-known promoters and is able to select a promoter suitable for use in expressing a peptide or protein in a particular environment, such as in a specified cell type.

[0091] For expression in a yeast host cell, suitable promoters include, but are not limited to, an ADH1 promoter, a PGK1 promoter, an ENO promoter, a PYK1 promoter and the like; or a regulatable promoter such as a GAL1 promoter, a GAL 10 promoter, an ADH2 promoter, a PH05 promoter, a CUP1 promoter, a GAL7 promoter, a MET25 promoter, a MET3 promoter, a CYC1 promoter, a HIS 3 promoter, an ADH1 promoter, a PGK promoter, a GAPDH promoter, an ADC1 promoter, a TRP1 promoter, a URA3 promoter, a LEU2 promoter, an ENO promoter, a TP1 promoter, and AOX1.

20 [0092] For expression in a prokaryotic host cell include, suitable promoters include, but are not limited to, a bacteriophage T7 RNA polymerase promoter; a trp promoter; a lac operon promoter; a trc promoter; a tac promoter; an araBAD promoter; an ssaG promoter; a pagC promoter, a sigma70 promoter, a dps promoter, an spv promoter, an SPI-2 promoter; an actA promoter, an rps M promoter; a tetracycline promoter, an SP6 promoter, a bacteriophage T3 promoter, a gpt promoter and a bacteriophage lambda P promoter.

25 [0093] Additional suitable bacterial and eukaryotic promoters are well-known, for example as described in Sambrook et al., Molecular Cloning, A Laboratory Manual, 2nd ed. 1989; and 3rd ed., 2001; Kriegler, Gene Transfer and Expression: A Laboratory Manual (1990); and Ausubel et al., Current Protocols in Molecular Biology, 2014.

30 [0094] For expression in an eukaryotic cell, promoters that can be included in an expression construct include, but are not limited to, cytomegalovirus immediate early promoter; herpes simplex virus thymidine kinase promoter; early and late SV40 promoters; a phosphoglycerate kinase (PGK) promoter; a promoter present in long terminal repeats from a retrovirus; and a mouse

metallothionein-I promoter, a beta-actin promoter, a ROSA26 promoter, a heat shock protein 70 (Hsp70) promoter, an EF-1 alpha gene encoding elongation factor 1 alpha (EF1) promoter, an eukaryotic initiation factor 4A (eIF-4A1) promoter, a chloramphenicol acetyltransferase (CAT) promoter and the long terminal repeat region of Rous Sarcoma virus (RSV promoter).

5 **[0095]** In addition to a promoter, one or more enhancer sequences may be included such as, but not limited to, cytomegalovirus (CMV) early enhancer element and an SV40 enhancer element.

[0096] Additional included sequences include an intron sequence such as the beta globin intron or a generic intron, a transcription termination sequence, and an mRNA polyadenylation (pA) sequence such as, but not limited to SV40-pA, beta-globin-pA and SCF-pA.

10 **[0097]** An expression construct may include sequences necessary for amplification in bacterial cells, such as a selection marker (e.g. kanamycin or ampicillin resistance gene) and a replicon.

[0098] An internal ribosome entry site (IRES) is an optionally included nucleic acid sequence that permits translation initiation at an internal site in an mRNA. IRES are well-known in the art, 15 for example as described in Pelletier, J. et al., Nature, 334:320-325, 1988; Vagner, S. et al., EMBO Rep., 2:893-898, 2001; and Hellen, C. U. et al, Genes Dev. 15:1593-1612, 2001.

[0099] The term “transcription termination site” refers to a DNA sequence operable to terminate transcription by an RNA polymerase. A transcription termination site is generally positioned downstream (3') of a nucleic acid encoding a peptide or protein to be expressed.

20 **[00100]** A leader sequence is optionally included in an expression construct.

[00101] Codon optimization of a nucleic acid encoding a desired protein may be used to improve expression in a particular expression system, for example by improving the efficiency of translation. A selected nucleic acid encoding a desired protein may be codon optimized for expression in any designated host cell, prokaryotic or eukaryotic, such as, but not limited to, 25 bacteria, insect cells, yeast, fungus, bird eggs and mammalian cells.

[00102] An expressed protein optionally includes an N-terminal element such as a leader sequence and/or N-terminal methionine.

[00103] In addition to one or more nucleic acids encoding a desired reprogramming protein, one or more nucleic acid sequences encoding additional proteins can be included in an expression vector. For example, a nucleic acid sequence encoding a reporter, including, but not limited to, 30 beta-galactosidase, green fluorescent protein and antibiotic resistance reporters is optionally included. In a further example, a his-tag, GST-tag or MBP-tag is optionally included.

[00104] A nucleic acid encoding a reprogramming protein can be cloned into an expression vector for transformation into prokaryotic or eukaryotic cells and expression of the encoded peptides and/or protein(s). As used herein, "expression vectors" are defined as polynucleotides which, when introduced into an appropriate host cell, an expression system, can be transcribed and translated, producing the encoded polypeptide(s).

[00105] Expression vectors are known in the art and include plasmids, cosmids, viruses and bacteriophages, for example. Expression vectors can be prokaryotic vectors, insect vectors, or eukaryotic vectors, for example.

[00106] For example, an expression construct including, in operable linkage: a promoter, a DNA sequence encoding a desired protein and a transcription termination site, is included in a plasmid, cosmid, virus or bacteriophage expression vector.

[00107] Particular vectors are known in the art and one of skill in the art will recognize an appropriate vector for a specific purpose.

[00108] Any suitable expression vector/host cell system can be used for expression of a transcription factor for administration to a subject according to aspects of the present invention.

[00109] Expression of a reprogramming protein using a recombinant expression vector is accomplished by introduction of the expression vector into an eukaryotic or prokaryotic host cell expression system such as an insect cell, mammalian cell, yeast cell, fungus, bird egg, bacterial cell or any other single or multicellular organism recognized in the art.

[00110] Host cells containing the recombinant expression vector are maintained under conditions wherein the desired protein is produced. Host cells may be cultured and maintained using known cell culture techniques such as described in Celis, Julio, ed., 1994, Cell Biology Laboratory Handbook, Academic Press, N.Y. Various culturing conditions for these cells, including media formulations with regard to specific nutrients, oxygen, tension, carbon dioxide and reduced serum levels, can be selected and optimized by one of skill in the art.

[00111] Bacterial cells can be used as the host cells to produce reprogramming proteins. Recombinant protein expression in bacterial cells and purification of the produced protein may be performed using known protocols, such as described in Paulina Balbás, Argelia Lorence ed., 2004, Recombinant Gene Expression: Reviews and Protocols, Humana Press, New Jersey; Peter E. Vaillancourt, 2003, E. Coli Gene Expression Protocols, Springer Science & Business Media.

[00112] Optionally, recombinantly produced reprogramming proteins are purified to remove endotoxin when an endotoxin producing host cell type is used. For example, an additional washing

step can be added during protein purification stage using 10 column volume of 0.2% of Triton X114 to remove endotoxin from bacterially expressed recombinant reprogramming proteins.

[00113] Alternatively, in order to produce recombinant reprogramming proteins which do not trigger endotoxic response in human cells, a genetically modified bacterial strain, ClearColi™ BL21(DE3) can be used as host cells such that no endotoxin removal is required.

[00114] For expression in a host cell, any of the well-known procedures for introducing recombinant nucleic acids into host cells may be used, such as calcium phosphate transfection, polybrene, protoplast fusion, electroporation, sonoporation, liposomes and microinjection, examples of which are described in Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 2001; and Ausubel, F. et al., (Eds.), Current Protocols in Molecular Biology, 2014.

[00115] Cell free expression systems is optionally used to express a reprogramming protein, such as described in Ausubel, F. et al., (Eds.), Current Protocols in Molecular Biology, 2014.

[0001] Human reprogramming proteins shown herein as SEQ ID NOs: 1-63, and encoded by the nucleic acid sequences of SEQ ID NOs:65-127, and variants thereof, can be used in methods according to aspects described herein.

[00116] As used herein, the term "variant" refers to a variation of a nucleic acid sequence encoding a reprogramming protein or a variation of a reprogramming protein in which one or more nucleotides or amino acid residues have been modified by nucleotide or amino acid substitution, addition, or deletion while retaining the function of the reference nucleic acid sequence or reprogramming protein. Variants of a nucleic acid sequence or reprogramming protein described herein are characterized by conserved functional properties compared to the corresponding nucleic acid sequence or reprogramming protein.

[00117] Mutations can be introduced using standard molecular biology techniques, such as chemical synthesis, site-directed mutagenesis and PCR-mediated mutagenesis.

[00118] One of skill in the art will recognize that one or more amino acid mutations can be introduced without altering the functional properties of a desired reprogramming protein. For example, one or more amino acid substitutions, additions, or deletions can be made without altering the functional properties of a desired reprogramming protein.

[00119] Biological activity of a reprogramming protein variant is readily determined by one of skill in the art, for instance using any of the functional assays described herein or other functional assays known in the art.

[00120] Variants of a reprogramming protein described herein are characterized by conserved

functional properties compared to the corresponding reprogramming protein and have 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or greater identity to the amino acid sequence of a reference reprogramming protein.

[00121] When comparing a reference reprogramming protein to a variant, amino acid similarity may be considered in addition to identity of amino acids at corresponding positions in an amino acid sequence. "Amino acid similarity" refers to amino acid identity and conservative amino acid substitutions in a putative homologue compared to the corresponding amino acid positions in a reference protein.

[00122] Variants of a reprogramming protein described herein are characterized by conserved functional properties compared to the corresponding reprogramming protein and have 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or greater similarity to the amino acid sequence of a reference reprogramming protein.

[00123] Conservative amino acid substitutions can be made in reference proteins to produce variants.

[00124] Conservative amino acid substitutions are art recognized substitutions of one amino acid for another amino acid having similar characteristics. For example, each amino acid may be described as having one or more of the following characteristics: electropositive, electronegative, aliphatic, aromatic, polar, hydrophobic and hydrophilic. A conservative substitution is a substitution of one amino acid having a specified structural or functional characteristic for another amino acid having the same characteristic. Acidic amino acids include aspartate, glutamate; basic amino acids include histidine, lysine, arginine; aliphatic amino acids include isoleucine, leucine and valine; aromatic amino acids include phenylalanine, glycine, tyrosine and tryptophan; polar amino acids include aspartate, glutamate, histidine, lysine, asparagine, glutamine, arginine, serine, threonine and tyrosine; and hydrophobic amino acids include alanine, cysteine, phenylalanine, glycine, isoleucine, leucine, methionine, proline, valine and tryptophan; and conservative substitutions include substitution among amino acids within each group. Amino acids may also be described in terms of relative size, alanine, cysteine, aspartate, glycine, asparagine, proline, threonine, serine, valine, all typically considered to be small.

[00125] A variant can include synthetic amino acid analogs, amino acid derivatives and/or non-standard amino acids, illustratively including, without limitation, alpha-aminobutyric acid, citrulline, canavanine, cyanoalanine, diaminobutyric acid, diaminopimelic acid, dihydroxy-phenylalanine, djenkolic acid, homoarginine, hydroxyproline, norleucine, norvaline, 3-

phosphoserine, homoserine, 5-hydroxytryptophan, 1-methylhistidine, 3-methylhistidine, and ornithine.

[00126] Percent identity is determined by comparison of amino acid or nucleic acid sequences, including a reference amino acid or nucleic acid sequence and a putative homologue amino acid or nucleic acid sequence. To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino acid or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity=number of identical overlapping positions/total number of positions X 100%). The two sequences compared are generally the same length or nearly the same length.

[00127] The determination of percent identity between two sequences can also be accomplished using a mathematical algorithm. Algorithms used for determination of percent identity illustratively include the algorithms of S. Karlin and S. Altshul, PNAS, 90:5873-5877, 1993; T. Smith and M. Waterman, Adv. Appl. Math. 2:482-489, 1981, S. Needleman and C. Wunsch, J. Mol. Biol., 48:443-453, 1970, W. Pearson and D. Lipman, PNAS, 85:2444-2448, 1988 and others incorporated into computerized implementations such as, but not limited to, GAP, BESTFIT, FASTA, TFASTA; and BLAST, for example incorporated in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Drive, Madison, Wis.) and publicly available from the National Center for Biotechnology Information.

[00128] A non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul, 1990, PNAS 87:2264-2268, modified as in Karlin and Altschul, 1993, PNAS. 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul et al., 1990, J. Mol. Biol. 215:403. BLAST nucleotide searches are performed with the NBLAST nucleotide program parameters set, e.g., for score=100, word length=12 to obtain nucleotide sequences homologous to a nucleic acid molecules of the present invention. BLAST protein searches are performed with the XBLAST program parameters set, e.g., to score 50, word length=3 to obtain amino acid sequences homologous to a protein molecule of the present invention. To obtain gapped alignments for comparison purposes, Gapped BLAST are utilized as described in Altschul et al., 1997, Nucleic

Acids Res. 25:3389-3402. Alternatively, PSI BLAST is used to perform an iterated search which detects distant relationships between molecules. When utilizing BLAST, Gapped BLAST, and PSI Blast programs, the default parameters of the respective programs (e.g., of XBLAST and NBLAST) are used. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, 1988, CABIOS 4:11-17. Such an algorithm is incorporated in the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 is used.

[00129] The percent identity between two sequences is determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, typically only exact matches are counted.

[00130] One of skill in the art will recognize that one or more nucleic acid or amino acid mutations can be introduced without altering the functional properties of a given nucleic acid or protein, respectively.

[00131] As noted, human reprogramming proteins shown herein as SEQ ID NOs: 1-63, are encoded by the nucleic acid sequences of SEQ ID NOs:65-127. It is appreciated that due to the degenerate nature of the genetic code, alternate nucleic acid sequences encode a particular reprogramming protein, and that such alternate nucleic acids may be expressed to produce the desired reprogramming protein.

[00132] The term “complementary” refers to Watson-Crick base pairing between nucleotides and specifically refers to nucleotides hydrogen bonded to one another with thymine or uracil residues linked to adenine residues by two hydrogen bonds and cytosine and guanine residues linked by three hydrogen bonds. In general, a nucleic acid includes a nucleotide sequence described as having a “percent complementarity” to a specified second nucleotide sequence. For example, a nucleotide sequence may have 80%, 90%, or 100% complementarity to a specified second nucleotide sequence, indicating that 8 of 10, 9 of 10 or 10 of 10 nucleotides of a sequence are complementary to the specified second nucleotide sequence. For instance, the nucleotide sequence 3'-TCGA-5' is 100% complementary to the nucleotide sequence 5'-AGCT-3'. Further, the nucleotide sequence 3'-TCGA-5' is 100% complementary to a region of the nucleotide sequence 5'-TTAGCTGG-3'.

[00133] The terms “hybridization” and “hybridizes” refer to pairing and binding of complementary nucleic acids. Hybridization occurs to varying extents between two nucleic acids

depending on factors such as the degree of complementarity of the nucleic acids, the melting temperature, T_m , of the nucleic acids and the stringency of hybridization conditions, as is well known in the art. The term “stringency of hybridization conditions” refers to conditions of temperature, ionic strength, and composition of a hybridization medium with respect to particular common additives such as formamide and Denhardt’s solution. Determination of particular hybridization conditions relating to a specified nucleic acid is routine and is well known in the art, for instance, as described in J. Sambrook and D.W. Russell, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press; 3rd Ed., 2001; and F.M. Ausubel, Ed., *Short Protocols in Molecular Biology*, Current Protocols; 5th Ed., 2002. High stringency hybridization conditions are those which only allow hybridization of substantially complementary nucleic acids. Typically, nucleic acids having about 85-100% complementarity are considered highly complementary and hybridize under high stringency conditions. Intermediate stringency conditions are exemplified by conditions under which nucleic acids having intermediate complementarity, about 50-84% complementarity, as well as those having a high degree of complementarity, hybridize. In contrast, low stringency hybridization conditions are those in which nucleic acids having a low degree of complementarity hybridize.

[00134] The terms “specific hybridization” and “specifically hybridizes” refer to hybridization of a particular nucleic acid to a target nucleic acid without substantial hybridization to nucleic acids other than the target nucleic acid in a sample.

[00135] Stringency of hybridization and washing conditions depends on several factors, including the T_m of the probe and target and ionic strength of the hybridization and wash conditions, as is well-known to the skilled artisan. Hybridization and conditions to achieve a desired hybridization stringency are described, for example, in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, 2001; and Ausubel, F. et al., (Eds.), *Short Protocols in Molecular Biology*, Wiley, 2002.

[00136] An example of high stringency hybridization conditions is hybridization of nucleic acids over about 100 nucleotides in length in a solution containing 6X SSC, 5X Denhardt’s solution, 30% formamide, and 100 micrograms/ml denatured salmon sperm at 37°C overnight followed by washing in a solution of 0.1X SSC and 0.1% SDS at 60°C for 15 minutes. SSC is 0.15M NaCl/0.015M Na citrate. Denhardt’s solution is 0.02% bovine serum albumin/0.02% FICOLL/0.02% polyvinylpyrrolidone.

[00137] A reprogramming protein modified by QQ is an isolated protein according to aspects of the present invention. The term “isolated protein” indicates that the protein has been separated

from biological materials, such as cells, cellular debris and other proteins, which may be present in the system in which the protein is produced. The term “isolated protein” may, but does not necessarily, indicate that the protein is purified. Purified protein included in methods and compositions of the present invention contains least about 1 – 100% of the mass, by weight, such as about 25%, 50%, 75%, 85%, 95%, 99% or greater than about 99% of the mass, by weight, of the protein included.

[00138] The term “subject” refers to an individual in need of treatment for a disease or injury responsive to the beneficial effects of cell reprogramming, and generally includes mammals and birds, such as, but not limited to, humans, other primates, cats, dogs, sheep, cows, goats, horses, pigs, poultry, rabbits and rodents, such as rats, mice and guinea pigs. According to aspects of the present invention, the subject is human.

[00139] Condition characterized by damaged, and/or defective cells.

[00140] The terms “treating” and “treatment” used to refer to treatment of a condition characterized by damaged, and/or defective cells such as a disease or injury of a subject include: inhibiting or ameliorating the disease or injury in the subject, such as slowing progression of the disease and/or reducing or ameliorating a sign or symptom of the disease or injury.

[00141] Conditions characterized by angiogenesis are treated according to aspects of methods of the present invention and are characterized by effective delivery of QQ-modified reprogramming proteins via the enhanced permeability and retention effect (EPR effect) of the angiogenic blood vessels.

[00142] Conditions characterized by damaged, and/or defective cells treated according to aspects of the present invention are various human conditions, including, but not limited to, cancer; cardiovascular disease or injury such as acute and chronic myocardial infarction, ischemia, heart injury, coronary artery disease, congenital heart disease, cardiomyopathies such as alcoholic cardiomyopathy, arrhythmogenic right ventricular cardiomyopathy, hypertensive cardiomyopathy, dilated cardiomyopathy, hypertrophic cardiomyopathy, inflammatory cardiomyopathy, ischemic cardiomyopathy, cardiomyopathy secondary to a systemic metabolic disease, myocardiodystrophy, noncompaction cardiomyopathy, restrictive cardiomyopathy, and valvular cardiomyopathy; vascular disease or blood vessel damage; anemias; ischemic and hemorrhagic stroke; metabolic diseases or conditions, such as diabetes, type I and type II, and obesity; neurological diseases and injuries such as spinal cord injury, traumatic brain injury, Huntington's disease, schizophrenia, Alzheimer's disease, amyotrophic lateral sclerosis, ataxias, autism, Lyme disease, meningitis, migraine, motor neuron diseases, movement disorders such as Parkinson's disease, neuropathy,

pain, neuropathic pain, spinal cord disorders, peripheral and central nerve disorders, autonomic nervous system disorders, seizure disorders such as epilepsy, sleep disorders, dementias such as Alzheimer's disease, muscular dystrophy, Charcot-Marie-Tooth Neuropathy Type 1A and demyelinating diseases such as acute disseminated encephalomyelitis, adrenoleukodystrophy, adrenomyeloneuropathy, multiple sclerosis, neuromyelitis optica, optic neuritis, and transverse myelitis; wounds; inflammatory conditions; liver diseases and injuries such as liver disease such as familial hyper-cholesterolaemia (FH), Crigler-Najjar syndrome, hereditary tryosinemia I, fulminant hepatic failure, viral hepatitis, drug-induced liver injury, cirrhosis, inherited hepatic insufficiency such as due to Wilson's disease, Gilbert's syndrome, or α 1-antitrypsin deficiency, hepatobiliary carcinoma, autoimmune liver disease, such as autoimmune chronic hepatitis or primary biliary cirrhosis; autoimmune diseases; osteoarthritis; rheumatoid arthritis; cartilage disease or injury such as due to joint disorder, osteoarthritis, cartilage injury, traumatic rupture or detachment, achondroplasia, costochondritis, spinal disc herniation, relapsing polychondritis, tumor, chondroma, chondrosarcoma, and pleomorphic adenoma; Crohn's disease and genetic abnormalities.

[00143] Particular cancers treated using methods and compositions described herein are characterized by abnormal cell proliferation including, but not limited to, pre-neoplastic hyperproliferation, cancer in-situ, neoplasms and metastasis. The terms "treating" and "treatment" used to refer to treatment of a cancer in a subject include: inhibiting or ameliorating the cancer in the subject, such as slowing progression of the cancer and/or reducing or ameliorating a sign or symptom of the cancer.

[00144] A therapeutically effective amount of a composition including an anti-cancer composition of the present invention is an amount which has a beneficial effect in a subject being treated. In subjects having cancer, such as a condition characterized by abnormal cell proliferation including, but not limited to, pre-neoplastic hyperproliferation, cancer in-situ, neoplasms, metastasis, a tumor, a benign growth or other condition responsive to an inventive composition, a therapeutically effective amount of a QQ-modified protein is effective to ameliorate one or more signs and/or symptoms of the cancer. For example, a therapeutically effective amount of a composition is effective to detectably decrease proliferation of cells of a cancer characterized by abnormal cell proliferation including, but not limited to, pre-neoplastic hyperproliferation, cancer in-situ, neoplasms, metastasis, a tumor, a benign growth or other condition responsive to an administered QQ-modified protein.

[00145] Such cancers include solid and non-solid cancers such as cancer of the bladder; breast; colorectal; cervical; esophagus; head and neck; kidney; lung; cancers of the nervous system such as

glioblastoma, astrocytoma, ependymoma, neuroblastoma, retinoblastoma, meningiomas, granular cell tumors and nerve sheath tumors; ovary; pancreas; prostate; skin; stomach; testicle; throat cancer; urachus; or vagina.

[00146] A pharmaceutical composition according to aspects of the present invention includes a QQ-modified reprogramming protein and a pharmaceutically acceptable carrier.

[00147] The term “pharmaceutically acceptable carrier” as used herein refers to a carrier or diluent that is generally non-toxic to an intended recipient and which does not significantly inhibit activity of a QQ-modified protein or other active agent included in the composition.

[00148] A composition according to the present invention generally includes about 0.1-99% of a QQ-modified reprogramming protein.

[00149] Pharmaceutical compositions suitable for delivery to a subject may be prepared in various forms illustratively including physiologically acceptable sterile aqueous or nonaqueous solutions, dispersions, suspensions or emulsions, and sterile powders for reconstitution into sterile injectable solutions or dispersions.

[00150] Pharmaceutical compositions optionally include a buffer, a solvent, or a diluent.

[00151] Examples of suitable aqueous and nonaqueous carriers include water, ethanol, polyols such as propylene glycol, polyethylene glycol and glycerol; vegetable oils such as olive oil; and injectable organic esters such as ethyloleate; and suitable mixtures of any two or more thereof.

[00152] Such formulations are administered by a suitable route including parenteral administration. Optionally, administration includes systemic or local administration.

[00153] These compositions may also contain adjuvants such as preserving, wetting, emulsifying, and dispensing agents. Prevention of the action of microorganisms can be ensured by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, and the like. One or more isotonic agents is optionally included, for example, sugars and or salts such as sodium chloride.

[00154] In particular aspects, QQ-modified proteins are administered by topical application.

[00155] A topical formulation can be an ointment, lotion, cream or gel in particular aspects. Topical dosage forms such as ointment, lotion, cream or gel bases are described in Remington: The Science and Practice of Pharmacy, 21st Ed., Lippincott Williams & Wilkins, 2006, p.880-882 and p.886-888; and in Allen, L. V. et al., Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems, 8th Ed., Lippincott Williams & Wilkins, 2005, p.277-297.

[00156] Pharmaceutically acceptable carriers and formulation of pharmaceutical compositions are known in the art, illustratively including, but not limited to, as described in Remington: The

Science and Practice of Pharmacy, 21st Ed., Lippincott, Williams & Wilkins, Philadelphia, PA, 2006; and Allen, L.V. et al., Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems, 8th Ed., Lippincott, Williams & Wilkins, Philadelphia, PA, 2005.

[00157] A pharmaceutical composition including a QQ-modified protein is suitable for administration to a subject by a variety of systemic and/or local routes including, but not limited to, parenteral, oral, rectal, nasal, pulmonary, epidural, ocular, otic, intraarterial, intracardiac, intracerebroventricular, intracranial, intradermal, intravenous, intramuscular, intraperitoneal, intraosseous, intrathecal, intratumoral, intravesical, subcutaneous, topical, transdermal, and transmucosal, such as by sublingual, buccal, vaginal, and inhalational routes of administration.

[00158] Administration of pharmaceutical composition

[00159] An inventive composition may be administered acutely or chronically. For example, a composition as described herein may be administered as a unitary dose or in multiple doses over a relatively limited period of time, such as seconds – hours. In a further embodiment, administration may include multiple doses administered over a period of days – years, such as for chronic treatment of cancer.

[00160] A therapeutically effective amount of a QQ-modified protein will vary depending on the route of administration and form of the composition being administered and the particular composition administered, the severity and type of condition being treated in the subject, the species of the subject, the age and sex of the subject and the general physical characteristics of the subject to be treated. One of skill in the art could determine a therapeutically effective amount in view of these and other considerations typical in medical practice without undue experimentation in view of the present disclosure and what is known in the art. In general it is contemplated that a therapeutically effective amount would be in the range of about 0.001 ng/kg – 100 mg/kg body weight, optionally in the range of about 0.01 ng/kg – 1 mg/kg, and further optionally in the range of about 0.1 ng/kg – 0.1 mg/kg. Further, dosage may be adjusted depending on whether treatment is to be acute or continuing.

[00161] Usually between 1 and 100 doses of a QQ-modified protein are administered to treat a subject in need thereof, although more doses can be given. A QQ-modified protein can be administered twice a day, daily, biweekly, weekly, every other week, monthly or at some other interval, for a treatment course extending one day, 1 week, 2 weeks, 4 weeks, 1 month, 2 months, 3-6 months or longer. A course of treatment is optionally repeated and may extend to chronic treatment if necessary.

[00162] Administration of a QQ-modified protein according to aspects of a method of the present invention includes administration according to a dosage regimen to produce a desired response. A suitable schedule for administration of doses depends on several factors including age, weight, gender, medical history and health status of the subject, type of composition used and route of administration, for example. One of skill in the art is able to readily determine a dose and schedule of administration for a particular subject.

[00163] Methods according to embodiments of the present invention include administration of a QQ-modified protein as a pharmaceutical formulation, such as by systemic or local administration. Exemplary routes of administration include, but are not limited to, parenteral, oral, rectal, nasal, pulmonary, epidural, ocular, otic, intraarterial, intracardiac, intracerebroventricular, intracranial, intradermal, intravenous, intramuscular, intraperitoneal, intraosseous, intrathecal, intratumoral, intravesical, subcutaneous, topical, transdermal, and transmucosal, such as by sublingual, buccal, vaginal, and inhalational routes of administration.

[00164] The QQ-modified protein may be administered parenterally, for example, by injection such as intravenous injection, intramuscular injection, intraperitoneal injection, subcutaneous injection, transdermal injection, intrathecal injection, intracranial injection, intracerebrospinal injection, and/or continuous infusion such as by an intravenous or intracerebrospinal continuous infusion device.

[00165] According to aspects, a protein transduction reagent-modified reprogramming protein is a DNA binding transcription factor. Administration of a protein transduction reagent-modified reprogramming protein is effective to reprogram one or more cell types in situ at the site of disease or damage to treat the disease or damage in vivo according to aspects described herein. Administration of a protein transduction reagent-modified reprogramming protein is effective to reprogram one or more cell types in situ at the site of disease or damage, generating transient stem cells which then differentiate into normal cells in situ at the site of disease or damage in vivo to treat the disease or damage according to aspects described herein.

[00166] Methods of treating cancer according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog. Methods of treating cancer according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog. Methods of treating cancer according to aspects of the present invention include administering: protein transduction reagent-modified Sox2, protein

transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog. Such methods produce new normal cells at the site of the cancer by reprogramming cancer cells at the site.

5 [00167] Methods of treating brain tumor according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog. Methods of treating brain tumor according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog. Methods of treating brain tumor according to aspects of the present invention include administering: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog. Such methods produce new normal cells at the site of the brain tumor by reprogramming brain tumor cells at the site.

10 [00168] Methods of treating brain tumor according to aspects of the present invention including administering one or more of: protein transduction reagent-modified Sox2 and protein transduction reagent-modified NeuroD. Methods of treating pancreatic cancer according to aspects of the present invention including administering: protein transduction reagent-modified Sox2 and protein transduction reagent-modified NeuroD. Such methods produce new normal cells at the site of pancreatic cancer by reprogramming brain tumor cells at the site.

15 [00169] Methods of treating breast cancer according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog. Methods of treating breast cancer according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog. Methods of treating breast cancer according to aspects of the present invention include administering: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog. Such methods produce new normal cells at the site of the breast cancer by reprogramming breast cancer cells at the site.

20 [00170] Methods of treating breast cancer according to aspects of the present invention including administering one or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Slug, protein transduction reagent-modified Gata3, protein transduction reagent-modified Brca-1 and protein transduction reagent-modified State5a. Methods

of treating breast cancer according to aspects of the present invention including administering two or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Slug, protein transduction reagent-modified Gata3, protein transduction reagent-modified Brca-1 and protein transduction reagent-modified State5a. Methods of treating breast cancer according to aspects of the present invention including administering three or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Slug, protein transduction reagent-modified Gata3, protein transduction reagent-modified Brca-1 and protein transduction reagent-modified State5a. Methods of treating breast cancer according to aspects of the present invention including administering four or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Slug, protein transduction reagent-modified Gata3, protein transduction reagent-modified Brca-1 and protein transduction reagent-modified State5a. Methods of treating breast cancer according to aspects of the present invention including administering: protein transduction reagent-modified Sox9, protein transduction reagent-modified Slug, protein transduction reagent-modified Gata3, protein transduction reagent-modified Brca-1 and protein transduction reagent-modified State5a. Such methods produce new normal cells at the site of breast cancer by reprogramming breast cancer cells at the site.

[00171] Methods of treating pancreatic cancer according to aspects of the present invention including administering one or more of: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog. Methods of treating pancreatic cancer according to aspects of the present invention including administering two or more of: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog. Methods of treating pancreatic cancer according to aspects of the present invention including administering: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog. Such methods produce new normal cells at the site of pancreatic cancer by reprogramming pancreatic cancer cells at the site.

[00172] Methods of treating pancreatic cancer according to aspects of the present invention including administering one or more of: protein transduction reagent-modified PDX1, protein transduction reagent-modified PAX4, protein transduction reagent-modified MafA and protein transduction reagent-modified Ngn3. Methods of treating pancreatic cancer according to aspects of the present invention including administering two or more of: protein transduction reagent-modified PDX1, protein transduction reagent-modified PAX4, protein transduction reagent-modified MafA and protein transduction reagent-modified Ngn3. Methods of treating pancreatic

cancer according to aspects of the present invention including administering three or more of: protein transduction reagent-modified PDX1, protein transduction reagent-modified PAX4, protein transduction reagent-modified MafA and protein transduction reagent-modified Ngn3. Methods of treating pancreatic cancer according to aspects of the present invention including administering:
5 protein transduction reagent-modified PDX1, protein transduction reagent-modified PAX4, protein transduction reagent-modified MafA and protein transduction reagent-modified Ngn3. Such methods produce new normal cells at the site of pancreatic cancer by reprogramming pancreatic cancer cells at the site.

[00173] Methods of treating a heart disease or heart damage, such as damage due to acute or
10 chronic myocardial infarction, according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Gata4, protein transduction reagent-modified Hand2, protein transduction reagent-modified MEF2c and protein transduction reagent-modified Tbox5. Methods of treating a heart disease or heart damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified
15 Gata4, protein transduction reagent-modified Hand2, protein transduction reagent-modified MEF2c and protein transduction reagent-modified Tbox5. Methods of treating a heart disease or heart damage according to aspects of the present invention include administering three or more of: protein transduction reagent-modified Gata4, protein transduction reagent-modified Hand2, protein transduction reagent-modified MEF2c and protein transduction reagent-modified Tbox5. Methods
20 of treating a heart disease or heart damage according to aspects of the present invention include administering: protein transduction reagent-modified Gata4, protein transduction reagent-modified Hand2, protein transduction reagent-modified MEF2c and protein transduction reagent-modified Tbox5. Such methods produce new cardiomyocytes, smooth muscle cells and endothelial cells at the site of the disease or damage by reprogramming fibroblasts at the site.

[00174] Methods of treating a liver disease or liver damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Gata4, protein transduction reagent-modified Hnf1a and protein transduction reagent-modified Foxa3. Methods of treating a liver disease or liver damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Gata4, protein
30 transduction reagent-modified Hnf1a and protein transduction reagent-modified Foxa3. Methods of treating a liver disease or liver damage according to aspects of the present invention include administering: protein transduction reagent-modified Gata4, protein transduction reagent-modified Hnf1a and protein transduction reagent-modified Foxa3. Such methods produce new hepatocytes

at the site of the disease or damage by reprogramming fibroblasts at the site. Examples of such liver disease or liver damage include familial hyper-cholesterolaemia (FH), Crigler-Najjar syndrome and hereditary tryosinaemica I.

[00175] Methods of treating a liver disease or liver damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Hnfla, protein transduction reagent-modified Foxa1, protein transduction reagent-modified Foxa2 and protein transduction reagent-modified Foxa3. Methods of treating a liver disease or liver damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Hnfla, protein transduction reagent-modified Foxa1, protein transduction reagent-modified Foxa2 and protein transduction reagent-modified Foxa3. Methods of treating a liver disease or liver damage according to aspects of the present invention include administering: protein transduction reagent-modified Hnfla, protein transduction reagent-modified Foxa1, protein transduction reagent-modified Foxa2 and protein transduction reagent-modified Foxa3. Such methods produce new hepatocytes at the site of the disease or damage by reprogramming fibroblasts at the site. Examples of such liver disease or liver damage include familial hyper-cholesterolaemia (FH), Crigler-Najjar syndrome and hereditary tryosinaemica I.

[00176] Methods of treating atherosclerosis according to aspects of the present invention include administering one or more of: protein transduction reagent-modified CEBP/ α/β and protein transduction reagent-modified PU.1. Methods of treating atherosclerosis according to aspects of the present invention include administering both of: CEBP/ α/β and protein transduction reagent-modified PU.1. Such methods produce new foam cells and macrophages at the site of the atherosclerosis by reprogramming fibroblasts at the site.

[00177] Methods of treating a neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Sox2 and protein transduction reagent-modified Foxg1. Methods of treating a neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Sox2 and protein transduction reagent-modified Foxg1. Methods of treating neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified Brn2, protein transduction reagent-modified Sox2 and protein transduction reagent-modified Foxg1. Such methods produce new neurons at the site of the disease or damage by reprogramming fibroblasts at the site.

[00178] Methods of treating a neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Lmx1a, protein transduction reagent-modified Nurr1, protein transduction reagent-modified Brn2 and protein transduction reagent-modified Myt1l. Methods of treating neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Lmx1a, protein transduction reagent-modified Nurr1, protein transduction reagent-modified Brn2 and protein transduction reagent-modified Myt1l. Methods of treating neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering three or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Lmx1a, protein transduction reagent-modified Nurr1, protein transduction reagent-modified Brn2 and protein transduction reagent-modified Myt1l. Methods of treating neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering four or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Lmx1a, protein transduction reagent-modified Nurr1, protein transduction reagent-modified Brn2 and protein transduction reagent-modified Myt1l. Methods of treating neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified Asc1, protein transduction reagent-modified Lmx1a, protein transduction reagent-modified Nurr1, protein transduction reagent-modified Brn2 and protein transduction reagent-modified Myt1l. Such methods produce new glutamatergic neurons at the site of the disease or damage by reprogramming fibroblasts at the site.

[00179] Methods of treating a neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Myt1l and protein transduction reagent-modified NeuroD1. Methods of treating neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Myt1l and protein transduction reagent-modified NeuroD1. Methods of treating neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering three or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Myt1l and protein transduction reagent-modified NeuroD1.

Methods of treating neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified Ascl1, protein transduction reagent-modified Brn2 , protein transduction reagent-modified Mytl1 and protein transduction reagent-modified NeuroD1. Such methods produce new glutamatergic neurons at the site of the disease or damage by reprogramming fibroblasts at the site.

[00180] Methods of treating a neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified Ngn2. Such methods produce new glutamatergic neurons at the site of the disease or damage by reprogramming astrocytes at the site.

[00181] Methods of treating a neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Sox2 and protein transduction reagent-modified NeuroD1. Methods of treating neurodegenerative disease or neuronal tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified Sox2 and protein transduction reagent-modified NeuroD1. Such methods produce new neural stem cells at the site of the disease or damage.

[00182] Methods of treating a neurological disease or neuronal tissue damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Ascl1, protein transduction reagent-modified Mytl1, protein transduction reagent-modified Lhx3 and protein transduction reagent-modified Hb9. Methods of treating neurological disease or neuronal tissue damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Ascl1, protein transduction reagent-modified Mytl1, protein transduction reagent-modified Lhx3 and protein transduction reagent-modified Hb9.

Methods of treating neurological disease or neuronal tissue damage according to aspects of the present invention include administering three or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Ascl1, protein transduction reagent-modified Mytl1, protein transduction reagent-modified Lhx3 and protein transduction reagent-modified Hb9. Methods of treating neurological disease or neuronal tissue damage according to aspects of the present invention include administering four or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Ascl1, protein transduction reagent-modified Mytl1, protein transduction reagent-modified Lhx3 and protein transduction reagent-modified Hb9. Methods of treating neurological disease or neuronal tissue damage according to aspects of the

present invention include administering: protein transduction reagent-modified Brn2, protein transduction reagent-modified Ascl1, protein transduction reagent-modified Mytl1, protein transduction reagent-modified Lhx3 and protein transduction reagent-modified Hb9. Such methods produce new motor neurons at the site of the disease or damage by reprogramming fibroblasts at the site. Examples of such neurological diseases include amyotrophic lateral sclerosis (ALS).

[00183] Methods of treating a neurological disease or neuronal tissue damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Ascl1, protein transduction reagent-modified Mytl1, protein transduction reagent-modified Lhx3, protein transduction reagent-modified Hb9, protein transduction reagent-modified Lsl1 and protein transduction reagent-modified Ngn2.

Methods of treating neurological disease or neuronal tissue damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Ascl1, protein transduction reagent-modified Mytl1, protein transduction reagent-modified Lhx3, protein transduction reagent-modified Hb9, protein transduction reagent-modified Lsl1 and protein transduction reagent-modified Ngn2.

Methods of treating neurological disease or neuronal tissue damage according to aspects of the present invention include administering three or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Ascl1, protein transduction reagent-modified Mytl1, protein transduction reagent-modified Lhx3, protein transduction reagent-modified Hb9, protein transduction reagent-modified Lsl1 and protein transduction reagent-modified Ngn2.

Methods of treating neurological disease or neuronal tissue damage according to aspects of the present invention include administering four or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Ascl1, protein transduction reagent-modified Mytl1, protein transduction reagent-modified Lhx3, protein transduction reagent-modified Hb9, protein transduction reagent-modified Lsl1 and protein transduction reagent-modified Ngn2.

Methods of treating neurological disease or neuronal tissue damage according to aspects of the present invention include administering five or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Ascl1, protein transduction reagent-modified Mytl1, protein transduction reagent-modified Lhx3, protein transduction reagent-modified Hb9, protein transduction reagent-modified Lsl1 and protein transduction reagent-modified Ngn2.

Methods of treating neurological disease or neuronal tissue damage according to aspects of the present invention include administering six or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Ascl1, protein transduction reagent-modified Mytl1, protein

transduction reagent-modified Lhx3, protein transduction reagent-modified Hb9, protein transduction reagent-modified Lsl1 and protein transduction reagent-modified Ngn2. Methods of treating neurological disease or neuronal tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified Brn2, protein transduction reagent-modified Ascl1, protein transduction reagent-modified Myt1l, protein transduction reagent-modified Lhx3, protein transduction reagent-modified Hb9, protein transduction reagent-modified Lsl1 and protein transduction reagent-modified Ngn2. Such methods produce new motor neurons at the site of the disease or damage by reprogramming fibroblasts at the site. Examples of such neurological diseases include amyotrophic lateral sclerosis (ALS).

[00184] Methods of treating a neurological disease or neuronal tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified. Such methods produce new GABA neurons at the site of the disease or damage by reprogramming fibroblasts at the site. Examples of such neurological diseases include spinal muscular atrophy (SMA).

[00185] Methods of treating a neurological disease or neuronal tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified Dlx2. Such methods produce new GABA neurons at the site of the disease or damage by reprogramming fibroblasts at the site. Examples of such neurological diseases include amyotrophic lateral sclerosis (ALS), spinal muscular atrophy (SMA) and Parkinson's disease.

[00186] Methods of treating a neurological disease or neuronal tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified Dlx2 and protein transduction reagent-modified Ascl1. Such methods produce new GABA neurons at the site of the disease or damage by reprogramming astrocytes at the site. Examples of such neurological diseases include spinal muscular atrophy (SMA).

[00187] Methods of treating a neurological disease or neuronal tissue damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Ascl1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Myt1l, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Foxa2. Methods of treating neurological disease or neuronal tissue damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Ascl1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Myt1l, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Foxa2. Methods of treating neurological disease or neuronal tissue damage according to aspects of the

present invention include administering three or more of: protein transduction reagent-modified Ascl1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Mytl1, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Foxa2. Methods of treating neurological disease or neuronal tissue damage according to aspects of the present invention include administering four or more of: protein transduction reagent-modified Ascl1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Mytl1, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Foxa2. Methods of treating neurological disease or neuronal tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified Ascl1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Mytl1, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Foxa2. Such methods produce new dopamine neurons at the site of the disease or damage by reprogramming fibroblasts at the site. Examples of such neurological diseases include Parkinson's disease.

[00188] Methods of treating a neurological disease or neuronal tissue damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Ascl1, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Nurr1. Methods of treating a neurological disease or neuronal tissue damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Ascl1, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Nurr1. Methods of treating neurological disease or neuronal tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified Ascl1, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Nurr1. Such methods produce new neurons at the site of the disease or damage by reprogramming fibroblasts at the site. Examples of such neurological diseases include Parkinson's disease.

[00189] Methods of treating a disease or disorder of the blood according to aspects of the present invention include administering: protein transduction reagent-modified Oct4. Such methods produce new hematopoietic cells at the site of the disease or damage by reprogramming fibroblasts at the site.

[00190] Methods of treating diabetes, a pancreatic disease or pancreatic tissue damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Ngn3, protein transduction reagent-modified Pdx1 and protein transduction reagent-modified Pax4. Methods of treating diabetes, a pancreatic disease or

pancreatic tissue damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Ngn3, protein transduction reagent-modified Pdx1 and protein transduction reagent-modified Pax4. Methods of treating diabetes, a pancreatic disease or pancreatic tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified Ngn3, protein transduction reagent-modified Pdx1 and protein transduction reagent-modified Pax4. Such methods produce new pancreatic beta-cells at the site of the disease or damage by reprogramming fibroblasts at the site.

[00191] Methods of treating diabetes, a pancreatic disease or pancreatic tissue damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Ngn3, protein transduction reagent-modified Pdx1 and protein transduction reagent-modified MafA. Methods of treating diabetes, a pancreatic disease or pancreatic tissue damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Ngn3, protein transduction reagent-modified Pdx1 and protein transduction reagent-modified MafA. Methods of treating diabetes, a pancreatic disease or pancreatic tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified Ngn3, protein transduction reagent-modified Pdx1 and protein transduction reagent-modified MafA. Such methods produce new pancreatic beta-cells at the site of the disease or damage by reprogramming pancreatic exocrine cells at the site.

[00192] Methods of treating obesity according to aspects of the present invention include administering one or both of: protein transduction reagent-modified Prdm16 and protein transduction reagent-modified C/EBPb. Methods of treating obesity according to aspects of the present invention include administering both: protein transduction reagent-modified Prdm16 and protein transduction reagent-modified C/EBPb. Such methods produce new brown adipocytes at the site of the disease or damage by reprogramming white adipocytes at the site.

[00193] Methods of treating a muscle disease or muscle damage according to aspects of the present invention include administering: protein transduction reagent-modified MyoD. Such methods produce new muscle cells at the site of the disease or damage by reprogramming fibroblasts at the site.

[00194] Methods of treating arthritis, osteoarthritis, cartilage degeneration and/or cartilage injury according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Runx2, protein transduction reagent-modified Sox5 and protein transduction reagent-modified Sox6. Methods of

treating arthritis, osteoarthritis, cartilage degeneration and/or cartilage injury according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Runx2, protein transduction reagent-modified Sox5 and protein transduction reagent-modified Sox6. Methods of treating arthritis, osteoarthritis, cartilage degeneration and/or cartilage injury according to aspects of the present invention include administering three or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Runx2, protein transduction reagent-modified Sox5 and protein transduction reagent-modified Sox6. Methods of treating arthritis, osteoarthritis, cartilage degeneration and/or cartilage injury according to aspects of the present invention include administering: protein transduction reagent-modified Sox9, protein transduction reagent-modified Runx2, protein transduction reagent-modified Sox5 and protein transduction reagent-modified Sox6. Such methods produce new chondrocytes at the site of the disease or damage by reprogramming fibroblasts at the site.

[00195] Methods of treating a breast disease or breast tissue damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Slug, protein transduction reagent-modified Stat5a, protein transduction reagent-modified Gata3 and protein transduction reagent-modified Brca-1. Methods of treating a breast disease or breast tissue damage according to aspects of the present invention include administering protein transduction reagent-modified Brca-1. Methods of treating a breast disease or breast tissue damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Slug, protein transduction reagent-modified Stat5a, protein transduction reagent-modified Gata3 and protein transduction reagent-modified Brca-1. Methods of treating a breast disease or breast tissue damage according to aspects of the present invention include administering both: protein transduction reagent-modified Sox9 and protein transduction reagent-modified Slug. Methods of treating a breast disease or breast tissue damage according to aspects of the present invention include administering both: protein transduction reagent-modified Stat5a and protein transduction reagent-modified Gata3. Methods of treating a breast disease or breast tissue damage according to aspects of the present invention include administering three or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Slug, protein transduction reagent-modified Stat5a, protein transduction reagent-modified Gata3 and protein transduction reagent-modified Brca-1. Methods of treating a breast disease or breast tissue damage according to aspects of the present invention include administering four or more of: protein

transduction reagent-modified Sox9, protein transduction reagent-modified Slug, protein transduction reagent-modified Stat5a, protein transduction reagent-modified Gata3 and protein transduction reagent-modified Brca-1. Methods of treating a breast disease or breast tissue damage according to aspects of the present invention include administering: protein transduction reagent-modified Sox9, protein transduction reagent-modified Slug, protein transduction reagent-modified Stat5a, protein transduction reagent-modified Gata3 and protein transduction reagent-modified Brca-1. Such methods produce new mammary duct cells at the site of the disease or damage by reprogramming fibroblasts at the site.

[00196] Methods of treating a vascular disease or blood vessel damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Erg1, protein transduction reagent-modified Er71, protein transduction reagent-modified Fli1 and protein transduction reagent-modified Gata2. Methods of treating a vascular disease or blood vessel damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Erg1, protein transduction reagent-modified Er71, protein transduction reagent-modified Fli1 and protein transduction reagent-modified Gata2. Methods of treating a vascular disease or blood vessel damage according to aspects of the present invention include administering three or more of: protein transduction reagent-modified Erg1, protein transduction reagent-modified Er71, protein transduction reagent-modified Fli1 and protein transduction reagent-modified Gata2. Methods of treating a vascular disease or blood vessel damage according to aspects of the present invention include administering: protein transduction reagent-modified Erg1, protein transduction reagent-modified Er71, protein transduction reagent-modified Fli1 and protein transduction reagent-modified Gata2. Such methods produce new endothelial cells at the site of the disease or damage by reprogramming cells at the site.

[00197] Methods of treating a vascular disease or blood vessel damage according to aspects of the present invention include administering one or more of: protein transduction reagent-modified Hey1, protein transduction reagent-modified Hey2, protein transduction reagent-modified FoxC1 and protein transduction reagent-modified FoxC2. Methods of treating a vascular disease or blood vessel damage according to aspects of the present invention include administering two or more of: protein transduction reagent-modified Hey1, protein transduction reagent-modified Hey2, protein transduction reagent-modified FoxC1 and protein transduction reagent-modified FoxC2. Methods of treating a vascular disease or blood vessel damage according to aspects of the present invention include administering three or more of: protein transduction reagent-modified Hey1, protein transduction reagent-modified Hey2, protein transduction reagent-modified FoxC1 and protein

transduction reagent-modified FoxC2. Methods of treating a vascular disease or blood vessel damage according to aspects of the present invention include administering: protein transduction reagent-modified Hey1, protein transduction reagent-modified Hey2, protein transduction reagent-modified FoxC1 and protein transduction reagent-modified FoxC2. Such methods produce new arterial endothelial cells at the site of the disease or damage by reprogramming cells at the site.

[00198] Methods of treating a vascular disease or blood vessel damage according to aspects of the present invention include administering one or both of: protein transduction reagent-modified Sox7 and protein transduction reagent-modified Sox18. Methods of treating a vascular disease or blood vessel damage according to aspects of the present invention include administering both: protein transduction reagent-modified Sox7 and protein transduction reagent-modified Sox18. Such methods produce new venous endothelial cells at the site of the disease or damage by reprogramming cells at the site.

[00199] Combination Treatments

[00200] Combinations of therapeutic agents are administered according to aspects of the present invention.

[00201] In some aspects, two or more QQ-modified proteins are administered to a subject to treat a disorder in a subject in need thereof.

[00202] In further aspects, at least one QQ-modified protein and at least one additional therapeutic agent are administered to a subject to treat a disorder in a subject in need thereof.

[00203] In still further aspects, at least one QQ-modified protein and at least two additional therapeutic agents are administered to a subject to treat a disorder in a subject in need thereof.

[00204] In some aspects, two or more QQ-modified proteins are administered to a subject to treat cancer in a subject in need thereof. In further aspects, at least one QQ-modified protein and at least one additional therapeutic agent are administered to a subject to treat cancer in a subject in need thereof. In still further aspects, at least one QQ-modified protein and at least two additional therapeutic agents are administered to a subject to treat cancer in a subject in need thereof.

[00205] The term “additional therapeutic agent” is used herein to denote a chemical compound, a mixture of chemical compounds, a biological macromolecule (such as a nucleic acid, an antibody, a protein or portion thereof, e.g., a peptide), or an extract made from biological materials such as bacteria, plants, fungi, or animal (particularly mammalian) cells or tissues which is a biologically, physiologically, or pharmacologically active substance (or substances) that acts locally or systemically in a subject.

[00206] Additional therapeutic agents included in aspects of methods and compositions of the present invention include, but are not limited to, antibiotics, antivirals, antineoplastic agents, analgesics, antipyretics, antidepressants, antipsychotics, anti-cancer agents, antihistamines, anti-osteoporosis agents, anti-osteonecrosis agents, antiinflammatory agents, anxiolytics, 5 chemotherapeutic agents, diuretics, growth factors, hormones, non-steroidal anti-inflammatory agents, steroids and vasoactive agents.

[00207] Combination therapies utilizing one or more QQ-modified proteins and one or more additional therapeutic agents may show synergistic effects, e.g., a greater therapeutic effect than would be observed using either the one or more QQ-modified proteins or one or more additional 10 therapeutic agents alone as a monotherapy.

[00208] According to aspects, combination therapies include: (1) pharmaceutical compositions that include one or more QQ-modified proteins in combination with one or more additional therapeutic agents; and (2) co-administration of one or more QQ-modified proteins with one or more additional therapeutic agents wherein the one or more QQ-modified proteins and the one or 15 more additional therapeutic agents have not been formulated in the same composition. When using separate formulations, the one or more QQ-modified proteins may be administered at the same time, intermittent times, staggered times, prior to, subsequent to, or combinations thereof, with reference to the administration of the one or more additional therapeutic agents.

[00209] Combination treatments can allow for reduced effective dosage and increased 20 therapeutic index of the one or more QQ-modified proteins and the one or more additional therapeutic agents used in methods of the present invention.

[00210] Optionally, a method of treating a subject having cancer or at risk of having cancer further includes an adjunct anti-cancer treatment. An adjunct anti-cancer treatment can be administration of an anti-cancer agent.

25 [00211] Anti-cancer agents are described, for example, in Goodman et al., Goodman and Gilman's The Pharmacological Basis of Therapeutics, 8th Ed., Macmillan Publishing Co., 1990.

[00212] Anti-cancer agents illustratively include acivicin, aclarubicin, acodazole, acronine, adozelesin, aldesleukin, alitretinoin, allopurinol, altretamine, ambomycin, ametantrone, amifostine, aminoglutethimide, amsacrine, anastrozole, anthramycin, arsenic trioxide, asparaginase, asperlin, 30 azacitidine, azetepa, azotomycin, batimastat, benzodepa, bicalutamide, bisantrene, bisnafide dimesylate, bizelesin, bleomycin, brequinar, bropiramine, busulfan, cactinomycin, calusterone, capecitabine, caracemide, carbetimer, carboplatin, carmustine, carubicin, carzelesin, cedefingol, celecoxib, chlorambucil, cirolemycin, cisplatin, cladribine, crisnatol mesylate, cyclophosphamide,

cytarabine, dacarbazine, dactinomycin, daunorubicin, decitabine, dexormaplatin, dezaguanine,
 dezaguanine mesylate, diaziquone, docetaxel, doxorubicin, droloxifene, dromostanolone,
 duazomycin, edatrexate, eflomithine, elsamitrucin, enloplatin, enpromate, epipropidine, epirubicin,
 5 erbulozole, esorubicin, estramustine, etanidazole, etoposide, etoprine, fadrozole, fazarabine,
 fenretinide, floxuridine, fludarabine, fluorouracil, flurocitabine, fosquidone, fostriecin, fulvestrant,
 gemcitabine, hydroxyurea, idarubicin, ifosfamide, ilmofosine, interleukin II (IL-2, including
 recombinant interleukin II or rIL2), interferon alfa-2a, interferon alfa-2b, interferon alfa-n1,
 interferon alfa-n3, interferon beta-1a, interferon gamma-1b, iproplatin, irinotecan, lanreotide,
 letrozole, leuprolide, liarozole, lometrexol, lomustine, losoxantrone, masoprocol, maytansine,
 10 mechlorethamine hydrochloride, megestrol, melengestrol acetate, melphalan, menogaril,
 mercaptopurine, methotrexate, metoprine, meturedopa, mitindomide, mitocarcin, mitocromin,
 mitogillin, mitomalcin, mitomycin, mitosper, mitotane, mitoxantrone, mycophenolic acid,
 nelarabine, nocodazole, nogalamycin, ormaplatin, oxisuran, paclitaxel, pegaspargase, peliomycin,
 pentamustine, peplomycin, perfosfamide, pipobroman, piposulfan, piroxantrone hydrochloride,
 15 plicamycin, plomestane, porfimer, porfiromycin, prednimustine, procarbazine, puromycin,
 pyrazofurin, riboprine, rogletimide, safingol, semustine, simtrazene, sparfosate, sparsomycin,
 spirogermanium, spiromustine, spiroplatin, streptonigrin, streptozocin, sulofenur, talisomycin,
 tamoxifen, tecogalan, tegafur, teloxantrone, temoporfin, teniposide, teroxirone, testolactone,
 thiamiprine, thioguanine, thiotepa, tiazofurin, tirapazamine, topotecan, toremifene, trestolone,
 20 triciribine, trimetrexate, triptorelin, tubulozole, uracil mustard, uredepa, vapreotide, verteporfin,
 vinblastine, vincristine sulfate, vindesine, vinepidine, vinglycinat, vinleurosine, vinorelbine,
 vinrosidine, vinzolidine, vorozole, zeniplatin, zinostatin, zoledronate, and zorubicin.

[00213] An adjunct anti-cancer treatment can be a radiation treatment of a subject or an affected area of a subject's body.

25 **[00214]** Embodiments of inventive compositions and methods are illustrated in the following
 examples. These examples are provided for illustrative purposes and are not considered limitations
 on the scope of inventive compositions and methods.

[00215] Examples

[00216] Example 1

30 **[00217]** Reprogramming protein expression and preparation

[00218] Plasmid construction

[00219] DNA encoding each of Oct4, Sox2, Klf4, and c-Myc, Nanog, GATA4, Hand2, Mef2c and Tbx5 reprogramming proteins were separately subcloned into an sHT-pET30a bacterial

expression vector, in which a short his-tag: 'HHHHHHSS' (SEQ ID NO:64) replaced the long his-tag in the pET30a expression vector from Novagen. A factor Xa (IEGR) cleavage site is placed between the short his-tag and the coding genes. The sequences of the bacterial expression vectors were confirmed by DNA sequencing.

5 **[00220]** Protein expression and purification

[00221] The DNA constructs of reprogramming proteins were transformed into E.Coli strain BL-21(DE3) or ClearColi™ BL21(DE3) bacterial cells individually. The ClearColi™ BL21(DE3) bacterial cells were used for production of recombinant reprogramming proteins which are endotoxin-free. A single colony was selected for bacterial protein expression. After brief
10 optimization, protein expression was induced by 0.5-1.2 mM IPTG depending on different proteins and continued to culture at 18 C° for 12-16 hours. The cells were harvested in the binding buffer containing 6M urea and sonicated several times to extract proteins. The recombinant reprogramming proteins were purified using a His-Bind Resin column (Novagen) according to the manual with modifications. Usually the protein extraction supernatant was loaded on the column
15 twice, washed with 5 x column volume of the binding buffers. When using regular bacterial strain, BL21(DE3) to produce recombinant reprogramming proteins, an additional washing step can be added during protein purification stage using 10 column volume of 0.2% of Triton X-114 to remove endotoxin from bacterially expressed recombinant reprogramming proteins. The column loaded with recombinant proteins is washed again using 10 x column volume washing buffer
20 containing 15-50 mM imidazole. The purified proteins were eluted from the column using elution buffer containing 500 mM imidazole. The purified proteins were dialyzed extensively against water and lyophilized into protein powders.

[00222] Example 2

[00223] QQ-modification: Modification of expressed reprogramming proteins with protein
25 transduction reagent to produce protein transduction reagent-modified reprogramming proteins

[00224] The protein transduction reagent (QQ reagent) can be adjusted by altering the composition to include reagents as shown in Table 1 to obtain the best protein transduction efficiency for the particular reprogramming protein and cell type.

[00225] For in vivo applications, to make total volume 1ml:

Table 1								
Protein transduction reagent (QQ reagent)	PEI 1.2K	PEI 2K	PEI 4K	PEI 8K	DOTAP or DOTMA	DOPE or DOGS	POPC	DMPE

QQ1a	10-200 μ l	-	-	-	25-100 μ l	25-100 μ l	-	-
QQ2a	10-200 μ l	-	10-100 μ l	-	25-100 μ l	25-100 μ l	-	-
QQ3a	10-200 μ l	-	10-100 μ l	10-100 μ l	25-100 μ l	25-100 μ l	-	-
QQ4a	10-200 μ l	10-100 μ l	-	10-100 μ l	25-100 μ l	25-100 μ l	-	-
QQ5a	10-200 μ l	10-100 μ l	10-100 μ l	10-100 μ l	25-100 μ l	25-100 μ l	-	-
QQ6a	10-200 μ l	10-100 μ l	10-100 μ l	10-100 μ l	-	--	25-100 μ l	25-100 μ l
QQ7a	10-200 μ l	10-100 μ l	10-100 μ l	10-100 μ l	25-100 μ l	25-100 μ l	25-100 μ l	25-100 μ l
QQ8a	10-200 μ l	-	-	10-100 μ l	25-100 μ l	25-100 μ l	25-100 μ l	25-100 μ l
QQ9a	10-200 μ l	-	10-100 μ l	10-100 μ l	-	-	25-100 μ l	25-100 μ l

The polyethylenimine (PEI) concentration for the stock solution:

5 1.2K 2K 4K 8K
5mg/ml 2mg/ml 2mg/ml 2mg/ml

Lipid concentration for the stock solution:

10 DOTAP DOTMA DOPE DOGS POPC DMPE
1mg/ml 1mg/ml 1mg/ml 1mg/ml 1mg/ml 1mg/ml

PBS buffer containing reprogramming proteins in 1-6 M urea and specified PEI, lipids is used to make 1 ml total volume.

[00226] The reprogramming protein(s) is first dissolved in sodium phosphate buffer (pH7.0, NaCl 50 mM) at concentrations of 0.5-10 mg/ml, depending on protein solubility. Protein solubility was found to influence cationization efficiency. To completely dissolve proteins, an overnight stir of the protein solution at room temperature is performed (with or without DTT at 3 mM for overnight, depending on if the reprogramming protein has cysteine residues). Proteins can also be dissolved in 1-6 M urea to improve protein solubility.

20 [00227] A lipid DOTAP/DOPE (1:1) emulsion was prepared using a method as the following: 1 mg of DOTAP/DOPE (0.5 mg:0.5 mg=1:1) mixture was dissolved in chloroform and dried under N₂ gas. The dried lipid film was then dissolved in PBS buffer, pH7.0 and the lipid solution was

sonicated for 3x30 seconds using a power of 7-8 on a sonicator from Fisher Scientific (Sonic Dismembrator, Model 100) with micro probe. The lipid solution was further incubated at 37°C for 2 hours until the suspension becomes semi-clear. The prepared emulsion was stored at 4°C and is stable for one month.

5 **[00228]** The other ingredients of the QQ reagents (not including the lipid emulsion or the optional Ca or DMSO) were mixed in a tube, according to the recipe described above. The QQ reagent is then titrated into the protein solution very slowly, drop by drop, while stirring and then add the lipid emulsion. Once this is completed, the resulting protein solution is left at room temperature for 4 hours before use. During this period, gentle stirring is necessary to mix the QQ
10 reagent with protein solution and also to allow the protein modification reaction to complete. If precipitation is observed, the protein solution can be centrifuged at 14,000 rpm for 15 minutes to remove the precipitate. If the precipitate occurs, a BCA protein assay will be carried out using the supernatant to check the amount of protein remaining in solution. To ensure the efficiency of protein transfer into the cells, the concentration of modified protein has to be high enough at >0.1
15 mg/ml.

[00229] Typically, QQ-modified proteins are prepared at 0.5-1.5 mg/ml concentration, depending on protein solubility, for in vivo administration.

[00230] QQ modification may be performed on each protein individually or on a mixture of proteins to be administered together to a subject.

20 **[00231]** Optionally, the QQ modification is performed on each protein individually, the QQ-modified proteins mixed together for several hours and then aliquoted into small tubes at 1ml/tube and stored at -20°C where they are stable for several months.

[00232] If the majority of the reprogramming protein is precipitated, another QQ reagent can be used for protein modification. The QQ series reagents cover a wide range of cationization
25 reagents along with different lipids and enhancers, thus any precipitation problem is solved. The above procedure can be repeated to prepare higher concentrations of protein transduction reagent-modified reprogramming protein.

[00233] The protein transduction reagent-modified reprogramming protein is passed through a desalting column to separate the protein transduction reagent-modified reprogramming protein
30 from remaining unreacted materials. The purified proteins are passed through a filter (0.22 µm cutoff, for sterilization before in vivo administration. The purified protein fractions can be concentrated before or after sterilization, such as by using a spin column, and are stable and can be stored at -20°C for between a few weeks to a few months.

[00234] Different QQ reagents can also be used for the best efficiency of protein transfer as well as the least cell toxicity. In addition, different proteins are modified with different QQ reagents for best efficiency of protein transfer into cells. In general, for better in vivo delivery efficiency, QQ5a-QQ9a are used. However, this may cause larger in vivo toxicity. When using
5 QQ5a-QQ9a, use of lower concentrations of larger PEI and lipids is emphasized.

[00235] For good in vivo delivery efficiency, QQ1a-QQ4a is used with less in vivo toxicity. QQ1-QQ4 can be used with higher lipid concentrations.

[00236] Example 3

[00237] The reprogramming proteins were dissolved in 50 mM sodium phosphate pH 7.4 with
10 2 M urea. Protein transduction reagents (QQ-reagents) were freshly prepared based on the recipe. QQ-reagent used in this example is a cocktail of polyethylenimine (PEI) 1,200 (1.2K, 0.05-1.0 mg/ml) and DOTAP/DOPE (25-100µg/ml). The QQ-modification of reprogramming proteins was performed by mixing the QQ-reagent with one or more reprogramming proteins, such as Oct4/Sox2/Nanog: 1 mg/ml, for 4-hours at room temperature or overnight in a cold room.

15 [00238] Example 4

[00239] QQ-Protein delivery into brain tumors in rats by an intravenous injection

[00240] Ferritin is a ubiquitous iron-containing protein useful as a negative contrast reagent for magnetic resonance imaging (MRI).

[00241] In this example, ferritin or ferritin treated with QQ reagent to produce QQ-ferritin is
20 intravenously injected into rats having brain tumors generated by implantation of rat gliosarcoma cell line 9L cells.

[00242] Cells of a rat gliosarcoma cell line 9L were freshly prepared and adjusted to 1×10^6 cells/ml before implantation. The intracranial xenografts were performed according to standard protocols. Fisher rats were anesthetized and placed in a stereotactic frame, and the skull was
25 exposed. A hole was made 3 mm to the right and 1 mm anterior of the bregma, and 9L cells (5×10^4 cells in 5 µl) were injected using a 10-µl Hamilton syringe with a 26s-gauge needle mounted in a stereotactic holder (Bonaduz, GR, Switzerland). The syringe was lowered to a depth of 3.5 mm and then raised to a depth of 3.0 mm. The tumor cells were injected at a rate of 0.5 µl/10 s, and the craniotomy covered with Horsley's bone wax. 12-days after 9L-cell implantation, QQ-modified
30 ferritin or ferritin alone, 100µg/100µl/rat, was injected via the rat tail vein. Four hours after injection, the rats were anesthetized and the rat brains were imaged by MRI. After MRI, the animals were sacrificed. The extracted brains were fixed in 4% paraformaldehyde overnight and then embedded in paraffin for further analysis. Six micrometer thin sections were cut from each of

the blocks and were stained with hematoxylin and eosin. The sizes of brain tumors were microscopically determined.

[00243] Figures 1A and 1B show MRI evidence of targeted delivery of QQ-ferritin into 9L-brain tumor in rats and no detectable ferritin in 9L-brain tumor of rats injected with ferritin alone. Figure 1A is a representative MRI image of a rat with 9L-brain tumor injected with QQ-ferritin via tail-vein. The arrow points to a 9L-brain tumor and the black color is indicative of ferritin delivered to the tumor. Figure 1B: A representative MRI image of a rat with 9L-brain tumor injected with ferritin without QQ modification. No ferritin is observed in the brain tumor in this case although the rat was sacrificed and confirmed to have a large 9L-brain tumor.

[00244] Double immunostains of the QQ-ferritin injected 9L-brain tumor tissue sections were performed using an antibody for a blood vessel marker, CD31, and an antibody for ferritin. The nuclei of the 9L-cancer cells in the sections were stained with DAPI. Overlapping immunostaining for the CD31 and ferritin were observed in blood vessels of the brain tumors, indicating that QQ-ferritin leaked out from the blood vessels by the enhanced permeability and retention effect (EPR effect) of the angiogenic blood vessels associated with brain tumor. In the QQ-ferritin treated animal, immunostaining for ferritin was observed inside the cytosol of the tumor cells near or next to the blood vessels, indicating that the QQ-ferritin penetrated into tumor cells. In contrast, ferritin immunostaining in the 9L-tumor sections treated with ferritin without QQ-modification showed that the ferritin without QQ modification did not penetrate into tumor cells, although immunostaining indicated that this unmodified ferritin also leaked out of the tumor-associated angiogenic blood vessels. These unmodified ferritin nanoparticles were observed between cells and likely did not accumulate due to the fluid between cells inside the tissues so that no tumor was observed by MRI without QQ-modification of the ferritin.

[00245] These data provide MRI evidence of the targeted delivery of ferritin to brain tumor in rats by the QQ-protein delivery technology, which is confirmed by histology analysis and immunostaining of the brain tissue sections.

[00246] Example 5

[00247] QQ-protein delivery into breast cancer in mice by an intravenous injection.

[00248] In this example, ferritin or ferritin treated with QQ reagent to produce QQ-ferritin is intravenously injected into mice having orthotopic breast tumors generated by implantation of mouse metastatic breast cancer cell line 4T1 cells. Implantation of mouse metastatic breast cancer cell line 4T1 cells in mice is a mouse model for spontaneous metastatic breast cancer.

[00249] 4T1 cells were implanted into the breast (#4 fat pad) of 2 month old female BALB/c mice (20,000 4T1 cells in 50 μ l/mouse). At day 18 after 4T1 cell implantation, 4T1 tumors had grown in the mouse breast to a size of ~1.0-1.5 cm in diameter. QQ-modified ferritin (100 μ g/50 μ l/mouse) was injected via tail-vein. The mice were anesthetized and subjected to MRI imaging before injection of the QQ-modified ferritin and at 0.5 hours, 2 hours, 3.5 hours, 8 hours and 18 hours following the injection.

[00250] Figures 2A-F are representative MRI images from this time course showing QQ-delivered ferritin in the primary tumor of breast cancer bearing mouse as darkened areas in the primary tumor. This primary tumor darkness was time-dependent and reached a maximum at 8-hours, then started to decrease in intensity.

[00251] MRI results were confirmed by histology analysis using Prussian blue stain and ferritin immunostaining. Immunohistochemical stain and Prussian blue stain of the primary tumor tissue sections of 4T1-breast cancer bearing mouse at 18-hour after QQ-ferritin injection showed positive ferritin and Prussian blue stains inside the primary tumor. This histological result confirmed the MRI observation of QQ-delivered ferritin inside the primary tumor of 4T1-tumor bearing mice.

[00252] Example 6

[00253] QQ-protein delivery of Oct4 and Klf4 protein into the nuclei of 4T1-breast cancer cells of the primary tumor of 4T1 breast cancer bearing mice

[00254] 4T1 cells (2×10^4 cells/50 μ l/mouse) were implanted into female BALB/c mice. 15-days after tumor cell implantation Alexa Fluor594 labeled Klf4 was prepared. QQ-modified unlabeled Oct4 (QQ-Oct4) and QQ-modified Alexa Fluor594 labeled Klf4 were then prepared. Intra-tumoral injection of 50 μ g QQ-Oct4 and QQ-Alexa Fluor594 labeled Klf4 proteins was performed. The mice were sacrificed 5 hours after injection and tumor tissue sections were prepared for immunostaining and fluorescence microscopy. Immunostaining for Oct4 in the primary 4T1-breast cancer tissue sections demonstrated nuclear localization of Oct4 in the 4T1-cancer cells and co-localization with DAPI. Fluorescence microscopic imaging also showed nuclear localization of Klf4 in the primary 4T1-breast cancer tissue sections and co-localization with DAPI.

[00255] Example 7

[00256] QQ-protein delivery of GATA4, OCT4 and SOX2 to the injured hearts of rats after acute myocardial infarction

[00257] QQ-modified and fluorescence labeled GATA4, Sox2 and Oct4 are administered by an intraperitoneal injection to rats after myocardial infarction and the QQ-modified proteins and fluorescence labeled were delivered to injured heart tissues in the animals.

[00258] GATA4, Sox2 and Oct4 are each labeled with Alexa Fluor488 according to methods specified by the manufacturer. The labeled GATA4, Sox2 and Oct4 proteins were purified with small spin columns. The Alexa Fluor488-GATA4, Alexa Fluor488-Sox2 and Alexa Fluor488-Oct4 are then QQ-modified to generate QQ-modified Alexa Fluor488-GATA4, QQ-modified Alexa Fluor488-Sox2 and QQ-modified Alexa Fluor488-Oct4.

[00259] Lewis rats were anesthetized and coronary artery ligation surgery was performed to create myocardial infarction in the Lewis rats. 24-hours after coronary artery ligation surgery, QQ-modified Alexa Fluor488-labeled GATA4, QQ-modified Alexa Fluor488-labeled Sox2 and QQ-modified Alexa Fluor488-labeled Oct4 (200 µg/100 µl/rat) were intraperitoneally injected into Lewis rats. 7-hours after fluorescence labeled protein injections, the rats were sacrificed and heart tissue sections were prepared for either Trichrome staining for infarction sizes or for immunostaining with an α -actinin antibody, which is a heart muscle marker indicative of mature cardiomyocytes.

[00260] Fluorescence microscopic analysis showed colocalization of Alexa Fluor488-GATA4 protein with α -actinin immunofluorescence, indicating that intraperitoneal injection of QQ-fluorescence labeled GATA4 resulted in delivery of the GATA4 protein into the scar zone of the injured heart tissue after coronary artery ligation. Similarly, fluorescence microscopic analysis of Alexa Fluor488-Sox2 and α -actinin immunofluorescence indicated that intraperitoneal injection of QQ-modified Alexa Fluor488 labeled-Sox2 resulted in location of the Sox2 protein in the border zone of the injured heart tissue after coronary artery ligation. Finally, fluorescence images of Alexa Fluor488-Oct4 showed localization of injected QQ-modified Alexa Fluor488-Oct4 in the remote normal heart tissue zone. These results indicate that fluorescence labeled and QQ-modified GATA4, Sox2 and Oct4 administered by intraperitoneal injection reached injured heart tissues in the scar zone, the border zone and the normal heart tissue zones. These data demonstrate that QQ-protein delivery targeted deliver these proteins into the injured heart tissues.

[00261] Example 8

[00262] Malignant glioma cells including 9L-, U251-, U87- and a primary GBM cell line from a patient were reprogrammed into protein induced pluripotent stem cells (piPSCs) using QQ-modified Sox2, Oct4 and Nanog (SON) proteins as follows: At day 0, glioma cells were seeded for 24-hours to allow them to grow in a 50 mm petri dish. At day 2, QQ-SON proteins were added to

the culture medium at 0.5-2.0 $\mu\text{g/ml}$ and cultured for 24 hours. Next day, fresh QQ-SON proteins were added into new culture medium and cultured for 24 hours. Such cycles were repeated for 5-7 cycles depending on glioma cells used for cell reprogramming. At the end of cell reprogramming, the culture media were changed to maintaining medium for 2 days. During cell reprogramming, piPSC colonies appeared. At day 8, whole dish passage was performed to expand the generated piPSCs. Immunostaining of the whole dish passaged glioma-piPSCs including both monolayer and colony of the 9L-piPSCs using pluripotency markers including nuclear markers Oct4, Nanog, Rex-1 and surface markers ALP, Tra1-60, and Tra1-81 showed nearly all cells were positive for these pluripotency markers. In contrast, immunostains of 9L-cells showed negative stains for these markers. These data indicate high conversion efficiency of cell reprogramming of 9L-cells into 9L-piPSCs using the QQ-SON proteins.

[00263] This is confirmed by immunostaining of whole well (96-well) of 9L-piPSCs using Nanog and Oct4. When counting positive 9L-piPSC clones compared with negative clones, an average of $96\pm 2\%$ conversion efficiency is found, see Table 2. A control immunostaining of 9L cells showed negative stains for these two pluripotency markers Nanog and Oct4. The conversion efficiency of 9L-piPSCs from 9L cells was calculated by positively stained colonies and monolayer 9L-piPSCs versus total colonies and monolayer cells as stained with DAPI using the software provided by EVOS Auto fluorescence microscope. Immunostaining of the 9L-piPSCs of duplicated 96-wells were used to calculate average conversion efficiency and standard deviation.

Table 2: Conversion efficiency of QQ-SON protein-induced cell reprogramming of 9L-gliosacoma cells			
Marker	Oct4	Nanog	Average
Whole Wells Counted	2	2	2
Conversion Efficiency (%)	94.0 ± 1.0	98.0 ± 0.1	96.0 ± 2.0

Cell reprogramming of 9L cells to generate 9L-piPSCs using the QQ-SON proteins (0.5ug/ml. Sox2:Oct4:Nanog=1:1:1, 5 continue cycles).

[00264] Using a similar cell reprogramming protocol, 4T1-piPSCs were generated from 4T1 mouse breast cancer cells. The reprogramming conversion efficiency of 4T1 cells into 4T1-piPSCs is $96\pm 1.3\%$ using the same whole well counting method (96-wells), see Table 3.

Table 3: Conversion efficiency of QQ-SON protein-induced cell reprogramming of 4T1-breast cancer cells

Pluripotent Marker	Number of Wells	Number of Colonies/Well	Conversion Efficiency (%)
Nanog	2	1344±200	98±1
Rex 1	2	1943±234	96±1
Sox 2	2	1862±198	94±2
Average			96±1.3

[00265] Example 9

[00266] Cells of human glioblastoma cell line U251 are reprogrammed into U251-piPSCs using QQ-modified Sox2, Oct4 and Nanog (SON) proteins as follows: At day 0, U251 glioblastoma cells were seeded for 24-hours to allow them to grow in a 50 mm petri dish. At day 2, QQ-SON proteins were added to the culture medium at 0.5-2.0 µg/ml and cultured for 24 hours. Next day, fresh QQ-SON proteins were added into new culture medium and cultured for 24 hours. Such cycles were repeated for 5-7 cycles. At the end of cell reprogramming, the culture media were changed to maintaining medium for 2 days. During cell reprogramming, U251-piPSCs colonies appeared. At day 8, whole dish passage was performed to expand the generated U251-piPSCs.

[00267] Proliferative activity of U251 and U251-piPSCs was assayed by determining the number of Ki67-positive cells. Figure 3A is a graph showing results of this assay indicating that cell reprogramming of glioma cells into piPSCs significantly reduces their proliferation, error Bar represents standard error of three independent experiments.

[00268] Example 10

[00269] Cells of a patient-derived primary human glioblastoma multiforme (GBM) cell line, GBM (12-14), is reprogrammed into GBM-piPSCs using QQ-modified Sox2, Oct4 and Nanog (SON) proteins as follows: At day 0, GBM cells were seeded for 24-hours to allow them to grow in a 50 mm petri dish. At day 2, QQ-SON proteins were added to the culture medium at 0.5-2.0 µg/ml and cultured for 24 hours. Next day, fresh QQ-SON proteins were added into new culture medium and cultured for 24 hours. Such cycles were repeated for 5-7 cycles. At the end of cell reprogramming, the culture media were changed to maintaining medium for 2 days. During cell reprogramming, GBM1-piPSCs colonies appeared. At day 8, whole dish passage was performed to expand the generated GBM-piPSCs.

[00270] Dose-dependent chemotherapeutic drug sensitivity of patient-derived primary human GBM (12-14) cells (squares) and GBM (12-14)-piPSCs (circles) against the alkylating agent temozolomide was determined and results are shown in Figure 3B. Experiments were done in

triplicate, mean $\pm$ SD, $p < 0.01$. Drug sensitivity of the GBM-piPSCs was significantly enhanced to temozolomide as compared with the parental glioblastoma cells.

[00271] Example 11

[00272] Cells of rat gliosarcoma cell line 9L are reprogrammed into 9L-piPSCs using QQ-modified Sox2, Oct4 and Nanog (SON) proteins as follows: At day 0, 9L glioma cells were seeded for 24-hours to allow them to grow in a 50 mm petri dish. At day 2, QQ-SON proteins were added to the culture medium at 0.5-2.0 $\mu\text{g/ml}$ and cultured for 24 hours. Next day, fresh QQ-SON proteins were added into new culture medium and cultured for 24 hours. Such cycles were repeated for 5-7 cycles. At the end of cell reprogramming, the culture media were changed to maintaining medium for 2 days. During cell reprogramming, 9L-piPSCs colonies appeared. At day 8, whole dish passage was performed to expand the generated 9L-piPSCs.

[00273] Dose-dependent chemotherapeutic drug sensitivity of 9L-cells (squares) and 9L-piPSCs (circles) against carboplatin was determined by an MTT assay and results are shown in Figure 3C. Experiments were done in triplicate, mean $\pm$ SD, $p < 0.01$. Drug sensitivity of the 9L-piPSCs was significantly enhanced to carboplatin as compared with the parental 9L cells.

[00274] Example 12

[00275] Cells of mouse mammary tumor cell line 4T1 are reprogrammed into 4T1-piPSCs using QQ-modified Sox2, Oct4 and Nanog (SON) proteins as follows: At day 0, 4T1 cells were seeded for 24-hours to allow them to grow in a 50 mm petri dish. At day 2, QQ-SON proteins were added to the culture medium at 0.5-2.0 $\mu\text{g/ml}$ and cultured for 24 hours. Next day, fresh QQ-SON proteins were added into new culture medium and cultured for 24 hours. Such cycles were repeated for 5-7 cycles. At the end of cell reprogramming, the culture media were changed to maintaining medium for 2 days. During cell reprogramming, 4T1-piPSCs colonies appeared. At day 8, whole dish passage was performed to expand the generated 4T1-piPSCs.

[00276] Proliferative activity of 4T1-cells and 4T1-piPSCs was assayed by determining the number of Ki67-positive cells and it was found that cell reprogramming of 4T1 mammary tumor cells into 4T1-piPSCs significantly reduces their proliferation.

[00277] Mammary sphere formation of 4T1-piPSCs was reduced as compared with that of 4T1-cells.

[00278] Dose-dependent chemotherapeutic drug sensitivity of 4T1-cells and 4T1-piPSCs against doxorubicin and paclitaxel was determined using an MTT assay. Drug sensitivity of the 4T1-piPSCs was significantly enhanced to doxorubicin and paclitaxel as compared with the parental 4T1 cells.

[00279] Example 13

[00280] Glioma cell-derived piPSC differentiation into three germ layers in vitro

[00281] Embryonic bodies (EBs) were prepared for glioma-piPSCs using the hanging-drop method. The glioma-piPSCs EBs were placed into a spontaneous differentiation medium for 10-
5 days, followed by immunostaining with markers of ectoderm, mesoderm and endoderm, showing positive immunostains of PAX6 for ectoderm, positive immunostains of desmin for mesoderm and positive immunostains of α -fetoprotein (AFP) for endoderm. Immunostaining of control 9L cells was negative for these markers. Similar results have been also obtained for 4T1-piPSCs, indicating that the 4T1-piPSCs also differentiated into three germ layers.

10 [00282] Example 14

[00283] Glioma cell-derived piPSC differentiation into neural-lineage in vitro

[00284] Glioma-piPSCs EBs were placed into a specific neural-lineage inducing differentiation medium for 14-days, followed by immunostaining with neural lineage markers, showing positive
immunostains of Tuj I for neurons, positive immunostains of GFAP for astrocytes and positive
15 immunostains of nestin for neural stem cells. These neural-specifically differentiated cells also displayed neural cell morphology. Immunostaining of control 9L cells was negative for these markers.

[00285] Example 15

[00286] Cell reprogramming causes a mesenchymal-to-epithelial transition (MET) and
20 reduction of tumorigenicity in vitro.

[00287] Results of Western blot analysis of several epithelial and mesenchymal markers of U251 cells and U251-piPSCs, including E-cadherin (E-cad), β -catenin (β -cat), vimentin (VMT), see Figure 4A, indicate an MET during cell reprogramming which is confirmed by cell morphology changes and qRT-PCR results. GFAP protein level is also significantly enhanced after
25 cell reprogramming into U251-piPSCs from U251 cells, indicating that cell reprogramming caused a significant reduction of tumorigenicity of U251 cells in vitro during cell reprogramming. Actin protein expression was used as an internal control. Immunostaining of U251 cells and U251-piPSCs using an anti-GFAP antibody, confirmed up-regulated protein expression of GFAP in U251-piPSCs as shown in Western blot. In addition, the GFAP-expressing cells displayed
30 morphologies of neural cells.

[00288] The tumorigenicity reduction of U251 cells after cell reprogramming has been confirmed by in vitro assays for cell migration, Figure 4B, and invasion, Figure 4C, of U251-cells

and U251-piPSCs, indicating significantly reduced migration and invasion of U251-piPSCs ($p < 0.01$).

[00289] Similarly, immunostainings of 4T1-piPSCs and 4T1 cells using an epithelial marker: E-cadherin and a mesenchymal marker fibronectin were compared and showed that while 4T1 cells display positive E-cadherin immunostaining, 4T1-piPSCs showed much stronger E-cadherin immunostaining. In contrast, 4T1 cells displayed stronger fibronectin immunostaining and 4T1-piPSCs exhibited nearly negative fibronectin immunostaining. This results confirm a MET of 4T1 cells during cell reprogramming into 4T1-piPSCs. In addition, 4T1-piPSCs also exhibited reduced proliferation, Figure 4D, reduced mammary sphere formation, Figure 4E, reduced migration, Figure 4F, and reduced invasion, Figure 4G, as compared to those properties of the parental 4T1 cells, indicating a significantly reduced tumorigenicity of 4T1 cells after cell reprogramming into 4T1-piPSCs in vitro.

[00290] Example 16

[00291] Co-culture between glioma cells and glioma-piPSCs indicated that glioma-piPSCs displayed a bystander effect to change the phenotypes of their surrounding glioma cells and reduced tumorigenicity in vitro.

[00292] Cell morphology changes of U251 cells indirectly co-cultured with U251-piPSCs for 40-64 hours are observed indicating a mesenchymal-to-epithelial transition (MET) of the co-cultured glioma cells. Figure 5A is an image showing cell morphology of U251 cells and Figure 5B is an image showing changes in cell morphology of U251 cells indirectly co-cultured with U251-piPSCs for 40-64 hours.

[00293] This MET was confirmed by immunostaining of U251 cells, U251-piPSCs, first (1st) co-cultured U251 cells and second (2nd) co-cultured U251 cells using an anti-pan-cadherin antibody, showing cytosol/nuclear localization of pan-cadherin in U251 cells and membrane/nuclear localization of U251-piPSCs and 1st/2nd co-cultured U251 cells, suggesting a major enhancement of cell-cell adhesion among U251-piPSCs as well as among the 1st/2nd co-cultured U251 cells. This is an indication of MET during cell reprogramming and the 1st/2nd co-cultured U251 cells. A 2nd co-culture is the indirect co-culture experiment that places the 1st co-cultured cells into the transwell insert and fresh U251 cells in the basolateral chamber. Double immunostains of the U251 cells and indirectly co-cultured U251 cells using distinctly labeled anti-GFAP and anti-Tuj1 antibodies, showing significant GFAP protein expression in the co-cultured U251 cells with neural morphology whereas U251 cells (spindle morphology) do not express GFAP.

[00294] Results of proliferation assay of U251 cells, U251-piPSCs and indirectly co-cultured U251 cells by Ki-67 immunostaining show significantly reduced proliferation of U251-piPSCs and the co-cultured U251 cells, Figure 5C. Results of migration assays, Figure 5D, and invasion assays, Figure 5E, of U251 cells, U251-piPSCs, directly co-cultured U251 cells at 1:1 and 8:1 ratios (U251:U251-piPSCs), show significantly reduced migration and invasion of the U251 cells after co-cultures with U251-piPSCs ($p < 0.005$).

[00295] Following a similar co-culture experiment, MET of co-cultured 4T1 cells with 4T1-piPSCs was observed. Immunostains of co-cultured 4T1 cells and 4T1-piPSCs showed that both cells had high protein expression of E-cadherin and lower protein expression of fibronectin, confirming a MET of the co-cultured 4T1 cells. In addition, the co-cultured 4T1 cells also displayed a significantly reduced proliferation, Figure 5F, migration, Figure 5G, and invasion, Figure 5H, indicating a significantly reduced tumorigenicity of the 4T1 cells after co-cultured with 4T1-piPSCs for 40-hours.

[00296] Co-culture of patient-derived primary brain tumor cells, GBM (12-14), with U251-piPSCs also causes a mesenchymal-to-epithelial transition and reduced tumorigenicity of the co-cultured patient-derived primary brain tumor cells – a bystander effect of U251-piPSCs.

[00297] These results indicate that a bystander effect of piPSCs changes the phenotypes of surrounding cancer cells for tumorigenicity and malignancy reduction. This bystander effect indicates that a smaller number of piPSCs can change malignant phenotype of a large numbers of surrounding tumor cells via this bystander effect, serving a cellular mechanism of a cell-converting cancer therapy.

[00298] Example 17

[00299] Significantly reduced tumorigenicity and metastasis/infiltration inhibition in vivo of piPSCs generated from malignant cancer cells.

[00300] In a first set of animals, equal numbers of 9L cells or 9L-piPSCs (5×10^4 cells/5 μ l) were implanted into Fisher rats intracranially. In a second set of animals, equal numbers of 9L cells or 9L-piPSCs (1×10^6 cells/100 μ l) were implanted into Fisher rats subcutaneously.

[00301] Tumor volume (mm^3) of intracranial-implanted rats with either 9L cells or 9L-piPSCs, $n = 10$ was measured at day 14, see Figure 6A. Tumor weight in grams of subcutaneously-implanted rats with 9L cells and 9L-piPSCs was measured at day 25 after 4T1-cell implantation, $n = 6$, see Figure 6B. Results showed a significantly reduced volume of the intracranial 9L-piPSC tumors and a significantly reduced weight of the subcutaneous 9L-piPSC tumors as compared with those of 9L-tumors, indicating a significantly reduced tumorigenicity of 9L-cells after cell

reprogramming into 9L-piPSCs. This result is also confirmed by hematoxylin and eosin (H&E) staining of the tumors from 9L-cell and 9L-piPSC implanted animals, indicating a major inhibition of 9L-cell infiltration, since the 9L-piPSC intracranial tumor showed a clear intact border, whereas 9L-intracranial tumor showed a broken border with a major infiltration of the 9L-tumor cells into normal brain tissues.

[00302] Example 18

[00303] Significantly reduced tumorigenicity and metastasis/infiltration inhibition in vivo of piPSCs generated from malignant breast cancer cells.

[00304] Equal numbers of 4T1 cells and 4T1-piPSCs (2×10^4 cells/50 μ l) were implanted into the #4 fat pads of 2-3 month old female BALB/c mice (20 grams). 4T1 breast tumors were observed at day 5-7 after cell implantation. The tumor volume and body weight of the mice were monitored every day. At day 25, the mice were sacrificed and the tumors were weighed. For the survival experiment, animals were sacrificed once they reached tumor burden.

[00305] 4T1 and 4T1-piPSC tumor growth were measured by tumor volume, Figure 6C, and tumor weight, Figure 6D, at day 25 after 4T1- and 4T1-piPSC implantation, indicating a significant tumor stasis of 4T1 cells after cell reprogrammed into 4T1-piPSCs.

[00306] Lung metastases of the mice that were implanted with 4T1 cells and 4T1-piPSCs were analyzed at day 25 after cell implantation. While all the mice implanted with 4T1 cells displayed many metastatic lesions, no lung metastases were observed in the mice implanted with 4T1-piPSCs at day 25 after cell implantation, Figure 6E, indicating a major metastasis inhibition caused by cell reprogramming of 4T1 cells into 4T1-piPSCs. This result was confirmed by H&E stain of the lung tissue sections of 4T1-bearing mice and 4T1-piPSC bearing mice, showing no lung metastasis in the mice implanted with 4T1-piPSCs.

[00307] A Kaplan-Meier survival curve of the mice implanted with 4T1 cells and 4T1-piPSCs in the #4 fat pads, Figure 6F, demonstrates a significantly prolonged survival of the 4T1-piPSC bearing mice during this 250-day survival experiment. These results demonstrate that cell preprogramming of malignant cancer cells into piPSCs significantly reduces tumorigenicity and metastatic property of the parental cancer cells, thus significantly prolonging the survival of the mice implanted with 4T1-piPSCs.

[00308] Example 19

[00309] Tumor cell derived piPSCs differentiate into different normal tissues in vivo depending on different differentiating tissue environments.

[00310] H&E stains of the brain tissue sections of 9L-intracranial implanted rats showed major glioma angiogenesis into normal brain tissues and severe bleeding inside brain tumors, whereas much less angiogenesis and bleeding were observed in the 9L-piPSC-intracranial implanted rats. In addition, 9L-piPSC-intracranial implanted rats also showed newly differentiated neural rosettes near the tumors in the brain. Further, H&E stain of the tumor tissue slides of subcutaneously-implanted rats with 9L-piPSCs, showed immature sweat glands with cuboidal and low columnar epithelial cells and melanocytes with brown color of melanin within the tumors. H&E stains of tumor tissue slides of rats subcutaneously implanted with 9L-piPSCs also showed skin epithelium with two layers of epidermis and early simple columnar epithelial cells with longer shape and nuclei in the base of the cells.

[00311] Similarly, H&E staining of the tumor masses of the 4T1-piPSC-bearing mice showed in vivo differentiation of the implanted 4T1-piPSCs into normal breast tissues including adipocytes including both brown and white adipocytes and mature and immature mammary ducts.

[00312] These results indicate that tumor cell derived-piPSCs differentiate into different normal tissue cells in vivo depending on different differentiating tissue environments.

[00313] Example 20

[00314] Lineage tracing shows tumor cell derived piPSCs differentiate into normal tissues.

[00315] To ensure that tumor cell derived piPSCs differentiate into normal tissues, GFP-expressing 9L- and 4T1-cells were prepared.

[00316] GFP-9L-piPSCs and GFP-4T1-piPSCs were prepared.

[00317] Lineage-tracing experiments were performed using both cell types. For GFP-9L-piPSCs, the cells were subcutaneously implanted into left flanks of Fisher rats. For GFP-4T1-piPSCs, the cells were implanted into the #4 fat pad of female BALB/c mice. Small tumor mass were observed at day 10-15 for GFP-9L-piPSC-implanted rats and at day 20-25 days for GFP-4T1-piPSC-implanted mice. The animals were sacrificed at different days after cell implantation and tumor masses were collected and used to make tissue sections for H&E stains and immunostaining. Immunostaining with Tuj I and Nestin neural marker antibodies, adiponectin (an adipocytes marker), cytokeratin 8, 14, 18 (CK8, 14, 18) or cytokeratin 5,8 (CK5,8) (mammary duct markers) were performed using differentially detectable labels.

[00318] H&E stain of both the center and edges of tumor mass of the 4T1-piPSC-bearing mice showed appearance of potential adipocytes and immature mammary ducts in the tumor mass.

[00319] Immunostains of the tissue section of the GFP-9L-piPSCs tumor mass obtained from the lineage tracing experiment using an anti-Tuj I antibody showed positive Tuj I immunostaining

that overlapped with GFP fluorescence of GFP-expressing cells of the same tissue sections. Only overlapping Tuj I and GFP fluorescence images were observed in the tissue sections, indicating that these Tuj I-positive cells originated from the GFP-expressing 9L-piPSCs. The Tuj I/GFP-positive cells displayed neuronal cell morphology. This data demonstrates that the GFP-expressing 9L-piPSCs differentiated in vivo into Tuj I-positive neuronal lineage cells.

[00320] Immunostains of the tissue section of the GFP-9L-piPSCs tumor mass obtained from the lineage tracing experiment using an anti-nestin antibody showed positive nestin immunostaining that overlapped with GFP fluorescence of GFP-expressing cells of the same tissue sections. Only overlapping nestin and GFP fluorescence images were observed in the tissue sections, indicating that these nestin-positive cells originated from the GFP-expressing 9L-piPSCs. This data demonstrates that the GFP-expressing 9L-piPSCs differentiated in vivo into nestin-positive neural lineage cells.

[00321] Immunostains of the tumor tissue sections of the 4T1-piPSC-bearing mice with an anti-adiponectin antibody showed adiponectin-positive immunostaining that overlapped with green fluorescence from the GFP-expressing cells. Only overlapping adiponectin and GFP fluorescence were observed in the tissue sections in cells with adipocyte morphology, indicating presence of adipocytes in the middle of the tumor mass of the 4T1-piPSC-bearing mice and these adipocytes originated from differentiation of the implanted GFP-4T1-piPSCs, since these adipocytes expressed GFP.

[00322] Immunostains of the tumor tissue sections of the 4T1-piPSC-bearing mice with anti-CK8,18,14 and anti-CK5,8 antibodies showed CK8,18,14- and anti-CK5,8-positive immunostaining that overlapped with green fluorescence from the GFP-expressing cells. Only overlapping CK8,18,14- or anti-CK5,8-positive and GFP fluorescence were observed in the tissue sections in cells with morphology of mammary ducts, indicating presence of immature mammary ducts that originated from differentiation of the implanted GFP-4T1-piPSCs, since these mammary ducts expressed GFP.

[00323] Example 21

[00324] QQ-SON protein-induced cell reprogramming treatment of tumors in vivo

[00325] To generate 9L-tumor bearing rats, 9L cells (1×10^6 cells/100 μ l) were implanted into Fisher rats subcutaneously. 5-days after 9L-implantation, QQ-SON protein was administered by intra-tumor injection every day at 1 μ g/day ($n = 5$), 5 μ g/day ($n = 5$), or 10 μ g/day ($n = 10$) for 18 daily treatments. QQ-reagent in PBS buffer was administered by intra-tumor injection to control

rats (n = 10). The tumor volume and body weight of the rats were monitored every day. At day 23, the rats were sacrificed and the tumors were weighed.

[00326] Tumor growth was measured by volume, see Figure 7A and tumor weight in grams, see Figure 7B. The rats were sacrificed at day 23 and tumors were weighed. This data indicates that QQ-SON treatment induced in situ cell reprogramming that results in tumor stasis.

[00327] To determine the effect of QQ-SON protein treatment on the survival of the rats subcutaneously implanted with 9L-cells, at day 5 after 9L cell implantation rats were treated with QQ-reagents in PBS (n = 8; median survival = 21 days) or QQ-SON proteins (10µg/day, n = 8; median survival = 127 days) for 30 daily treatments. Animals were sacrificed once they reached tumor burden (< 12 cm³). Kaplan–Meier survival curve (130-days) is shown in Figure 7C. This data indicated that QQ-SON protein treatment significantly prolonged the survival of 9L-tumor bearing rats.

[00328] Example 22

[00329] QQ-SON protein-induced cell reprogramming treatment of tumors in vivo

[00330] To generate orthotopic 4T1 breast cancer-bearing mice, 4T1 cells (2 x 10⁴ cells/50µl) were implanted into the #4 fat pads of 2-3 month old female BALB/c mice (20 grams). 4T1 breast tumors were observed at day 5-7 after cell implantation. The tumor volume and body weight of the mice were monitored every day. At day 25, the mice were sacrificed and the tumors were weighed. For the survival experiment, animals were sacrificed once they reached endpoints including tumor burden (< 2 gram), metastasis causing labored breath, uncontrollable pain, etc.

[00331] Various dosages of QQ-SON proteins were administered, 0.5 µg QQ-modified SON proteins/mouse, 1.25 µg QQ-modified SON proteins/mouse, or 2.5 µg QQ-modified SON proteins/mouse, and compared with QQ-PBS control. Efficacy was determined by measurement of tumor volume over a 25-day time course, Figure 7D, and measurement of tumor weight in grams, Figure 7E, at day 25.

[00332] In a further experiment, tumor volume of the mice treated either with QQ-SON proteins (n = 8) or QQ-PBS as the control (n = 8) during a 35-day time course was determined by MRI, see Figure 7F, and tumor weights of both groups were determined by weighing at day 35, see Figure 7G. QQ-SON protein treatment caused major tumor stasis without primary tumor removal.

[00333] Mice were analyzed to determine number and percentage of metastases in the lung, Figures 7I and 7M, lymph nodes, Figures 7J and 7N, liver, Figures 7K and 7O, and spleen, Figures 7L and 7P, of the QQ-SON and QQ-PBS treated 4T1-bearing mice without primary tumor removal as observed by MRI at indicated days. These results demonstrate that, as compared to QQ-PBS

treated mice, QQ-SON treated mice displayed metastasis at much later dates after 4T1-cell implantation and much less metastatic lesions in the lung, lymph nodes and no metastasis lesion was observed in the liver and spleen. This data indicates a major metastasis inhibition in the 4T1-bearing mice caused by QQ-SON protein treatment without primary tumor removal.

5 **[00334]** Example 23

[00335] Histological analysis of the tumor tissue sections treated with QQ-PBS and QQ-SON proteins without primary tumor removed.

[00336] H&E stains of a 9L-tumor tissue slides of a rat treated with QQ-reagents in PBS buffer for 18 daily treatments, showing uniform tumor cells with extensive angiogenesis. Immunostains of a nearby tissue slide with anti-VE-cadherin antibody confirmed angiogenesis. This rat was sacrificed at day 26 after subcutaneous 9L cell implantation with a tumor of 13.5 cm³ (around 15.5 g). In contrast, H&E stains of a tumor tissue slide of a rat treated with QQ-SON protein for 18 daily treatments showed different cell zones including tumor cell zone, connective tissue zone and fibroblast zone. This rat was sacrificed at day 49 after subcutaneous 9L cell implantation and its tumor was significantly reduced in size to 0.2 cm³ (around 0.6 g). Immunostain of the fibroblast zone of a nearby tumor tissue slide of the same rat using an anti- α -procollagen, a fibroblast marker, antibody, confirmed the fibroblast zoon (positively stained) and tumor cell zone (negatively stained). Immunostains of the tumor tissue slides of the same rat treated with QQ-SON proteins for 18 daily treatments using anti-GFAP and anti-Tuj1 antibodies showed positive stains of both markers with neuronal cell morphology and neural rosette formation. This data indicates that the injected QQ-SON proteins induced in situ generation of transient 9L-piPSCs that differentiate into non-cancerous cells including neuronal lineage cells.

[00337] Similarly, H&E stains of 4T1-tumor tissue slides of the mice treated with QQ-SON protein also showed adipocytes and mammary duct, which were immunostained positive with adiponectin and cytokeratin 5/8/14, indicating that the injected QQ-SON protein also induced in situ 4T1-tumor cell reprogramming inside the primary tumor into 4T1-piPSCs that differentiate into breast tissue.

[00338] Example 24

[00339] QQ-SON treatment significantly enhances the genome stability of the treated 9L-cells in vivo.

[00340] SKY genome analysis of explanted cells of two subcutaneous 9L-tumor-bearing rats, both treated with QQ-SON protein for 18-days was performed. One rat (#12) displayed major response to the QQ-SON treatment and tumor started to shrink. Before the tumor disappeared, this

rat was sacrificed and the tumor collected for both explants and tumor tissue slides. Another rat (#15) displayed no response to the QQ-SON protein treatment and tumor continued to grow to reach the ending point ($> 12 \text{ cm}^3$). This rat was also sacrificed and explants and tumor tissue slides were prepared.

[00341] The cells from the explant of #15 rat grew rapidly and reached confluency in 3-days. The cells displayed typical 9L-cell spindle morphology. However, the cells from the explant of #12 rat grew very slowly and only two-weeks later some cells with neural morphology with some colonies were observed. Immunostains of the tumor tissue slides with GFAP and Tuj I showed positive stains for both markers for those cells that displayed neural morphology in #12 rat, but negative stains for these two markers for #15 rat. Additional immunostains with β -catenin, CK5/6, E-cadherin, Lefty, Nodal and Cripto-1 showed opposite results for #12 and #15 rats. While the tumor tissue slides of #12 rat showed positive stains of CK5/6, E-cadherin, Lefty and β -catenin, a negative stain was observed in the tumor tissue slides of #15 rat. Interestingly, Lefty and E-cadherin co-stained in the same areas of the tumor tissue sections as well as CK5/6 and E-cadherin. In contrast, the tumor tissue slides of #15 rat showed positive stains for Nodal and Cripto-1, but those tissue sections of #12 showed negative stains for these two markers. These results indicate tumor cell conversion into non-cancerous cells in the tissue section of #12 rat, whereas the tissue sections of #15 rat remained 9L-tumor cells.

[00342] To assess genome stability of the explant cells, molecular cytogenetic analyses were performed. 20 mitotic images were collected for each explant of rat #15 and rat #12 and the average number of chromosomes for rat #15 and for rat #12 were analyzed. Table 4 shows a comparison of chromosomal aberrations of rat #15 and rat #12. Molecular cytogenetic analyses indicated 35% non-clonal chromosomal aberrations (NCCAs) and 20% of clonal chromosomal aberrations (CCAs) for rat #12 and 70% NCCAs and 15% CCAs for rat #15 (Table 4). Since the frequencies of NCCAs represent the level of genome instability while CCAs represent relative stability, these data suggest that the tumor that did not respond to QQ-SON protein treatment (rat #15) displayed higher levels of genome heterogeneity than those of the tumor (rat #12) that showed a good response to the QQ-SON treatment. This result indicates that the protein-induced *in situ* cell reprogramming significantly enhances genome stability of the treated cancer cells, indicating cell conversion of malignant cancer cells into non-cancerous cells.

Table 4: tumors that respond to the QQ-SON protein treatment (#12) display higher genome stability than a tumor that did not respond (#15); NCCA: non-clonal chromosomal aberrations; CCA: clonal chromosomal aberrations
--

Rat	#mitotic images	# chromosomes	# NCCAs	% NCCAs	# CCAs	% CCAs
#12	20	64	7	35	4	20
#15	20	60	14	70	3	15

[00343] Example 25

[00344] QQ-SON induced cell-converting cancer therapy mediated by protein-induced in situ cell reprogramming caused cancer cure of subcutaneous 9L-tumor bearing rats and late stage 4T1-bearing mice after surgical removal of primary tumors.

[00345] Subcutaneous implantation of 9L-cells into Fisher rats was performed to generate tumor-bearing rats. Five days after tumor cell implantation, daily intratumor injections of QQ-modified SON (QQ-SON) proteins were performed. A proper control of QQ-reagents in PBS (QQ-PBS) was also performed. A 90-day survival experiment was performed following a 18 daily intratumoral treatment regimen using QQ-SON proteins at 10 µg/mouse/day. During treatments, significantly reduced tumor growth in all treated rats was observed. After 18 daily treatments, 50% of the treated rats displayed diminished tumor volume over time, and no palpable tumor was present in these animals. The remaining treated rats displayed significantly slower tumor growth. The median survival of the treated group was 49±20 days (n = 8) whereas survival in the control group was 21±4 days (n = 6). Tumor recurrence was observed in three treated rats at day 32, 43 and 62, with a median recurrence of 45±13 days. The recurrent tumors grew very aggressively and reached a volume of > 12 cm³ in 5-7 days (Table 5, Treatment 1). The fourth glioma-cured rat remained tumor-free for more than 30 months without evidence of teratoma formation.

[00346] When the daily intratumor treatment regimen was expanded to 30 days, 100% glioma-cured rats within the first 73 days during a 400-day survival experiment (n = 8) were obtained. Again, tumor recurrence for three treated rats at day 73, 78, and 92 (median recurrence: 81±8 days) was observed, which was much later than the tumor recurrence observed in rats treated for only 18 days. The remaining 5 rats remained tumor-free for 400-days. The three rats with recurrent tumors were treated for an additional 30 days (daily intra-tumor injection, 10µg/day QQ-SON). Of the three rats, two had very slow progression and survived for an additional 66 days (tumor occurrence on day 78) or 68 days (tumor reoccurrence on day 92). Only one rat displayed slow tumor growth and was sacrificed at day 109 (36 days after tumor recurrence was identified) when the tumor reached 12 cm³ in volume. The median survival of the treated rats was 280±155 days (n = 8) compared to 21±4 days for the control group (n = 6) (Table 5, Treatment 2).

[00347] To ensure that this result was reproducible, the above 30 daily treatment experiment was repeated and achieved 80% glioma-cured rats within the first 79-days. Two rats displayed slow tumor growth but had reached a tumor volume of 12 cm³ at day 29 and 48, respectively (Table 5, Treatment 3). However, two glioma-cured rats displayed tumor recurrence at day 79 and 111 (median recurrence: 95±16 days). These two rats were treated with the QQ-SON proteins (10 µg/day) for an additional 30 days. One of the two rats displayed a slow tumor growth and reached a volume of 12 cm³ in 30 days after the tumor recurred on day 79. The other rat survived for an additional 69-days. The median survival of the treated rats was 276±156 days (n = 10), whereas the control rats treated with QQ-PBS only survived for 23±3 days after tumor implantation (n = 6). The tumor-cured rats from both treatments 2 and 3 survived for more than 15-months so far without tumor recurrence and without teratoma formation.

Table 5						
Group	# of Rats	# injections	9L-Glioma Cured (%)	Mean Survival (Day)	Mean Recurrence (day / Recurrence %)	Recurrence day
Control	6	18 or 30 (QQ-PBS)	0	21±4	-	-
Treatment 1	8	18 (QQ-SON)	13 (90 day survival)	49±20	45±13 / 37.5	32/42/62
Treatment 2	8	30 (QQ-SON)	70 (400 day survival)	280±155	81±8 / 30	72/78/92
Treatment 3	10	30 (QQ-SON)	60 (400 day survival)	276±156	95±16 / 20	79/111

[00348] A similar result was observed for late stage of 4T1-breast cancer bearing mice after surgical removal of the primary tumors at day 18 after 4T1-cell implantation for both QQ-PBS and QQ-SON treated mice. The daily intra-tumor QQ-SON protein treatment was performed at day 5 and continued for 40-days. Previously, it was shown that lung metastasis started at day 7. At day 18, MRI observable lung metastatic lesions for every mouse were observed. A 250-day survival experiment was performed. The resultant data, Table 6, indicated that while the QQ-PBS control mice died between days 25-47, the QQ-SON treated mice survived much longer and 61% (n=11) of QQ-SON treated mice survived entire 250-days without tumor recurrence and teratoma formation (n=18), see Figure 8. MRI results showed disappearance of lung metastatic lesions. These data demonstrate high treatment efficacy of this QQ-SON-induced cell-converting cancer therapy.

Table 6							
Mice	Group	Tumor size at surgery (mm ³)	Primary Tumor Recurrence	Day detecting lung lesion	Day of lung lesions disappeared or cause of death	Survival Day	Alive 10-23-2014)
Red 1	Control	114	No	16	Metastasis	46	
Red 2	Control	125	No	16	Metastasis	46	
Red 3	Control	60.5	No	16	Metastasis	44	
Red 4	Control	25.5	No	16	Metastasis	44	
Blue 1	Treatment	6.5	No	16	Inflammation	47	
Blue 2	Treatment	20.5	No	16	Stasis		110
Blue 3	Treatment	3.5	No	16	74		267
Blue 4	Treatment	60.5	No	16	Metastasis	49	
Blue 5	Treatment	33.5	No	16	59		267
Green 1	Treatment	20	No	16	60		267
Green 2	Treatment	18.5	No	16	Inflammation	75	
Green 3	Treatment	15.5	No	16	60		267
Green 4	Treatment	54	No	16	Metastasis position	45	
Green 5	Treatment	53	No	16	Inflammation	54	

[00349] Figure 8 is a Kaplan-Meier survival curve (250-day survival) of 4T1-breast cancer bearing mice after surgical removal of the primary 4T1-breast cancer at day 18. These 4T1 breast cancer bearing mice were treated with QQ-PBS (Control) and QQ-SON proteins (Treatment) at day 6 after 4T1-cell implantation into the #4 fat pad of female BALB/c mice. The primary tumors were palpable around day 5 and surgically removed at day 18. The QQ-PBS/ QQ-SON treatment was performed daily by both intra-tumoral (5μg/mouse/day) and intraperitoneal injections (25μg/mouse/day).

[00350] Day 0 in the Kaplan-Meier survival curve of Figure 8 is the day when the surgery was performed at day 18. A small population of mice displayed tumor recurrence due to incomplete tumor removal. These mice were treated with QQ-SON protein or QQ-PBS via intra-tumoral injections and their survival were also compared (dotted survival curves). For those mice without tumor recurrence, their survival curves are shown in solid lines. Mice that survived for this entire 250-day survival experiment without tumor recurrence are considered tumor-cured mice. Tumor-cure has been achieved in 61% of population of treated mice.

[00351] Thus, QQ-SON-induced in situ cell reprogramming of the cancer cells inside the tumor to generate transient piPSCs that differentiate into different non-cancerous cells within that tissue is demonstrated. The types of the differentiated non-cancerous cells depend on specific tissue environment. This inter-play between QQ-SON-induced in situ cell reprogramming and tissue environment induced differentiation precisely regulates generation of transient piPSCs inside tumor and induced differentiation, preventing tumor and teratoma formation. This safe protein-induced in situ cell reprogramming technology in situ generates stem cells and which then differentiate into normal cells induced by tissue environment to replace diseased cells for treatment of many diseases and injuries.

[00352] Sequences

[00353] Homo sapiens SRY (sex determining region Y)-box 2 (SOX2) (NM_003106) SEQ ID NO:1 (protein) and SEQ ID NO:65 (DNA)

MYNMMETELKPPGPQQTSGGGGGNSTAAAAGGNQKNSPDRVKRPMNAFMVWSRGQRRKMAQENPK
MHNSEISKRLGAEWKLLSETEKRPFIDEAKRLRALHMKHEPDYKYRPRRKTCTLMKKDKYTLPGG
LLAPGGNSMASGVGVGAGLGAGVNQRMDSYAHMNGWSNGSYSMMQDQLGYPQHPGLNAHGAAQMQ
PMHRYDVSALQYNSMTSSQTYMNGSPTYSMSYSQQGTPGMALGSMGSSVVKSEASSSPVVTSSSH
SRAPCQAGDLRDMISMYLPGAIEVPEPAAPSRLHMSQHYQSGPVPGTAINGTLPPLSHM (SEQ ID
NO:1)

atgtacaacatgatggagacggagactgaagccgcgggcccgcagcaaaacttcggggggcgggcg
cggaactccaccgcggcgggcgggcggaaccagaaaaacagcccgaccgcgtcaagcggc
ccatgaatgccttcattggtgtggtcccgcgggcagcggcgaagatggcccaggagaacccaag
atgcacaactcggagatcagcaagcgccctgggcgcgcgagtggaacttttgcggagacggagaa
gcgggcgttcacgcagaggctaagcggctgcgagcgctgcacatgaaggagcaccgcgattata
aataccggcccccgcggaacaaagacgctcatgaagaaggataagtacacgctgcccggcggg
ctgctggcccccgcggaatagcatggcgagcggggtcggggtgggcgcgcgacctgggcgcggg
cgtgaaccagcgcattgacaggttacgcgcacatgaacggctggagcaacggcagctacagcatga
tgcaggaccagctgggctacccgcagcaccgggacctcaatgcgcacggcgagcgcagatgcag
cccatgcaccgctacgacgtgagcgccctgcagtacaactccatgaccagctcgcagacctacat
gaacggctcgccacctacagcatgtcctactcgcagcagggcaccctggcatggctcttggct
ccatgggttcgggtgggtcaagtcaggagccagctccagccccctgtggttacctcttccctccac
tccagggcgccctgccaggccggggacctccgggacatgatcagcatgtatctccccggcgccga
ggtgccggaacccgcgcggccccagcagacttcacatgtcccagcactaccagagcggccccggtgc
ccggcacggccattaacggcacactgcccctctcacacatgtga (SEQ ID NO:65)

[00354] Homo sapiens POU class 5 homeobox 1 (POU5F1) also known as Oct4 (NM_002701) SEQ ID NO:2 (protein) and SEQ ID NO:66 (DNA)

MAGHLASDFAFSPPPGGGDGPGGPEPGWVDPRTWLSFQGPPGGPGIGPGVGPSEVWGIPPCPP
PYEFCGGMAYCGPQVGVLVPQGGLTSQPEGEAGVGVESNSDGASPEPCTVTPGAVKLEKEKLE
5 QNPEESQDIKALQKELEQFAKLLKQKRITLGYTQADVGLTLGLVLFGKVFSQTTICRFEALQLSFK
NMCKLRPLLQKWVEEADNNENLQEICKAETLVQARKRKRTSIENRVRGNLENLFLQCPKPTLQQI
SHIAQQLGLEKDVVRVWFCNRRQKGRSSSDYAQREDFEAAGSPFSGGPVSFPLAPGPHFGTPGY
GSPHFTALYSSVPFPEGEAFPPVSVTTLGSPMHSN (SEQ ID NO:2)

atggcgggacacctggccttcgcatcttcgccttctcgccccctccaggtgggtggaggtgatgggcc
10 aggggggcccggagccgggctgggttgatcctcggacctggctaagcttccaaggccctcctggag
ggccaggaatcgggcccgggggttgggcccaggtctcgaggtgtgggggattcccccatgcccccg
ccgtatgagttctgtggggggatggcgactgtgggccccaggttgagtggtgtagtgcccca
aggcggcttgagacctctcagcctgagggcggaagcaggagtcgggggtggagagcaactccgatg
gggcctccccggagccctgcaccgtcaccctgggtgccgtgaagctggagaaggagaagctggag
15 caaaaccggaggagtcaccagcatcaaagctctgcagaaagaactcgagcaatttgccaagct
cctgaagcagaagaggatcaccctgggatatacacaggccgatgtggggctcaccctgggggttc
tatttgggaaggtattcagccaaacgaccatctgccgctttgaggctctgcagcttagcttcaag
aacatgtgtaagctgcccgccttgctgcagaagtgggtggaggaagctgacaacaatgaaaatct
tcaggagatatgcaaagcagaaaccctcgtgcaggcccgaaagagaaagcgaaccagtatcgaga
20 accgagtgagaggcaacctggagaatttgttcctgcagtgcgccgaaaccacactgcagcagatc
agccacatcgcccagcagcttgggctcgagaaggatgtgggtccgagtggttctgtaaccggcg
ccagaagggcaagcgatcaagcagcgactatgcacaacgagaggattttgaggctgctgggtctc
ctttctcagggggaccagtgctcttctcctctggccccagggccccattttggtaccccaggctat
gggagccctcacttcactgcactgtactcctcgggtccctttccctgagggggaagcctttcccc
25 tgtctccgtcaccactctgggctctcccatgcattcaaactga (SEQ ID NO:66)

[00355] Homo sapiens Nanog homeobox (NANOG) (NM_024865) SEQ ID NO:3 (protein) and SEQ ID NO:67 (DNA)

MSVDPACPQSLPCFEASDCKESSMPVICGPEENYPQLQSSAEMPHTETVSPLPSSMDLLIQDS
PDSSTSPKKGQPTSAEKSVAKKEDKVPVKKQKTRTVFSSTQLCVLNDRFQRQKYLSQLQMQEELSN
30 IILNSYKQVKTWFQNRQMKSKRWQKNNWPKNNSNGVTQKASAPTYPSLYSSYHQGLVNPTGNLPM
WSNQTNWNSTWSNQTNQIWSNSHWNQTQWCTQSWNNQAWNSPFYNCGEESLQSCMQFQPNSPA
SDLEAALEAAGEGLNVIQQTTRYFSTPQTMDLFLNYSNMNPEDV (SEQ ID NO:3)

atgagtggtggatccagcttgccttgccttgccttgaagcatccgactgtaaagaatc
ttcacctatgcctgtgatttgtgggctgaagaaaactatccatccttgcaaatgtcttctgctg
35 agatgcctcacacggagactgtctctcctcttccctcctccatggatctgcttattcaggacagc
cctgattcttccaccagtcctcaaaggcaacaacccacttctgcagagaagagtgtcgcaaaaaa
ggaagacaaggtcccgggtcaagaacagaaagaccagaactgtgttctcttccaccagctgtgtg
tactcaatgatagatttcagagacagaaatacctcagcctccagcagatgcaagaactctccaac
atcctgaacctcagctacaaacaggtgaagacctggttcagaaccagagaatgaaatctaagag
40 gtggcagaaaaaacaactggccgaagaatagcaatggtgtgacgcagaaggcctcagcacctacct
acccagcctttactcttctaccaccagggatgcctggtgaaccgactgggaaccttccaatg
tggagcaaccagacctggaacaattcaacctggagcaaccagaccagaacatccagtcctggag
caaccactcctggaacactcagacctggtgcacccaatcctggaacaatcaggcctggaacagtc
ccttctataactgtggagaggaatctctgcagtcctgcagtcagttccagccaaattctcctgcc
45 agtgacttggaggctgccttgggaagctgctggggaaggccttaatgtaatacagcagaccactag
gtatttttagtactccacaaaccatggatttattcctaaactactccatgaacatgcaacctgaag
acgtgtga (SEQ ID NO:67)

[00356] Homo sapiens lin-28 homolog A (C. elegans) (LIN28A) (NM_024674) SEQ ID NO:4 (protein) and SEQ ID NO:68 (DNA)

MGSVSNQQFAGGCAKAAEEAPEEAPEDAARAADPEQLLHGAGICKWFNVRMGFGLSMTARAGVALDPPVDVVFVHQSKLHMEGFRSLKEGEAVEFTFKKSAKGLESIRVTGPGGVFCIGSERRPKGKSMQKRRSKGDRCYNCGGLDHHAKECKLPPQPKKCHFCQSISHMVASCPKLAQQGPSAQGKPTYFREEE E E I H S P T L L P E A Q N (SEQ ID NO:4)

5 atgggctccgtgtccaaccagcagtttgcaggtggctgcgccaaggcggcagaagaggcgcccgaggaggcgccggaggacgcggcccgggcgcgagcctcagctgctgcacgggtgcgggcatctgtaagtgggtcaacgtgcgcagtggggttcggcttcctgtccatgaccgcccgcgggggtcgcgctcgacccccagtggtatgtctttgtgcaccagagtaagctgcacatggaagggttcgggagcttgaaggagggtgaggcagtggttgcaccttaagaagtcagccaagggtctggaatccatccgtgtcaccggacctggtggagtattctgtattgggagtgagaggcgccaaaaggaaagagcatgcagaagcgcagatcaaaaggagacaggtgctacaactgtggaggtctagatcatcatgccaaaggaatgcaagctgccacccccagcccaagaagtgccacttctgccagagcatcagccatatggtagcctcatgtccgctgaaggcccagcagggccctagtgcacagggaaagccaacctactttcgagaggaagaa

15 **[00357]** Homo sapiens Krueppel-like factor 4 (Klf4) (NP_004226) SEQ ID NO:5 (protein) and SEQ ID NO:69 (DNA)

MRQPPGESDMAVSDALLPSFSTFASGPAGREKTLRQAGAPNNRWREELSHMKRLPPVLPGRPYDLAAATVATDLES GGAGAACGGSNLAPLPRRETEEFNDLLDLDLDFILSNLTHPPESVAATVSSSASASSSSSPSSSGPASAPSTCSFTYPIRAGNDPGVAPGGTGGGLLYGRESAPPPTAPFNLADINDVSPSGGFVAELLRPELDPVYIPPQQPQPPGGGLMGKFVLKASLSAPGSEYGPSVISVSKGSPDGSHPVVVAPYNGGPPRTCPKIKQEAVSSCTHLGAGPPLSNGHRPAAHDFPLGRQLPSRTTPTLGLEEVLSSRDCHPALPLPPGFHPHPGPNYPSFLPDQMOPQVPPLHYQELMPPGSCMPEEPKPKRGRRSWPRKRTATHTCDYAGCGKTYTKSSHLKAHLRHTGKPYHCDWDGCGWKFARSDDELTRHYRKHTGHRPFQCQKCDRAFSRSDHLALHMKRHF (SEQ ID NO:5)

25 atgaggcagccacctggcgagtctgacatggctgtcagcgacgcgctgctcccatctttctccacgttcgcgctctggccccggcggaaggaggagaagacactgcgtcaagcaggtgccccgaataaccgctggcgaggagctctcccatgaagcgacttccccagtgcttcccggccgcccctatgacctggcgggcgacacctggccacagacctggagagcgggcgagccggtgcggcttgcgggcggtagcaacctggcgccccctacctcgagagagaccgaggagttcaacgatctcctggacctggactttattctctccaattcgctgacctatcctccggagtcagtgcccgccaccgtgtcctcgtcagcgtcagcc

30 tctcttcgctcgctcgccgctgagcagcgccctgccagcgcgccctccacctgcagcttcacctatccgatccggggcggaacgacccggcggtggcgccggggcgacggggcgaggcctcctctatggcaggagtcgctccccctccgacggctcccttcaacctggcgagcatcaacgacgtgagcccc

35 tcggggcggttcgctggcgagctcctgcggccagaattggacccggtgtacattccgcccagcagccgcagcgccaggtggcgggctgatgggcaagttcgtgctgaaggcgctcgctgagcgccctggcagcgagtacggcagcccgctcggtcatcagcgctcagcaaggcagccctgacggcagccaccgggtggtggtggcgccctacaacggcgggcgccgcgcacgtgccccaaagatcaagcaggaggcggtctcttcgctgacccacttggcgctggacccccctctcagcaatggccaccggccggctgcacacgacttccccctggggcggcagctccccagcaggactaccccgaccctgggtcttgaggaagtgtgagcagcagggactgtcaccctgcctgccgttctcctccggcttccatccccaccggggcccaa

40 ttaccatccttctgcccgatcagatgcagccgcaagtcgccgctccattaccaagagctcatgccaccgggttctgcatgccagaggagcccaagccaaagaggggaagacgatcggtggccccgaaaaggacggccaccacacttgtgattacggggctgcggcaaacctacacaaagagttcccatctcaaggcacacctgcgaacccacacaggtgagaaaccttaccactgtgactgggacggctgtgatggaaattcgccccgctcagatgaactgaccaggcactaccgtaaacacacggggcaccgcccgtccagtgccaaaaatgcgaccgagcattttccaggtcggaccacctgccttacacatgaagag

45 gcatttttaa (SEQ ID NO:69)

[00358] Homo sapiens c-MYC (NP_002458) SEQ ID NO:6 (protein) and SEQ ID NO:70 (DNA)

10

15

20

25

35

50

[illegible]

ccaggcagcaagaatacagatgccatcagtgctctgaggatgtcgacctgcttttgaatcaaaggat
 aaataactcccagtcggctcagtcattggctaccccagtggtttccgtagcaactcctactttac
 caggacaaggaatgggaggatatccatcagccatttcaacaacatatggtaccgagtagtctctg
 agtagtgacagacctgtcatctctgtctgggtttaacaccgccagcgctcttcaccttggttcagt
 5 aactggctggcaacagcaacacctacataacatgccaccatctgccctcagtcagttgggagctt
 gcactagcactcatttatctcagagttcaaattctctccctgccttctactcaaagcctcaacatc
 aagtcagaacctgtttctcctcctagagaccgtaccaccaccccttcgagatacccacaacacac
 gcgccacgagggcgggagatctcctggtgacagcttgagcagctgtagcagttcgtacgacggga
 gcgaccgagaggatcaccggaacgaattccactccccattggactcaccagaccttcgccggac
 10 gaaagggaaagtcctcagtcgaagcgcagtcgactttctgaaggatgggcaacatga (SEQ ID
 NO: 73)

[00362] Human transcription factor TBX5 (U80987) SEQ ID NO:10 (protein) and SEQ ID
 NO:74 (DNA)

MADADEGFLAHTPLEPDAKDLPCDSKPESALGAPSKSPSSPQAAFTQQGMEGIKVFLHERELWL
 15 KFHEVGTEMIITKAGRRMFPSYKVKVTGLNPCKTYILLMDIVPADDHRYKFADNKWSVTGKAEP
 MPGRLYVHPDSPATGAHWMRQLVSFQKLKLTNNHLDPFHGHIILNSMHKYQPRLHIVKADENNGFG
 SKNTAFCTHVFPETAFAVTSYQNHKITQLKIENNPFAGKFRGSDDMELHRMSRMQSKEYPVVPR
 STVRQKVASNHSPFSSESRLSTSSNLGSQYQCENGVSGPSQDLLPPPNPYPLPQEHQSQIYHCTK
 RKGECDHPWSICFLSYLFLSLGWG (SEQ ID NO:10)

atggccgacgcagacgagggctttggcctggcgacacgcctctggagcctgacgcaaaagacct
 gccctgcgattcgaaacccgagagcgcgctcggggcccccagcaagtccccgtcgtccccgcagg
 ccgccttcacccagcagggcatggagggaatcaaagtgtttctccatgaaagagaactgtggcta
 aaattccacgaagtgggcacggaaatgatcataaccaaggctggaaggcggatgtttcccagtta
 caaagtgaaggtgacgggccttaatcccaaacgaagtacattcttctcatggacattgtacctg
 25 ccgacgatcacagatacaaatcgcagataataaatggtctgtgacgggcaaagctgagcccgcc
 atgcctggccgctgtacgtgcacccagactccccgccaccggggcgcatgttgatgaggcagct
 cgtctccttccagaaactcaagctcaccaacaaccacctggacccatttgggcatattattctaa
 attccatgcacaaataaccagcctagattacacatcgtgaaagcggatgaaaataatggatttggc
 tcaaaaaatacagcgttctgcactcacgtctttcctgagactgcgtttatagcagtgacttccta
 30 ccagaaccacaagatcacgcaattaaagattgagaataatccctttgcaaaggatttcggggca
 gtgatgacatggagctgcacagaatgtcaagaatgcaaagtaaagaatatcccggtggtccccagg
 agcaccgtgaggcaaaaagtggcctccaaccacagtcctttcagcagcgagctctcgagctctctc
 cacctcatccaatttgggggtcccaataaccagtgtgagaatggtgtttccggcccccctcccaggacc
 tcctgcctccaccaacccatacccactgccccaggagcatagccaaatttaccattgtaccaag
 35 aggaaaggtgagtgtgatcacccctggtcaatttgccttcttcttaccttttcccttcccttggg
 ttgggggtga (SEQ ID NO:74)

[00363] Homo sapiens neurogenin 3 (NEUROG3) (NM_020999.3) SEQ ID NO:11 (protein)
 and SEQ ID NO:75 (DNA)

MTPQPSGAPTVQVTRETERSFPRASEDEVTCPTSAPPSPTRTRGNCAEAEEGGCRGAPRKLRRR
 40 GGRSRPKSELALSKQRRSRKKANDRERNRMHNLNSALDALRGVLPFTPDDAKLTKIETLRF
 YIWALTQTLRIADHSLYALEPPAPHCHELSPGGSPGDWGSLSYSPVSQAGSLSPAASLEERPGLL
 GATFSACLSPGSLAFSDFL (SEQ ID NO:11)

atgacgcctcaaccctcgggtgcgcccactgtccaagtgacccgtgagacggagcgggtccttccc
 cagagcctcgggaagacgaagtgcctgccccacgtccgccccgcccagccccactcgcacacggg
 45 ggaactgcgcagaggcgaagaggaggctgccgagggggccccgaggaagctccggggcacggcgc
 gggggacgcagccggcctaagagcaggttgccactgagcaagcagcgacggagtcggcgaaagaa
 ggccaacgaccgcgagcgcaatcgaatgcacaacctcaactcggcactggacgccttgcgcggtg
 tcctgcccaccttcccagacgacgcgaagctcaccaagatcgagacgctgcgcttcgcccaaac
 tacatctgggcgctgactcaaacgctgcgcatacgcgaccacagcttgtacgcgctggagccgcc
 50 ggcgccgcactgcggggagctgggcagcccaggcggttcccccggggactgggggtccctctact

ccccagtctcccaggctggcagcctgagtcgccgcgcgtcgctggaggagcgacccgggctgctg
ggggccaccttttccgcctgcttgagcccaggcagtcctggctttctcagattttctgtga (SEQ
ID NO:75)

[00364] Homo sapiens paired box 4 (PAX4) (NM_006193.2) SEQ ID NO:12 (protein) and
SEQ ID NO:76 (DNA)

MNQLGGLFVNGRPLPLDTRQQIVRLAVSGMRPCDISRILKVSNGCVSKILGRYYRTGVLEPKGIG
GSKPRLATPPVVARIAQLKGECPALFAWEIQRLCAEGLCTQDKTPSVSSINRVLRALQEDQGLP
CTRLRSPAVLAPAVLTPHSGSETPRGTHPGTGHRNRTIFSPSQAEALEKEFQRGQYPDSVARGKL
ATATSLPEDTVRVWFSNRRAKWRRQEKWKWEMQLPGASQGLTVPRVAPGIIISAQQSPGSPVPTAAL
PALEPLGPSCYQLCWATAPERCLSDTPPKACLKPCWGHLPQPNSLDSGLLCLPCSSHCHLASL
SGSQALLWPGCPLLYGLE (SEQ ID NO:12)

atgaaccagcttggggggctctttgtgaatggccggccctgcctctggatacccggcagcagat
tgtgcggttagcagtcagtggaatgcggccctgtgacatctcacggatccttaaggtatctaag
gctgtgtgagcaagatcctagggcggtactaccgcacaggtgtcttgagccaaagggcattggg
ggaagcaagccacggctggctacacccctgtggtggctcgaattgccagctgaagggtagtg
tccagccctctttgcctgggaaatccaacgccagctttgtgctgaagggctttgcacccaggaca
agactcccagtgctcctccatcaaccgagtcctgcgggcattacaggaggaccagggaactaccg
tgcacacggctcaggtcaccagctgttttggtccagctgtcctcactccccatagtggtctga
gactccccggggtacccacccagggaaccggccaccggaatcgactatcttctcccaagccaag
cagaggcactggagaaagagttccagcgtgggcagtatcctgattcagtgggccgtggaaagctg
gctactgccacctctctgcctgaggacacgggtgagggctctgggttttccaacagaagagccaaatg
gcgtcggcaagagaagctcaagtgggaaatgcagctgccaggtgcttcccaggggctgactgtac
caaggggtgccccaggaatcatctctgcacagcagtcacctggcagtggtgccacagcagccctg
cctgccttgaaccactgggtccctcctgctatcagctgtgctgggcaacagcaccagaaaggtg
tctgagtgacacccacctaagcctgtctcaagccctgctggggccacttgccccacagccga
attccttggactcaggactgctttgccttcccttgcccttccctccactgtcacttgccagtcct
agtggctctcaggccctgctctggcctggctgccactactgtatggcttggaatga (SEQ ID
NO:76)

[00365] Homo sapiens pancreatic and duodenal homeobox 1 (PDX1) (NM_000209.3) SEQ ID
NO:13 (protein) and SEQ ID NO:77 (DNA)

MNGEEQYYAATQLYKDPCAFQRGPAPEFSASPPACLYMGRQPPPPPPHFPFGALGALEQGSPPDI
SPYEVPLADDDPAVAHLHHHLPAQLALPHPPAGPFPEGAEPGVLEENRVLQLPFPWMKSTKAHAW
KGQWAGGAYAAEPEENKRTRTAYTRAQLELEKEFLFNKYISRPRRVELAVMLNLTERHIKIWFQ
NRRMKWKKEEDKKRGGGTAVGGGGVAEPEQDCAVTSGEELLALPPPPPPGGAVPPAAPVAAREGR
LPPGLSASPQSSVAPRRPQEPR (SEQ ID NO:13)

atgaacggcgaggagcagtcactacgcggccacgcagctttacaaggacccatgcgcgttccagcg
aggcccgcgccggagttcagcgccagccccctgcgtgcctgtacatgggcccagccccgc
cgccgcgcgcgacccggtccctggcgccctgggcgcgtggagcagggcagcccccgacatc
tccccgtacgaggtgccccccctcgccgacgaccccgcggtggcgaccttcaccaccacctcc
ggctcagctcgcgctccccacccgcccgggccccttcccggaggagccgagccggcgctcc
tggaggagcccaaccgcgtccagctgcctttcccatggatgaagtctaccaaagctcacgcgtgg
aaaggccagtgggcaggcgcgctacgctgaggagccggaggagaacaagcggacgcgcacggc
ctacacgcgcgcacagctgctagagctggagaaggagttcctattcaacaagtacatctcacggc
cgcgccgggtggagctggctgtcatgttgaaacttgaccgagagacacatcaagatctgggttccaa
aaccgcccagatgaagtggaaaaaggaggaggacaagaagcgcgcgcgaggagagctgtcggggg
tggcggggtcgcgagcctgagcaggactgcgcgctgacctccggcgaggagcttctggcgctgc
cgccgcgcgcgcccccgagggtgctgtgccgcccgtgccccggttgcgcggcgagaggggccgc
ctgccgcttggccttagcgcgctgccacagccctccagcgtcgcgccctcggcggccgcaggaacc
acgatga (SEQ ID NO:77)

[00366] Human SOX9 protein (1-509) (Z46629.1) SEQ ID NO:14 (protein) and SEQ ID NO:78 (DNA)

MNLLDPFMKMTDEQEKGSLGAPSPMTSEDSAGSPCPSGSGSDTENTRPQENTFPKGEPDLKKESE
EDKFPVCIREAVSQVLKGYDWTLVPMPVRVNGSSKNKPHVKRPMNAFMVWAQAARRKLADQYPHL
HNAELSKTLGKLWRLLENESEKRPFVEEAERLRVQHKKDHPDYKYQPRRRKSVKNGQAEAEAEATEQ
THISPNAIFKALQADSPHSSSGMSEVHSPGEHSGQSQGPPTPPTTPKTDVQPGKADLKREGRPLP
EGGRQPPIDFRDVDIGELSSDVISNIETFDVNEFDQYLPPNGHPGVPATHGQVITYTGSYGISSTA
ATPASAGHVWMSKQQAPPPPPQPPQAPPAPQAPPQPAAPPQQAAPPQQAHTLTTLSEPG
QSQRTHIKTEQLSPSHYSEQQHSPPQIAYSFPNLPHYSPSYPPITRSQYDYTDHQNSSSYSHA
AGQGTGLYSTFTYMNPAQRPMYTPADTSGVPSIPQTHSPQHWEQPVYTQLTRP (SEQ ID
NO:14)

atgaatctcctggaccccttcatgaagatgaccgacgagcaggagaagggcctgtccggcgcccc
cagccccaccatgtccgaggactccgcgggctcgccctgcccgtcgggctccggctcggacaccg
agaacacgcggccccaggagaacacgttccccaaaggcgagcccgatctgaagaaggagagcgag
gaggacaagttccccgtgtgcatccgcgaggcggtcagccaggtgctcaaaggctacgactggac
gctggtgccccatgccggtgcgcgtcaacggctccagcaagaacaagccgcacgtcaagcggccca
tgaacgccttcatggtgtggcgcgaggcgcgcgaggaagctcgcggaaccagtaccgcacttg
cacaacgcgcgagctcagcaagacgctgggcaagctctggagacttctgaacgagagcgagaagcg
gcccttcgtggaggaggcgagcggtgctgcgcgtgcagcacaagaaggaccaccggattacaagt
accagccgcggcgagggaagtcggtgaagaacgggcaggcgaggcagaggaggccacggagcag
acgcacatctcccccaacgccatcttcaaggcgctgcaggccgactcgccacactcctcctccgg
catgagcgaggtgactcccccgcgagcactcggggcaatcccaggggccaccgacccccacca
ccacccccaaaaccgacgtgcagccgggcaaggctgacctgaagcgagaggggcgcccccttgcca
gaggggggagagacagccccctatcgacttccgcgacgtggacatcggcgagctgagcagcgacgt
catctccaacatcgagaccttcgatgtcaacgagtttgaccagtacctgcccgaacggccacc
cggggggtgcccggccacgcacggccaggtcacctacacgggcagctacggcatcagcagcaccgcg
gccacccccggcgagcgcgggccacgtgtgatgtccaagcagcaggcgccgcccacccccgca
gcagccccccacaggcccccgccggcccccgaggcgcccccgagccgcaggcgccgccccacagc
agccggcgggcaccccccgcagcagccacaggcgccacacgctgaccacgctgagcagcgagccgggc
cagtcccagcgaacgcacatcaagacggagcagctgagccccagccactacagcgagcagcagca
gcactcgcccccaacagatcgctacagcccccttaacctcccacactacagccccctcctaccgc
ccatcaccgcgtcacagtagactacaccgaccaccagaactccagctcctactacagccacgcg
gcaggccaggggcaccggccttactccaccttcacctacatgaaccccgctcagcgccccatgta
cacccccatcgccgacaccttgggggtcccttccatcccgcagacccacagccccagcactggg
aacaacccgcttacacacagctcactcgaccttga (SEQ ID NO:78)

[00367] Homo sapiens zinc finger protein SLUG (SLUG) gene (AF084243.1) SEQ ID NO:15 (protein) and SEQ ID NO:79 (DNA)

MPRSFLVKKHFNASKKPNYSELDTHTVIISPYLYESYSMPVIPQPEILSSGAYSPITVWTTAAPF
HAQLPNGLSPLSGYSSSLGRVSPPPPSDTSSKDHSGSESPISDEEERLQSKLSDPHAIEAEKFQC
NLCNKTYSTFSGLAKHKQLHCDAQSRKSFSCYCDKEYVSLGALKMHIRTHTLPCVCKICGKAFFS
RPWLLQGHIRTHTGEKPFSCPHCNRAFAADRSLRAHLQTHSDVKKYQCKNCSKTFSRMSLLHKHE
ESGCCVAH (SEQ ID NO:15)

atgccgcgctccttccctgggtcaagaagcatttcaacgcctccaaaaagccaaactacagcgaact
ggacacacatacagtgattatttccccgtatctctatgagagttactccatgctgtcataccac
aaccagagatcctcagctcaggagcatacagccccatcactgtgtggactaccgctgctccattc
cacgcccagctacccaatggcctctctcctctttccggatactcctcatctttggggcgagtgag
tccccctcctccatctgacacctcctccaaggaccacagtggctcagaaagccccattagtgatg
aagaggaaagactacagttcaagctttcagacccccatgccattgaagctgaaaagttttagtg
aatttatgcaataagacctattcaactttttctgggctggccaaacataagcagctgcactgcga
tgcccagttctagaaaatctttcagctgtaaatactgtgacaaggaatatgtgagcctgggcccc

tgaagatgcatatttcggacccacacattaccttggtgtttgcaagatctgcggcaaggcgttttcc
agaccctgggttgcttcaaggacacattagaactcacacgggggagaagcctttttcttgccctca
ctgcaacagagcatttgcagacaggtcaaactctgagggctcatctgcagacccattctgatgtaa
agaaataccagtgcaaaaactgctccaaaaccttctccagaatgtctctcctgcacaaacatgag
5 gaatctggctgctgtgtagcacactga (SEQ ID NO:79)

[00368] Homo sapiens v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog A (MafA) (NM_201589.3) SEQ ID NO:16 (protein) and SEQ ID NO:80 (DNA)

MAAELAMGAELPSSPLAIEYVNDLMDKFEVKKEPPEAERFCHRLPPGSLSSSTPLSTPCSSVPSS
PSFCAPSPGTGGGGGAGGGGSSQAGGAPGPPSGGPGAVGGTSGKPALEDLYWMSGYQHHLNPEA
10 LNLTPEDAVEALIGSGHHGAHHGAHHHPAAAAAYEAFRGPFGAGGGGADDMGAGHHHGAHHAAHHH
HAAHHHHHHHHHGGAGHGAGGAGHHVRLEERFSDDQLVSMVRELNRQLRGFSKEEVIRLKQKRR
TLKNRGYAQSCRFRVQQRHILESEKCQLQSQVEQLKLEVGRRLAKERDLYKEYEKLGRGGPGS
AGGAGFPREPSPPQAGPGGAKGTADFFL (SEQ ID NO:16)

atggccgcggagctggcgatggcgccgagctgccagcagcccgctggccatcgagtacgtcaa
15 cgacttcgacctgatgaagttcgaggtgaagaaggagcctcccagggccgagcgcttctgccacc
gcctgccgccaggtcgcgtgtcctcgacgcgcgtcagcacgccttgcctcctgcctcctcgc
cccagcttctgcgcgccagcccgggcaccggcgggcgggcgggcgggcgggcgggcgggcggtc
gtctcagggcgggggcgccccggggcgccgagcgggggccccggcgccgtcgggggcacctcgg
ggaagccggcgctggaggatctgtactggatgagcggtaccagcatcacctcaaccccgaggcg
20 ctcaacctgacgcccagggagcggtggaggcgctcatcggcagcgggccaccacggcgcgacca
cggcgcgcaccacccggcgggcgccgcgcagcctacgaggctttccgcgggccccgggttcgcggggc
gcggcgagcgggacgacatgggcgcggccaccaccacggcgcgccaccacgcggccccaccatcac
cacgcgcggccaccaccaccaccaccaccaccaccaccatggcgggcggggacacggcggtggcg
gggcccaccacgtgcgcctggaggagcgcttctccgacgaccagctggtgtccatgtcggtgcg
25 agctgaaccggcagctccgcggcttcagcaaggaggaggtcatccggctcaagcagaagcggcgc
acgtcaagaaccgcggctacgcgcagtcctgccgcttcaagcggtgcagcagcggcacattct
ggagagcgagaagtgcgaactccagagccaggtggagcagctgaagctggagggtggggcgccctgg
ccaaagagcgggacctgtacaaggagaaatacagaagctggcgggcgggggcgggccccgggagc
gcgggcgggggcggtttcccgcgggagccttcgcgcgcgcagggccggtcccggcgggggccaagg
30 cacggccgacttcttcctgtag (SEQ ID NO:80)

[00369] Homo sapiens neuronal differentiation 1 (NEUROD1)(NM_002500.4) SEQ ID NO:17 (protein) and SEQ ID NO:81 (DNA)

MTKSYSESLMGEPQPQGPPSWTDECLSSQDEEHEADKKEDDLEAMNAEEDSLRNGGEEDEDED
LEEEEEEEEEEDDDQKPKRRGPKKKKMTKARLERFKLRMKANARERNRMHGLNAALDNLKRVPC
35 YSKTQKLSKIETLRLAKNYIWALSEILRSGKSPDLVSFVQTLCKGLSQPTTNLVAGCLQLNPRTF
LPEQNQDMPPHLPTASASFVHPYSYQSPGLSPPYGTMDSSHVFHVKPPPHAYSAALEPFFESP
LTDCTSPSFDGPLSPPLSINGNFSFKHEPSAEFEKNYAFTMHYPAATLAGAQSHGSIFSGTAAPR
CEIPIDNIMSFDSHSHHERVMSAQLNAIFHD (SEQ ID NO:17)

atgaccaaatacgtacagcgagagtgggctgatgggcgagcctcagccccaaggtcctccaagctg
40 gacagacgagtgtctcagttctcaggacgaggagcacgaggcagacaagaaggaggacgacctcg
aagccatgaacgcagaggaggactcactgaggaaacgggggagaggaggaggacgaagatgaggac
ctggaagaggaggaagaagaggaagaggaggatgacgatcaaaagcccaagagacgcggccccaa
aaagaagaagatgactaaggctcgccctggagcggttttaaatgagacgcatgaaggctaacgccc
gggagcggaaccgcatgcacggactgaacgcggcgctagacaacctgcgcaagggtggtgccttg
45 tattctaagacgcagaagctgtccaaaatcgagactctgcgcttgccaagaactacatctgggc
tctgtcggagatcctgcgctcaggcaaaagcccagacctggtctccttcgttcagacgctttgca
agggcttatcccaaccaccaccaacctggttgcgggctgcctgcaactcaatcctcggactttt
ctgcctgagcagaaccaggacatccccccacctgccgacggccagcgcttcttccctgtaca
cccctactcctaccagtcgcctgggctgcccggtccgccttacggtaccatggacagctcccatg
50 tcttccacgttaagcctccgcccgcacgcctacagcgagcgctggagcccttctttgaaagcct

ctgactgattgcaccagcccttcctttgatggaccctcagcccgccgctcagcatcaatggcaa
 cttctcttttcaaacacgaaccgtccgcccagtttgagaaaaattatgcctttaccatgcactatc
 ctgcagcgacactggcaggggcccagccacggatcaatcttctcaggcaccgctgccctcgc
 tgcgagatccccatagacaatattatgtccttcgatagccattcacatcatgagcgagtcagag
 5 tgcccagctcaatgccatatttcatgattag (SEQ ID NO:81)

[00370] Homo sapiens myogenic factor 5 (MYF5) (NM_005593.2) SEQ ID NO:18 (protein) and SEQ ID NO:82 (DNA)

MDVMDGCQFSPSEYFYDGSCIPSPGEGFGEFVPRVAAFAGHKAELQGSDEDEHVRAPTGHHQAG
 HCLMWACKACKRKSTTMDRRKAATMRERRRLKKVNQAFETLKRCTTTNPNQRLPKVEILRNAIRY
 10 IESLQELLREQVENYYSLPGQSCSEPTSPTSNCS DGMPECNSPVWSRKSSTFDSIYCPDVS NVYA
 TDKNSLSSLDCLSNIVDRITSSEQPGLPLQDLASLSPVASTDSQPATPGASSSRLIYHVL (SEQ
 ID NO:18)

atggacgtgatggatggctgccagttctcaccttctgagtacttctacgacggctcctgcatacc
 gtcccccgaggggtgaatttggggacgagtttgtgccgagtggtgccttcggagcgcacaaag
 15 cagagctgcagggctcagatgaggacgagcacgtgcgagcgctaccggccaccaccaggctggt
 cactgcctcatgtgggctgcaaagcctgcaagaggaagtccaccaccatggatcggcggaaggc
 agccactatgcgcgagcggaggcgctgaagaaggtcaaccaggctttcgaaaccctcaagaggt
 gtaccacgaccaaccccaaccagaggctgcccaggtggagatcctcaggaatgccatccgctac
 atcgagagcctgcaggagttgctgagagagcaggtggagaactactatagcctgccgggacagag
 20 ctgctcggagcccaccagccccacctccaactgctctgatggcatgccgaatgtaacagtcctg
 tctggtccagaaagagcagtacttttgacagcatctactgtcctgatgtatcaaagtatatgcc
 acagataaaaactccttatccagcttgattgcttatccaacatagtggaccggatcacctcctc
 agagcaacctgggttgctctccagcatctggcttctctctctccagttgccagcaccgattcac
 agcctgcaactccaggggcttctagtctccaggttatctatcatgtgctatga (SEQ ID
 25 NO:82)

[00371] Homo sapiens PR-domain-containing protein 16 (PRDM16)(AF294278.1) SEQ ID NO:19 (protein) and SEQ ID NO:83 (DNA)

MRSKARARKLAKSDGDVNNMYEPNRDLLASHSAEDEAEDSAMSPIVGSPPPFPTSEDFTPKEG
 SPYEAPVYIPEDIPIPADFELRESSIPGAGLGWAKRKMEAGERLGPCVVVPRAAAKETDFGWEQ
 30 ILTDVEVSPQEGCITKISEDLGSEKFCVDANQAGAGSWLKYIRVACSCDDQNLTMCISEQVIYY
 KVIKDIEPGEELLVHVKEGVYPLGTVPPLGDEEPTFRCEDELFQSKLDLRHKKYTCGSVGAA
 LYEGLAELKPEGLGGSGQAHECKDCERMFNPKYSLEQHMVIHTEEREYKCDQCPKAFNWKSNF
 IRHQMSHDSGKRFECENCVKVFTDPSNLQRHIRSQHV GARAHACPDCKTFATSSGLKQHKHIHS
 TVKPFICEVCHKSYTQFSNLCHRKHMHADCRTOIKCKDCGMFSTTSSLNKHRRFCEGKNHYTPG
 35 GIFAPGLPLTPSPMMDKAKPSPSLNHASLGFN EYFPYRPHPGSLPFSTAPPTFPALTPGFPGIFP
 PSLYPRPPLLPTSLCLKSPLNHTQDAKLPSPLGNPALPLVSAVSNSSQGTAAAGPEEKFESRLE
 DSCVEKLKTRSSDMSDGSDFEDVNTTTGTDLDTTGTGSDLDSDVSDPDKDKGKSAEGQPKF
 GGGLAPPGAPNSVAEVPVFYSQHSFFPPDEQLLTATGAAGDSIKAIASIAEKYFGPGFMGMQEK
 KLGS LPHYSAFPFQFLPNFPHSLYPFTDRALAHNLLVKAEPKSPRDALKVGGPSAECPFDLTTKP
 40 KDVKPILPMPKGPSAPASGEEQPLDLSIGSRARASQNGG GREPRKNH VYGERKLGAGEGLPQVCP
 ARMPQQPPLHYAKPSFFMDPIYRVEKRVKTD PVGALKEKYL RPSPLL FHPQMSAIETMTEKLES
 FAAMKADSGSSLQPLPHHPFNFRSPPTLS DPILRK GKERYTCRYCGKIFPRSANLTRHLRHTTG
 EQPYRCKYCDRSFSSISNLQRHVRNIHNKEKPFKCHLCNRCFGQQTNLDRHLKKHEHENAPVSQH
 PGVLTNHLGTSASSPTSESDNHALLDEKEDSYFSEIRNF IANSEM NQASTRTEK RADMQIVD GSA
 45 QCPGLASEKQEDVEEEDDDDDLEEDDED SLAGKSQDDTVSPAPEPQAAYE DEEDEE PAASLAVGFD
 HTRCAEDHEGGLLALEPMPTFGKGLDLRAAEEAF EVKDV LNSTLDSEAL KHTLCRQAKN QAYA
 MMLS LSEDTP LHTPSQGS LDAWLKVTGATSESGAFHPINHL (SEQ ID NO:19)

atgcgatccaaggcgagggcgaggaagctagccaaaagtgcggtgacgttgtaaataatatgta
 tgagcccaaccgggacctgctggccagccacagcgcgaggacgaggccgaggacagtgccatgt
 50 cgcccatccccgtggggtcaccgcccccttccccaccagcagaggacttcaccccccaaggagggc

tcgccgtacgagggccctgtctacattcctgaagacattccgatcccagcagacttcgagctccg
agagtcctccatcccaggggctggcctgggggtctgggccaagaggaagatggaagccggggaga
ggctggggccctgctggtggtgccccggcgcggaagagacagacttcggatgggagcaa
atactgacggacgtggaagtgtcgccccaggaaggctgcatcacaagatctccgaagacctggg
5 cagtgagaagtctgctggtgatgcaaatcaggcggggctggcagctggctcaagtacatccgtg
tggcgtgctcctgcatgaccagaacctcaccatgtgtcagatcagtgagcaggtaatttactat
aaagtcattaaggacattgagccaggtgaggagctgctggtgcacgtgaaggaaggcgtctaccc
cctgggcacagtgccgcccggcctggacgaggagcccacgttccgctgtgacgagtgtgacgaac
tcttccagtccaagctggacctgcgcgccataagaagtacacgtgtggctcagtgggggctgcg
10 ctctacgagggcctggctgaggagctcaagcccaggggccttggcggtggcagcgcccaagccca
cgagtgaaggactgcgagcggatgttccccacaagtaacagcctggagcagcacatggtcatcc
acacggaggagcgcgagtacaaatgcgaccagtgtcccaaggccttcaactggaagtccaacttc
atccgccaccagatgtcccacgacagcggcaaacgcttcgaatgtgaaaactgctgaaagtgtt
cacggaccccagcaaccttcagcggcacatccgctcgcagcacgtgggcgctcggggcccacgcct
15 gccccgactgcggaagaccttcgccacgtcctccggcctcaagcagcacaagcatatccacagc
acggtgaagcctttcatatgtgaggtctgccacaagtcctacacgcagttctccaacctgtgccg
gcacaagcggatgcacgcccactgcccgcacgcagatcaagtgaaggactgtggccagatgttca
gcactacctcctccctcaacaagcacccggcgcttctgcgagggcaagaaccattacgcggggc
ggcatctttgccccgggcctgcccttgacccccagccccatgatggacaaggcaaaacctcccc
20 cagcctcaatcacgccagcctgggcttcaacgagtaactttccctacaggccgcacccggggagcc
tgcccttctccacggcgccctcccacgttccccgcactcacccccggcttcccgggcatcttccct
ccatccttgtacccccggcgccctctgctacctcccacatcgctgctcaagagccccctgaacca
caccagagcagccaagctccccagtcctcctggggaaccacagccctgcccctggtctccgccgtca
gcaacagcagccagggcacgacggcagctgcggggcccagggagaagttcgagagccgcctggag
25 gactcctgtgtggagaagctgaagaccaggagcagcgacatgtcggacggcagtgactttgagga
cgtcaacaccaccacggggaccgacctggacacgaccacggggacgggctcggacctggacagcg
acgtggacagcgacctgacaaggacaagggcaagggcaagtccgccgagggccagcccaagttt
gggggcggttggcgcccccgggggccccgaacagcgtggccgaggtgcctgtcttctattccca
gcaactcattcttcccgccacccgacgagcagctgctgactgcaacgggcgcgcgggggactcca
30 tcaaggccatcgcatccattgccgagaagtactttggccccggcttcatggggatgcaggagaag
aagctgggctcgctcccctaccactcggcgcttccccttccagttcctgcccacttccccactc
cctttaccccttcacggaccgagccctcgcccacaacttgcctggtcaaggccgagccaaagtcac
cccgggacgcccctcaaggtgggcgggccccagtgccgagtgcccctttgatctcaccaccaagccc
aaagacgtgaagcccatcctgcccatgcccgaagggccccctcgggccccgcacccggcgaggagca
35 gccgctggacctgagcatcggcagccgggcccgtgccagccaaaacggcgggcggggagcccc
gcaagaaccacgtctatggggaacgcaagctgggcgcggcgaggggctgcccaggtgtgcccg
gcgcggtatgccccagcagccccgcctccactacgccaagccctcgcccttcttcatggacccat
ctacagggtagaaaagcgggaaggtcacagaccccgtgggagccctgaaggagaagtaacctgcggc
cgtccccgctgctcttccacccccagatgtcagccatagagaccatgacagagaagctggagagc
40 tttgcagccatgaaggcggactcgggcagctccttgacgcccctccccaccaccccttcaactt
ccggtccccacccccaaacgtctcgcgaccccatcctcaggaagggaaggagcgatacacgtgca
ggtactgtgggaagatcttccccagatcagccaatctcaccagacacctgaggacgcacactggg
gagcagccgtacaggtgtaagtactgcgaccgctccttcagcatctcttccgaacctccagcgga
cgtccggaacatccacaacaaggagaagcctttcaagtgccacctgtgcaaccgctgcttcgggc
45 agcagaccaacctggaccggcacctcaagaagcacgagcacgagaacgcaccagtgcagccagcac
cccggggtcctcacgaaccacctggggaccagcgctcctctcccacctcagagtcgggacaacca
cgcacttttagacgagaaagaagactcttattttctcgaaatcagaaactttattgccaatagt
agatgaaccaagcatcaacgcgaacagagaaacgggcgacatgcagatcgtggacggcagtgcc
cagtgtccaggcctagccagtgcagaagcaggaggacgtggaggaggaggacgacgatgacctgga
50 ggaggacgatgaggacagcctggcggggaagtcgcaggatgacaccgtgtccccgcacccgagc

cccaggccgcctacgaggatgaggaggatgaggagccagccgcctccctggccgtgggctttgac
cacacccgaaggtgtgctgaggaccacgaaggcgggtctgttagctttggagccgatgccgacttt
tgggaaggggctggacctccgcagagcagctgaggaagcatttgaagttaagatgtgcttaatt
ccaccttagattctgaggctttaaaacatacactgtgcaggcaggctaagaaccaggcatatgca
5 atgatgctgtccctttccgaagacactcctctccacacccctcccagggttctctggacgcttg
gttgaaggtcactggagccacgtcggagctctggagcatttcaccccatcaaccacctctga
(SEQ ID NO:83)

[00372] Homo sapiens paired box 6 (PAX6) (NM_001604.5) SEQ ID NO:20 (protein) and
SEQ ID NO:84 (DNA)

10 MQNSHSGVNQLGGVFNVRPLPDSTRQKIVELAHSGARPCDISRILQTHADAKVQVLDNQNVSNQ
CVSKILGRYYETGSIRPRAIGGSKPRVATPEVVSKIAQYKRECPSIFAWEIRDRLLESEGVCNDN
IPSVSSINRVLRLNLASEKQQMGADGMYDKLRMLNGQTGSWGTRPGWYPGTSVPGQPTQDGCQQQE
GGGENTNSISSNGEDSDEAQMRLQLKRKLQRNRTSFTQEQIEALEKEFERTHYPDVFAERERLAAK
IDLPEARIQVWFSNRRRAKWRREEKLNRNRRQASNTPSHIPISSSFSTSVYQPIQPPTPVSSFTS
15 GSMLGRTDTALTNTYSALPPMPSFTMANNLPMQPPVPSQTSSYSCMLPTSPSVNGRSYDTYTPPH
MQTHMNSQPMGTSGTTSTGLISPGVSVPVQVPGSEPDMSQYWPRLQ (SEQ ID NO:20)

atgcagaacagtcacagcggagtgaatcagctcgggtggtgtctttgtcaacggggcggccactgcc
ggactccacccggcagaagattgtagagctagctcacagcggggcccgccgtgagacatttccc
gaattctgcagacccatgcagatgcaaaagtccaagtgtggacaatcaaaacgtgtccaacgga
20 tgtgtgagtaaaattctgggcaggtattacgagactggctccatcagacccagggaatcggtgg
tagtaaacccgagagtagcagactccagaagttgtaagcaaaatagcccagtataagcgggagtgcc
cgtccatctttgcttgggaaatccgagacagattactgtccgaggggtctgtaccaacgataac
ataccaagcgtgtcatcaataaacagagttcttcgcaacctggctagcgaaaagcaacagatggg
cgcagacggcatgtatgataaactaaggatgttgaaacgggcagaccggaagctggggcaccgcc
25 ctggttgggtatccggggacttcgggtgccaggggcaacctacgcaagatggctgccagcaacaggaa
ggagggggagagaataaccaactccatcagttccaacggagaagattcagatgaggctcaaatgcg
acttcagctgaagcgggaagctgcaaagaaatagaacatcctttaccaagagcaaattgaggccc
tggagaaagagtttgagagaacccattatccagatgtgtttgcccagaaaagactagcagccaaa
atagatctacctgaagcaagaatacaggtatggttttctaatacgaaggggccaaatggagaagaga
30 agaaaaactgaggaatcagagaagacaggccagcaacacacctagtcatttctatcagcagta
gtttcagcaccagtgcttaccacccaattccacaaccaccacacccggtttcctccttcacatct
ggctccatgttgggcccgaacagacacagccctcacaacacctacagcgtcttgccgcctatgcc
cagcttcaccatggcaaataacctgcctatgcaacccccagtcaccagccagacctcctcact
cctgcatgctgccaccagcccttcgggtgaatgggcggagttatgatacctacacccccccat
35 atgcagacacacatgaacagtcagccaatgggcacctcgggcaccacttcaacaggactcatttc
ccctggtgtgtcagttccagttcaagttcccgggaagtgaacctgatatgtctcaatactggccaa
gattacagtaa (SEQ ID NO:84)

[00373] SEQ ID NO:21: Homo sapiens HNF1 homeobox A (HNF1A) (NM_000545.5) SEQ
ID NO:21 (protein) and SEQ ID NO:85 (DNA)

40 MVSKLSQLQTELLAALLESGLSKEALIQALGEPGPYLLAGEGPLDKGESCGGGRGELAEPLNGLG
ETRGSEDETDDDGEDFTPPILKELENLSPEEAAHQKAVVETLLQEDPWRVAKMVKSYLQQHNIPQ
REVVDTTGLNQSHLSQHLNKGTPMKTQKRAALYTWYVRKQREVAQQFTHAGQGGLIEPTGDELP
TKKGRNRNRFKWGPASQQILFQAYERQKNPSKEERETLVEECNRAECIQRGVSPSQAQGLGSNLVT
EVRVYNWFANRRKEEAFRHKLAMDTYSGPPPGPGPALPAHSSPGLPPPALSPSKVHGVRYGQP
45 ATSETAEVPSSSGGPLVTVSTPLHQVSPTGLEPSHSLSTEAKLVSAAGGPLPPVSTLTALHSLE
QTSPGLNQPPQNLIMASLPGVMTIGPGEPASLGPTFTNTGASTLVIGLASTQAQSVPVINSMGSS
LTTLQPVQFSQPLHPSYQQPLMPVQSHVTQSPFMAQMAQLQSPHALYSHKPEVAQYTHTGLLPQ
TMLITDTTNLSALASLTPKQVFTSDTEASSESLHTPASQATTLHVPSQDPAGIQHLQPAHRLS
ASPTVSSSSLVLYQSSDSSNGQSHLLPSNHSVIETFISTQMASSSQ (SEQ ID NO:21)

atggttttctaaactgagccagctgcagacggagctcctggcgccctgctcgagtcagggctgag
 caaagaggcactgatccaggcactgggtgagccggggccctacctcctggctggagaaggccccc
 tggacaagggggagtcctgcggcgccggtcgaggggagctggctgagctgcccaatgggctgggg
 gagactcggggctccgaggacgagacggacgacgatggggaagacttcagcccaccatcctcaa
 5 agagctggagaacctcagccctgaggaggcgggccaccagaaagccgtgggtggagacccttctgc
 aggaggaccctgtggcgtgtggcgaagatggtcaagtccctacctgcagcagcacaacatcccacag
 cgggaggtggctgataccactggcctcaaccagtccacctgtcccaacacctcaacaagggcac
 tcccatgaagacgcagaagcgggccgcccctgtacacctggtaacgtccgcaagcagcagaggtgg
 cgcagcagttcacccatgcagggcagggagggctgattgaagagcccacaggtgatgagctacca
 10 accaagaagggcgagggaaccgtttcaagtggggcccagcatccacgcagatcctgttccaggc
 ctatgagaggcagaagaaccctagcaaggaggagcgagagacgctagtggaggagtgaataggg
 cggaatgcatccagagaggggtgtccccatcacaggcacaggggctgggctccaacctcgtcacg
 gaggtgctgtgtctacaactggtttgccaaaccggcgcaaagaagaagccttccggcacaagctggc
 catggacacgtacagcgggccccccccaggggccaggccccgggacctgcgctgcccgcctcacagct
 15 cccctggcctgcctccacctgcctctccccagtaaggtccacgggtgtgcgctatggacagcct
 gcgaccagtgcagactgcagaagtacctcaagcagcgggcggtcccttagtgacagtgtctacacc
 cctccaccaagtgtccccacgggcctggagcccagccacagcctgctgagtacagaagccaagc
 tgggtctcagcagctggggggccccctccccctgtcagcaccctgacagcactgcacagcttggag
 cagacatccccaggcctcaaccagcagccccagaacctcatcatggcctcacttccctggggctcat
 20 gaccatcgggcctgggtgagcctgcctccctgggtcctacgttcaccaacacaggtgcctccacce
 tgggtcatcggcctggcctccacgcaggcacagagtgtgccgggtcatcaacagcatgggcagcagc
 ctgaccaccctgcagcccgtccagttctcccagcgcgtgcacccctcctaccagcagccgctcat
 gccacctgtgcagagccatgtgaccagagcccccttcatggccaccatggctcagctgcagagcc
 cccacgcccctctacagccacaagcccaggtggcccagtagacccacacgggcctgctcccgcag
 25 actatgctcatcaccgacaccaccaacctgagcgcctggccagcctcacgcccaccaagcaggt
 cttcacctcagacactgaggcctccagtgagtcggggttcacacgcccgcctctcaggccacca
 ccctccacgtccccagccaggacctgcccggcatccagcacctgcagccggcccacgggctcagc
 gccagccccacagtgtcctccagcagcctgggtgctgtaccagagctcagactccagcaatggcca
 gagccacctgctgccatccaaccacagcgtcatcgagaccttcctcctccaccagatggcctctt
 30 cctcccagtaa (SEQ ID NO:85)

[00374] Homo sapiens forkhead box A3 (FOXA3) (NM_004497.2) SEQ ID NO:22 (protein) and SEQ ID NO:86 (DNA)

MLGSVKMEAHDLAEWSYYPEAGEVYSPVTPVPTMAPLNSYMTLNPLSSPYPPGGLPASPLPSGPL
 APPAPAAPLGPTFPGLGVSGSSSSGYGAPGGLVHGKEMPKGYRRPLAHAKPPYSYISLITMAI
 35 QQAPGKMLTLSEIYQWIMDLFPYYRENQQRWQNSIRHLSFNDCFVKVARSPDKPGKGSYWALHP
 SSGNMFENGCYLRRQKRFKLEEKVKKGSGAATTTTRNGTGSAASTTTTPAATVTSPPQPPPPAPEP
 EAQGGEDVGALDCGSPASSTPYFTGLELPGELKLDAPYNFNHPFSINNLMSEQTPAPPKLDVGFG
 GYGAEGGEPGVYYQGLYSRSLNAS (SEQ ID NO:22)
 atgctgggctcagtgaagatggaggcccatgacctggccgagtgagctactaccgggaggcggg
 40 cgaggtctactcgccgggtgaccccagtgcccaccatggccccctcaactcctacatgacctga
 atcctctaagctctccctatccccctggggggctccctgcctccccactgccctcaggaccctg
 gcacccccagcacctgcagccccctggggcccactttcccaggcctgggtgtcagcgggtggcag
 cagcagctccgggtacggggccccgggtcctgggctggtgcacgggaaggagatgccgaaggggt
 atcggcggccccctggcacacgccaaagccaccgtattcctatatctcactcatcaccatggccatc
 45 cagcaggcgccgggcaagatgctgaccttgagtgaatctaccagtggtatcatggacctcttccc
 ttactaccgggagaatcagcagcgtggcagaactccattcgccactcgctgtctttcaacgact
 gcttcgtcaaggtggcgcggttccccagacaagcctggcaagggctcctactgggccttacacccc
 agctcagggaaacatgtttgagaatggctgctacctgcgcgccagaaacgcttcaagctggagga
 gaaggtgaaaaaagggggcagcggggtgccaccaccaccaggaacgggacaggggtctgctgcct
 50 cgaccaccacccccgcggccacagtcacctccccgccccagccccgcctccagccctgagcct

gaggcccagggcggggaagatgtgggggctctggactgtgggtcaccgcgttcctccacacccta
tttactggcctggagctcccaggggagctgaagctggacgcgcctacaacttcaaccaccctt
tctccatcaacaacctaatagtcagaacagacaccagcacctcccaaactggacgtgggggttggg
ggctacgggggtgaaggtggggagcctggagctctactaccagggcctctattcccgcgtctttgct
5 taatgcatacctag (SEQ ID NO:86)

[00375] Homo sapiens forkhead box A1 (FOXA1) (NM_004496.3) SEQ ID NO:23 (protein)
and SEQ ID NO:87 (DNA)

MLGTVKMEGHETSDWNSYYADTQEAYSSVPVSNMNSGLGSMNSMNTYMTMNTMTTSGNMTPASFN
MSYANPGLGAGLSPGAVAGMPGGSAGAMNSMTAAGVTAMGTALSPSGMGAMGAQQAASMNGLGPY
10 AAAMNPMCSMPMAYAPSNLGRSRAGGGGDAKTFKRSYPHAKPPYSYISLITMAIQQAPSKMLTLSE
IYQWIMDLFPYRQNRQNRWQNSIRHSLSFNDCFVKVARSPDKPGKGSYWTLHPDSGNMFENG CYL
RRQKRKFCEKQPGAGGGGSGSGSGSAGKGPESRKDPGASNPSADSPLHRGVHGTGQLEGAPA
PGPAASPQTLDHSGATATGGASELKTPASSTAPPISSGPGALASVPASHPAHGLAPHESQLHLKG
DPHYSFNHPFSINNLMSSEQQHKLDFKAYEQALQYSPYGSTLPASLPLGSASVTTTRSPIEPSAL
15 EPAYYQGVYSRPVLNTS (SEQ ID NO:23)

atgctggggtcagtgaaatggaggcccatgacctggccgagtgagctactaccgggaggcggg
cgaggtctactcgccggtgaccccgagtgccaccatggccccctcaactcctacatgacctga
atcctctaagctctccctatccccctggggggtcctgctccccactgacctcaggacctctg
gcacccccagcacctgcagccccctggggcccaacttcccaggcctgggtgtcagcggtggcag
20 cagcagctccgggtacggggccccgggtcctggggtggtgcacgggaaggagatgccgaaggggt
atcggcggccccctggcacacgcccaagccaccgtattcctatatctcactcatcaccatggccatc
cagcaggcgccgggcaagatgctgaccttgagtgaatctaccagtgatcatggacctcttccc
ttactaccgggagaatcagcagcgctggcagaactccattcgccactcgctgtctttcaacgact
gcttcgtcaaggtggcgcggttccccagacaagcctggcaagggtcctactgggcctacacccc
25 agctcaggaacatgtttgagaatggctgctacctgcgcgccagaaacgcttcaagctggagga
gaaggtgaaaaaagggggcagcggggtgccaccaccaccaggaacgggacaggggtctgctgcct
cgaccaccaccccccgcgggccacagtcacctccccgcgccagccccgcctccagccccctgagcct
gaggcccagggcggggaagatgtgggggctctggactgtgggtcaccgcgttcctccacacccta
tttactggcctggagctcccaggggagctgaagctggacgcgcctacaacttcaaccaccctt
30 tctccatcaacaacctaatagtcagaacagacaccagcacctcccaaactggacgtgggggttggg
ggctacgggggtgaaggtggggagcctggagctctactaccagggcctctattcccgcgtctttgct
taatgcatacctag (SEQ ID NO:87)

[00376] Homo sapiens forkhead box A2 (FOXA2) (NM_021784.4) SEQ ID NO:24 (protein)
and SEQ ID NO:88 (DNA)

MHSASSMLGAVKMEGHEPSDWSSYYAEPEGYSSVSNMAGLGMNGMNTYMSMSAAAMGSGSGNMS
AGSMNMSSSYVGAGMSPSLAGMSPGAGAMAGMGGSAGAAGVAGMGPHLSPSLSPGGAAGAMGGL
APYANMNSMSPMYGQAGLSRARDPKTYRRSYTHAKPPYSYISLITMAIQQSPNKMLTLSEIYQWI
MDLFPFYRQNRQNRWQNSIRHSLSFND CFLKVPSPDKPGKGSFWTLHPDSGNMFENG CYLRRQKR
FKCEQLALKEAAGAAGSGKKAAGAQAQAQLGEAAGPASETPAGTESPHSSASPCQEHKRGG
40 GELKGTAAALSPPEPAPSPGQQQQAHAHLLGPPHHPGLPPEAHLKPEHHYAFNHPFSINNLMS
EQQHSHSHHHQPHKMDLKAYEQVMHYPGYGSPMPGSLAMGPVTNKTGLDASPLAADTSYYQGVY
SRPIMNSS (SEQ ID NO:24)

atgcactcggttcagtatgctgggagcggtgaagatggaagggcacgagccgtccgactggag
cagctactatgcagagccccgagggtactcctccgtgagcaacatgaacgcgcggtcggggatga
45 acggcatgaacacgtacatgagcatgtcggcgcccgccatgggcagcggtcgggcaacatgagc
gcggggtccatgaacatgtcgtcgtagctgggcgctggcatgagcccgctccctggcggggatgtc
ccccggcgcgggcgccatggcgggcatggcgggctcgccggggcgccggcggtggcgggcatgg
ggccgcacttgagtcccagcctgagcccgctcgggggggcaggcgccggggccatggcgggcctg
gccccctacgccaacatgaactccatgagccccatgtacgggcaggcgggcctgagccgcgccccg
50 cgaccccaagacctacaggcgagctacacgcacgcaaagccgcctactcgtagcatctcgctca

tcaccatggccatccagcagagccccaacaagatgctgacgctgagcagagatctaccagtggatc
 atggacctcttcccccttctaccggcagaaccagcagcgctggcagaactccatccgccactcgct
 ctcttcaacgactgtttcctgaaggtgccccgctcgcccgacaagcccggaagggtctcttct
 ggacctgcacctgactcgggcaacatgttcgagaacggctgctacctgcgccgccagaagcgc
 5 ttcaagtgcgagaagcagctggcgctgaaggaggccgcaggcgccggcgagcggaagaaggc
 ggccgcccggagcccaggcctcacaggctcaactcggggaggcccgccgggcccgcctccgagactc
 cggcgggacccgagtcgcctcactcgagcgctccccgtgccaggagcacaagcgagggggcctg
 ggagagctgaaggggacgccggtgcggcgctgagccccccagagccggcgccctctcccgggca
 gcagcagcaggccgcccacctgctgggcccgcaccaccgggctgccgctgagggcc
 10 acctgaagccggaacaccactacgccttcaaccaccggttctccatcaacaacctcatgtcctcg
 gagcagcagcaccaccacagccaccaccaccaccaacccccacaaaatggacctcaaggcctacga
 acaggtgatgcactaccccggtacgggtcccccatgcctggcagcttggccatgggcccgggtca
 cgaacaaaacgggctggagcgctcgccctggccgcagataacctcctactaccaggggggtgtac
 tccccggcccattatgaactcctcttaa (SEQ ID NO:88)

[00377] SEQ ID NO:25: Homo sapiens CCAAT/enhancer binding protein (C/EBP), alpha (CEBPA) (NM_001287435.1) SEQ ID NO:25 (protein) and SEQ ID NO:89 (DNA)

MSSHLQSPHPAPSSAAFGRGAGPAQPPAPPAPEPLGGICEHETSIDISAYIDPAAFNDEFLA
 DLFQHSRQEQEKAAVGP TGGGGGDFDYPGAPAGPGGAVMPGGAHGPPPGYGCAAAGYLDGRLE
 PLYERVGAPALRPLVIKQEPREDEAKQLALAGLFPYQPPPPPPSHPHPHPPAHLAAPHLQFQ
 20 IAHCGQTTMHLQPGHPTPPPTPVPSHPAPALGAAGLPGPGSALKGLGAHPDLRASGGSGAGKA
 KKSVDKNSNEYRVRERRNNIAVRKSRDKAKQRNVETQQKVLELTSDNDRLRKRVEQLSRELDTLR
 GIFRQLPESSLVKAMGNCA (SEQ ID NO:25)

atgagcagccacctgcagagcccccgacgcgcccagcagcgccgcttcggctttccccgggg
 cgccgggccccgcgcagcctcccgccccacctgcgcggccggagccgctgggcggcacatctgcgagc
 25 acgagacgtccatcgacatcagcgccctacatcgacccggccgccttcaacgacgagttcctggcc
 gacctgttccagcacagccggcagcaggagaaggccaaggcgccgtgggccccacgggcggcgg
 cggcgggcgggcacttttgactacccgggcgcgccccgcgggccccggcgggcgccgtcatgcccgggg
 gagcgacgggcccccgcccggtacggctgcgcggcgccggcgctacctggacggcaggctggag
 cccctgtacgagcgcgctcgggcgccggcgctgcggcgcgctggtgatcaagcaggagccccgcga
 30 ggaggatgaagccaagcagctggcgctggccggcctcttcccttaccagccgcccgcgcgcgcgc
 cgccctcgcacccgcacccgcacccgcgcgcgcgcacctggccgccccgcacctgcagttccag
 atcgcgactgcggccagaccacctgcacctgcagcccggtcacccacgcgcgcgcgcacgc
 cgtgccagcccgaccccgcgccccgcgctcggtgcgcggcgccctgcggggccctggcagcgcg
 tcaaggggtgggcgccgcgcaccccgacctccgcgcgagtgccggcagcgggcgcgggcaaggcc
 35 aagaagtcgggtggacaagaacagcaacgagtagccgggtgcggcgcgagcgcaacaacatcgcggt
 gcgcaagagccgcgacaaggccaagcagcgcaacgtggagacgcagcagaaggtgctggagctga
 ccagtgacaatgaccgcctgcgcaagcggtggaacagctgagccgcgaactggacacgctgcgg
 ggcattctccgagctgccagagagctccttggtaaggccatgggcaactgcgcgtga (SEQ
 ID NO:89)

[00378] Homo sapiens Spi-1 proto-oncogene (SPI1) (PU.1) (NM_001080547.1) SEQ ID NO:26 (protein) and SEQ ID NO:90 (DNA)

MLQACKMEGFPLVPPQPSDLVPYD TDLYQRQTHEYYPYLSSDGESHSDHYWDFPHPHVHSEFES
 FAENNFTELQSVQPPQLQQLYRHMELEQMHVLDTPMVPHPHSLGHQVSYLPRMCLQYPSLSPAQP
 SSDEEEGERQSPPLEVSDGEADGLEPGPGLLPGETGSKKKIRLYQFLLDLLRSGDMKDSIWWVDK
 45 DKGTFQFSSKHKEALHRWGIQKGNRKKMTYQKMARALRNYGKTGEVKVKVKKLTYQFSGEVLGR
 GGLAERRHPPH (SEQ ID NO:26)

atgttacaggcgctgcaaaatggaagggtttccccctcgtccccctcagccatcagaagacctggt
 gccctatgacacggatctataccaacgccaacgcacgagtagtattaccctatctcagcagtgatg
 50 gggagagccatagcgaccattactgggacttccacccccaccacgtgcacagcgagttcgagagc
 ttcgcccagagaacaacttcacggagctccagagcggtgcagccccgcagctgcagcagctctaccg

ccacatggagctggagcagatgcacgtcctcgatacccccattggtgccaccccatcccagtcttg
gccaccaggtctcctacctgccccggatgtgcctccagtacccatccctgtccccagcccagccc
agctcagatgaggaggagggcgagcgccagagccccccactggaggtgtctgacggcgaggcgga
5 tggcctggagccccgggctgggctcctgctggggagacaggcagcaagaagaagatccgcctgt
accagttcctgttggacctgctccgcagcgccgacatgaaggacagcatctggtgggtggacaag
gacaagggcaccttccagtctctcgctccaagcacaaggaggcgctggcgccacgctggggcatcca
gaagggcaaccgcaagaagatgacctaccagaagatggcgcgcgcgctgcgcaactacggcaaga
cgggcgaggtcaagaaggtgaagaagaagctcacctaccagttcagcggcggaagtgctggggcgcg
10 gggggcctggccgagcgccgcccacccgccccactga (SEQ ID NO:90)

[00379] Homo sapiens POU class 3 homeobox 2 (POU3F2) (Brn2) (NM_005604.3) SEQ ID
NO:27 (protein) and SEQ ID NO:91 (DNA)
MATAASNHYSLTSSASIVHAEPPGGMQQGAGGYREAQSLVQGDY GALQSN GHPLSHAHQWITAL
SHGGGGGGGGGGGGGGGGGGGGGGGGGSPWSTSPLGQPDIKPSVVVQQGGRGDELHGPGALQQQHQQ
15 QQQQQQQQQQQQQQQQQQQQQRPPLVHHAANHHPGPGAWRSAAAAAHLPPSMGASNGGLLYSQPSF
TVNGMLGAGGQPA GLHHHGLRDAHDEPHHADHHPHSHPHQQPPPPPPPPQGP PGHPGAHHDPHS
DEDTP TSD DLEQFAKQFKQRRIKLGFTQADVGLALGTLYGNVFSQTTICRFEALQLSFKNMCKLK
PLLNKWL E EADSSSGSPTSIDKIAAQGRKKRRTSIEVSVKGALESHFLKCPKPSAQEITSLADS
LQLEKEVVRVWFCNRRQKEKRMTPPGGTLP GAEDVYGGSRDTPPHHGVQTPVQ (SEQ ID
20 NO:27)

atggcgaccgcagcgtctaaccactacagcctgctcacctccagcgcctccatcgctgcacgccga
gccgccccggcgcatgcagcagggcgcggggggtaccgcgaagcgagagcctggtgcagggcg
actacggcgctctgcagagcaacggacacccgctcagccacgctcaccagtgatcaccgcgctg
tcccacggcgggcgggcgggggcggtggcgggcgggggggcgggggcgggcgggggcggg
25 cggcgacggctccccgtggtccaccagccccctgggccagccggacatcaagccctcggtggtgg
tgcagcagggcgggcgggagacgagctgcagggccagggcgccctgcagcagcagcatcagcag
cagcaacagcaacagcagcagcaacagcagcaacagcagcagcagcagcaacagcgccgcc
gcatctggtgcaccacgcccgttaaccaccacccgggacccggggcatggcgaggcgcgggcggtg
cagcgcacctcccaccctccatgggagcgtccaacggcggttgctctactcgcagcccagcttc
acggtgaacggcatgctggggcgccggcgggcgagccggcgccgtctgcaccaccacggcctgcggga
30 cgcgcacgacgagccacacatgccgaccaccacccgcacccgcactcgcaccacacaccagcagc
cgccgcccccgccgcccccgaggggtccgcctggccacccaggcgcgacccacgacccgcactcg
gacgaggacacgcccacctcgacgacctggagcagttcgccaagcagttcaagcagcgggcgat
caaactgggatttaccgaagcggaagctggggctggctctgggcacccctgtatggcaacgtgttct
cgagaccaccatctgcaggttttagggcctgcagctgagcttcaagaacatgtgcaagctgaag
35 cctttgttgaacaagtgggtggaggaggcggaactcgctcctcgggcagccccacgagcatagacaa
gatcgagcgcaagggcgcaagcggaagcggaagcggaagcggaagcggaagcggaagcggaag
tgagagccatttccctcaaatgccccgaagccctcgggccaggagatcacctccctcgcggaagc
ttacagctggagaaggaggtggtgagagtttggttttgaacaggagacagaaagagaaaaggat
gacccctcccggagggactctgccggggcgccgaggatgtgtacgggggagtagggacactccac
40 cacaccacgggggtgcagacgcccgtccagtga (SEQ ID NO:91)

[00380] Homo sapiens forkhead box G1 (FOXG1) (NM_005249.4) SEQ ID NO:28 (protein)
and SEQ ID NO:92 (DNA)
MLDMGDRKEVKMIPKSSFSINSLVPEAVQNDNHHASHGHNSHHPQH HHHHHHHHHHHPPPPAPQP
45 PPPPQQQQPPPPPPPPAPQPQTRGAPAADDDKGPQQLLLPPPPPPPPAAALDGA KADGLGKGKEP
GGGPGE LAPVGPDEKEKGAGAGGEEKKGAGEGGKDGEGGKEGEKNGKYEKPPFSYNALIMMAIR
QSPEKRLTLNGIYEFIMKNFPYYRENKQGWQNSIRHNL SLNKCFVKVPRHYDDPGKGN YWMLDPS
SDDVFIGGTTGKLRRRSTTSRAKLAFKRGARLTSTGLTFMDRAGSLYWPMS PFLSLHHPRASSTL
SYNGTTSAYPSHPMPYSSVLTQNSLGNHNSFSTANGLSVDRLVNGEI PYATHHLTAALAASVPC
GLSVPCSGTYSLNPCSVNLLAGQTSYFFPHVPHPSMTS QSSTSMSARAASSSTSPQAPSTLPCES
50 LRPSLPSFTTGLSGGLSDYFTHQNQGSSSNPLIH (SEQ ID NO:28)

[illegible]

25 **[00381]** Homo sapiens mRNA for ASC1 protein (SLC7A10 gene) (AJ277731.1) SEQ ID NO:29 (protein) and SEQ ID NO:93 (DNA)

30 MAGHTQQPSGRGNRPAPSPSPVPGTVPGASERVALKKEIGLLSACTIIIGNIIGSGIFISPKGV
LEHSGSVGLALFVWVLGGGVLTALGSLCYAELGVAIPKSGGDYAYVTEIFGGLAGFLLWSAVLIM
YPTSLAVISMTFSNYVLQPVPNCIPPTTASRVLSMACLMLLTWVNSSSVRWATRIQDMFTGGKL
LALSLIIGVGLLQIFQGHFEELRPSNAFAFWMTPSVGHLALAFLOGSFAFSGWNFLNYVTEEMVD
ARKNLPRAIFISIPLVTFVYFTFTNIAYFTAMSPQELLSSNAVAVTFGEKLLGYFSWMPVSVALS
TFGGINGYLFYTSRLCFSGAREGHLPSLLAMIHVRHCTPIPALLVCCGATAVIMLVGDYTLINY
VSFINYLCYGVTTILGLLLLRWRRPALHRPIKVNLLIPVAYLVFWAFLLVFSFISEPMVCGVGVII
ILTGVPIFFLGVFWRSKPKCVHRLTESMTHWGQELCFVVYPQDAPEEEENGPCPPSLLPATDKPS
KPQ (SEQ ID NO:29)

35 atggccgggccacacgcagcagcgcgagcgggcgcgggaaaccccaggcctgcgccctcgccctcccc
agtcacagggaaccgtccccggcgccctcggagcgggtggcgctcaagaaggagatcgggctgctga
gcgcctgcaccatcatcatcgggaacatcatcggtcgggcatcttcatctcgcccaaggggggtc
ctggagcactcagggtccgtgggtctggccctgttcgtctgggtcctgggtgggggcgtgacggc
40 tctgggctccctctgctatgcagagctgggagtcgccatccccaagtctggcggggactacgcct
acgtcacagagatcttcgggggcctggctggctttctgctgctctggagcgcgctcctcatcatg
taccaccaccagccttgctgtcatctccatgaccttctccaactacgtgctgcagcccgctgttccc
caactgcatcccccccaccacagcctcccgggtgctgtccatggcctgctgatgctcctgacat
gggtgaacagctccagtgtgcgctgggccacgcgcattccaggacatgttcacaggcggggaagctg
ctggccttgctccctcatcatcggcgctgggccttctccagatcttccaaggacacttcgaggagct
45 gaggcccagcaatgcctttgctttctggatgacgcctccgtgggacacctggccctggccttcc
tccagggtccttcgccttcagtggctggaaacttctcaactatgtcaccgaggagatggttgac
gcccgaaagaacctacctcgcgccatcttcatctccatcccactggtgaccttcgtgtacacgtt
caccaacattgcctacttcacggccatgtccccccaggagctgctctcctccaatgcggtggctg
tgaccttcggggagaagctgctgggctacttttcttgggtcatgcctgtctcgtggctctgtca
50 accttcggagggatcaatggttacctgttcacctactccaggctgtgcttctctggagcccgca

ggggcacctgccagcctgctggccatgatccacgtcagacactgcacccccatccccgcctcc
tcgtctgttgcggggccagccggtcatcatgctcgtgggcgacacgtacacgctcatcaactat
gtgtccttcatcaactacctctgtacggcggtcaccatcctgggcctgctgctgctgcgctggag
gcggcctgcactccacaggcccatcaaggtgaaccttctcatccccgtggcgtacttggtcttct
gggccttcctgctgggtcttcagcttcatctcagagcctatggtctgtggggtcggcgtcatc
atccttacgggggtgccattttctttctgggagtgttctggagaagcaaaccaaagtgtgtgca
cagactcacagagtcctatgacacactggggccaggagctgtgtttcgtggtctacccccaggacg
ccccgaagaggaggagaatggcccctgccaccctcctgctgctgccacagacaagccctcg
aagccacaatga (SEQ ID NO:93)

10 **[00382]** Homo sapiens achaete-scute family bHLH transcription factor 1 (ASCL1)
(NM_004316.3) SEQ ID NO:30 (protein) and SEQ ID NO:94 (DNA)

MESSAKMESGGAGQQPQPQPQQPFLPPAACFFATAAAAAAAAAAAAAAQSASQQQQQQQQQQQQQAPQ
LRPAADGQPSGGGHKSAPKQVKRQRSSSPELMRCKRRLNFSGFGYSLPQQQPAAVARRNERERNR
VKLVNLGFATLREHVPNGAANKKMSKVETLRSAYEYIRALQQLLDEHDAVSAAFQAGVLSPTISP
NYSNDLNSMAGSPVSSYSSDEGSYDPLSPPEEQELLDFTNWF (SEQ ID NO:30)

[illegible]

[00383] Homo sapiens Nurr1 gene (AB017586) SEQ ID NO:31 (protein) and SEQ ID NO:95 (DNA)

30 MPCVQAQYGSSPQGASPASQSYSYHSSGEYSSDFLTPEFVKFSMDLTNTEITATTSPLPSFSTFMD
NYSTGYDVKPPCLYQMPLSGQQSSIKVEDIQMHNYQQHSHLPPQSEEMMPHSGSVYYKPSSPPTP
TTPGFQVQHSPMWDPPGSLHNFHQNYVATTHMIEQRKTPVSRLSLFSFKQSPPGTPVSSCQMRFD
GPLHVPMPNPEPAGSHHVVDGQTFVNPPIRKPASMGFPGLQIGHASQLLDTQVPSPPSRGSPSNE
35 GLCAVCGDNAACQHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCPVDKRRRNRCQYCRFQKCLA
VGMVKEVVRTDSLKGRRGRLLPSKPKSPQEPSPPSPVSLISALVRAHVDSNPAMTSLDYSRFQAN
PDYQMSGDDTQHIQQFYDLLTGSMEIIRGWAEKIPGFADLPKADQDLLFESAFLELFLVRLAYRS
NPVEGKLIFCNGVVLHRLQCVRGFGIEWIDSIVEFSSNLQNMNIDISAFSCIAALAMVTERHGLKE
PKRVEELQNKIVNCLKDHVTFNNGGLNRPNYLSKLLGKLPELRTLCTQGLQRIFYLKLEDLVPPP
AIIDKLFLDTLPF (SEQ ID NO: 31)

40 atgccttgtgttcagggcgcagtatgggtcctcgccctcaaggagccagccccgcttctcagagcta
cagttaccactcttcgggagaatacagctccgatttcttaactccagagtttgtaagtttagca
tggacctcaccaacactgaaatcactgccaccacttctctcccagcttcagtaacctttatggac
aactacagcacaggctacgacgtcaagccaccttgcttgtaacaaatgccccgtgccggacagca
gtcctccattaaggtagaagacattcagatgcacaactaccagcaacacagccacctgcccccc
45 agtctgaggagatgatgccgcactccgggtcggtttactacaagccctctcgcgcccgacgccc
accaccccgggcttccagggtgcagcacagccccatgtgggacgacccgggatctctccacaactt
ccaccagaactacgtggccactacgcacatgatcgagcagaggaaaacgccagttctccgcctct
ccctcttctcctttaagcaatcgccccctggcaccccggtgtctagttgccagatgcgcttcgac
gggccccgtgcacgtccccatgaacccggagcccgccggcagccaccacgtggtggacgggcagac
50 cttcgctgtgcccaaccccattcgcaagcccgctccatgggcttcccgggctgcagatcgcc

acgcgtctcagctgctcgacacgcaggtgccctcaccgccgtcgcggggctccccctccaacgag
 gggctgtgcgctgtgtgtggggacaacgcggcctgccaaactacggcgtgcgcacctgtgaggg
 ctgcaaaggcttctttaagcgcacagtgcaaaaaaatgcaaaatacgtgtgttagcaaataaaa
 actgcccagtgagacaagcgtcgccggaatcgctgtcagtagtccgatttcagaagtgccctggct
 5 gttgggatgggtcaaagaagtgggttcgcacagacaggtttaaaggccggagaggtcggttgcctc
 gaaaccgaagagcccacaggagccctctcccccttcgcccccggtgagtctgatcagtgccctcg
 tcaggggcccatgtcgactccaaccggctatgaccagcctggactattccagggtccaggcgaac
 cctgactatcaaatgagtgagatgacaccagcatatccagcaattctatgatctcctgactgg
 ctccatggagatcatccggggctgggcagagaagatccctggcttcgcagacctgcccgaagccg
 10 accaagacctgctttttgaatcagcttttctagaactgtttgtccttcgattagcatacagggtcc
 aaccagtgaggaggtaaactcatcttttgcaatgggggtggctcttgacaggttgcaatgcgttcg
 tggctttggggaatggattgattccattgttgaaattctcctccaacttgacagaatatgaacatcg
 acatttctgccttctcctgcattgctgccttggtatggtcacagagagacacgggctcaaggaa
 cccaagagagtggaagaactgcaaaacaagattgtaaattgtctcaaagaccacgtgactttcaa
 15 caatgggggggtgaaccgcccccaattatttgtccaaactgttggggaagctcccagaacttcgta
 ccctttgcacacaggggctacagcgcattttctacctgaaattggaagacttggtgccaccgcca
 gcaataattgacaaacttttcttgacactttacctttctaa (SEQ ID NO:95)

[00384] SEQ ID NO:32: Homo sapiens neurogenin 2 (NEUROG2) (NM_024019) SEQ ID NO:32 (protein) and SEQ ID NO:96 (DNA)

MFVKSETLELKEEEDVLVLLGSASPALAALTPLSSSADEEEEEEPGASGGARRQRGAEGQGARG
 GVAAGAEGCRPARLLGLVHDCRRPSRARAVSRGAKTAETVQRIKKTRRLKANNRERNRMHNLNA
 ALDALREVLPTFPEDAKLTKIETLRFHNYIWALTETLRLADHCGGGGGGLPGALFSEAVLLSPG
 GASAALSSSGDSPASTWSCTNSPAPSSSVSSNSTSPYSCTLSPASPAGSDMDYWQPPPPDKHR
 YAPHLPIARDCI (SEQ ID NO:32)

atgttcgtcaaattccgagaccttgaggtgaaggaggaagaggacgtgttagtgctgctcggatc
 ggcctcccccgcttggcggccctgaccccgctgtcatccagcgccgacgaagaagaggaggagg
 agccgggcgcgctcaggcgggcgcgctcggcagcgcggggctgaggccgggcagggggcgcggggc
 ggcgtggctgcgggtgcggagggtgcggccccgcacggctgctgggtctggtacacgattgcaa
 acggcgcccttcccgggcgcgggcgctctcccagggcgccaagacggccgagacgggtgcagcgca
 30 tcaagaagaccctgtagactgaaggccaacaaccgcgagcgaaaccgcatgcacaacctcaacgcg
 gcactggacgcgctgcgcgaggtgctccccacgttccccgaggacgccaagctcaccaagatcga
 gacctgcgcttcgcccacaactacatctgggcactcaccgagacctgcgcctggcggatcact
 gcggggggcgcgggggggcctgcggggggcgctcttctcagaggcagtgttgctgagcccgga
 ggagccagcgccgctgagcagcagcgagacagccctcgcccgctccacgtggagttgcac
 35 caacagccccgcgcgctcctcctcctgctgctcctcaattccacctccccctacagctgcactttat
 cgcccgccagcccgccgggtcagacatggactattggcagccccacctcccgcacaagcaccgc
 tatgcacctcacctccccatagccagggttgtagtag (SEQ ID NO:96)

[00385] Homo sapiens LIM homeobox transcription factor 1, alpha (LMX1A) (NM_177398) SEQ ID NO:33 (protein) and SEQ ID NO:97 (DNA)

MLDGLKMEENFQSAIDTSASFSSLLGRAVSPKSVCEGCQRVILDRFLLRLNDSFWHEQCVQCASC
 KEPLETTCFYRDKKLYCKYDYEKLFVAVKCGGCFEAIAPNEFVMRAQKSVYHLSCFCCVCERQLQ
 KGDEFVLKEGQLLCKGDYKERELLSLVSPAASDSGKSDDEESLCKSAHGAGKGTAEEGKDHKRP
 KRPRITLTQQORAFKASFEVSSKPCRKVRETAAETGLSVRVVQVWFQNRQAKMKKLARRQQQQ
 QQDQQNTQRLSSAQTNNGGSAGMEGIMNPYTALPTPQQLLAIEQSVYSSDPFRQGLTPPQMPGDH
 45 MHPYGAEPLFHDLDSDDTSLNLGDCFLATSEAGPLQSRVGNPIDHLYSMQNSYFTS (SEQ
 ID NO:33)

atgctggacggcctaagatggaggagaacttccaaagcgcgatcgacacctcggcctccttctc
 ctgcgtgctgggcagagcggtgagccccaagtctgtctgcgagggctgtcagcgggctcatcttg
 acaggtttctgctgcggctcaacgacagcttctggcatgagcagtgctgcagtgcgctcctgc
 50 aaagagccccctggagaccacctgcttctaccgggacaagaagctgtactgcaagtatgactacga

gaagctgtttgctgttaaagtgtgggggctgcttcgagggccatcgctcccaatgagtttggtatgc
 gggcccagaagagtgatataccacctgagctgcttctgctgctgtgtctgagcgacagcttcag
 aagggtgatgagtttgctcctgaaggagggcagctgctctgcaaaggggactatgagaaggagcg
 ggagctgctcagcctgggtgagcccagcagcctcagactcaggtaaaagtgatgatgaagaaagtc
 5 tctgcaagtcagcccatggggcaggggaaaggaactgctgaggaaggcaaggaccataagcgcccc
 aaacgtccgagaaccatcttgacaactcaacagagggcgagcattcaaggcctcatttgaaagtatc
 ctccaagccctgcaggaaggtgagagagactctggctgcagagacagggctgagtgtccgtgtcg
 tccaggtgtggttccaaaaccagagagcggaagatgaagaagctggccaggcgacagcagcagcag
 cagcaagatcagcagaacacccagaggctgagctctgctcagacaaacgggtggaggagtgctgg
 10 gatggaaggaatcatgaaccctacacggctctgcccacccacagcagctcctggccatcgagc
 agagtgtctacagctcagatcccttccgacagggctcaccacccacccagatgcttgagaccac
 atgcacccttatggtgcccagcccccttttccatgacctggatagcgacgacacctccctcagtaa
 cctgggtgattgtttcctagcaacctcagaagctgggcctctgcagtccagagtgggaaacccca
 ttgaccatctgtactccatgcagaattcttacttcacatcttga (SEQ ID NO:97)

15 **[00386]** Homo sapiens HB9 homeobox gene, (U07663) SEQ ID NO:34 (protein) and SEQ ID NO:98 (DNA)

MEKSKNFRIEPCWRWTPHEPPLAERALAKVTSPPVPASGTGGGGGGGGASGGTSGSCSPASSEPP
 AAPADRLRAESPPRLAAHCALLPKPGFLGAGGGGGGTGGGHGGPHHHAHPGAAAAAAAAAAAA
 AAAAAGGLALGLHPGGAQGGAGLPAQAALYGHPVYGYSAALAGQHPALSYSPQVQGAHP
 20 AHPADPIKLGAGTFQLDQWLRATAGMILPKMPDFNSQAQSNLLGKCRRPRTAFTSQQLLELEHQF
 KFNKYLSRPKRFEVATSLMLTETQVKIWFQNRMKWKRSKKAKEQAAQEAQKQKGGGGGAGKGG
 EEPGAEELLGPPAPRDKSGPPADLRDSDPEEDEDEDEDHFPYSNGASVHAASSDCSSEDDSP
 PRPSHQAPQ (SEQ ID NO:34)

atggaaaaatccaaaatttccgcacatcgagccctgctggcgggtggacccccacgagccgcctct
 25 cgcagagcgcgctggccaaggtcacgtcgccgcccgtgcccgcacatctggcaccggaggtggcg
 gcgcgccgcgccggggcgagcgccgggactagcgccagctgcagccccgcgtcctcgagccgcgcg
 gctgcgccccgcgacccgcctgcgccgcgagagcccgctcgccgcgcgcctgctggccgcgcactg
 cgcgctgctgcccgaagccgggcttccctgggcgcggggcgccggcgccggcgccggcgccggcg
 acgggggggccccaccaccacgcgcacatccgggcgcagcgcccgctgcccgcgcgcgcgcgcgc
 30 gccgcgcgcgcgcgcctgggggcctggcgctggggctgcaccctgggggcgcgcagggcgccgcggg
 cctcccggcgccagggcgctctacggccacccgggtctacggctactccgcggcgccggcgccggg
 ctgctgctggcgggccagcaccggcgctctcctactcgtaccgcaggtgcaaggcgccgacccc
 ggcacccccgcgaccccatcaagctgggcgcgcggcaccttccagctggaccagtggctgcgcg
 gtccaccgcgggcatgatcctgctaagatgcccgaacttcaact (SEQ ID NO:98)

35 **[00387]** Homo sapiens LIM homeobox 3 (LHX3) (NM_178138) SEQ ID NO:35 (protein) and SEQ ID NO:99 (DNA)

MLLETGLERDRARPGAAAVCTLGGTREIPLCAGCDQHILDRFILKALDRHWHSKCLKCSDCHTPL
 AERCFSRGESVYCKDDFFKRFGTKAACQLGIPPTQVVRRAQDFVYHLHCFACVVCKRQLATGDE
 FYLMEDSRLVCKADYETAKQREAEATAKRPTTITAKQLETLSAYNTSPKPARHVREQLSSETG
 40 LDMRVVQVWFQNRRAKEKRLKKDAGRQRWGQYFRNMKRSRGGSKSDKDSVQEGQDSDAEVSFPDE
 PSLAEMGPANGLYGSLGEPTQALGRPSGALGNFSLHGGLAGPEQYRELRPGPSYGVPPSPAAPQ
 SLPGPQPLLSSLVYPDTSLGLVPSGAPGGPPPMRVLGNGPSSDLSTGSSGGYPDFPASPASWLD
 EVDHAQF (SEQ ID NO:35)

atgctgctggaaacggggctcgagcgcgaccgagcgagggcccgggcgccgcgcctctgcacctt
 45 gggcgggactcgggagatcccgcctgtgctgctggctgtgaccagcacatcctggaccgcttcatcc
 tcaaggctctggaccgccactggcacagcaagtgtctcaagtgcagcgactgccacacgccactg
 gccgagcgctgcttcagccgaggggagagcgcttactgcaaggacgactttttcaagcgcttcgg
 gaccaagtgcgcgcgcgtgccagctgggcaccccgccacgcaggtgggtgcgcgcgcgcgcaggact
 tcgtgtaccacctgcaactgctttgctgctgctgtgcaagcgccagctggccacgggcgacgag
 50 ttctacctcatggaggacagccggctcgtgtgcaaggcggaactacgaaaccgccaagcagcgaga

ggccgagggccacggccaagcgggcgcgacgaccatcaccgccaagcagctggagacgctgaaga
 gcgcttacaacacctcgcccaagccggcgcgccacgtgcgcgagcagctctcgccgagacgggc
 ctggacatgcgcgtggtgcaggtttggttcagaaccgcccgggccaaggagaagaggtgaagaa
 ggacgcccggccggcagcgtgggggcagtatcttcgcaacatgaagcgtcccgcgggcggtcca
 5 agtcggacaaggacagcgttcaggaggggaggacagcagcgtgaggtctccttccccgatgag
 ccttccttggcggaatgggcccggccaatggcctctacgggagcttgggggaacccaccagggc
 cttgggcccggccctcgggagccctgggcaacttctccctggagcatggaggcctggcaggccag
 agcagtaccgagagctgcgtcccggcagccctacgggtgtcccccatcccccgccgccccgcag
 agcctccctggccccagccctcctctccagcctggtgtaccagacaccagcttgggccttgt
 10 gccctcgggagccccggcgggccccaccatgaggggtgctggcagggaaacggaccagttctg
 acctatccacggggagcagcgggggttaccgcacttccctgccagccccgcctcctggtggtg
 gaggtagaccacgctcagttctga (SEQ ID NO:99)

[00388] Homo sapiens distal-less homeobox 2 (DLX2) (NM_004405) SEQ ID NO:36 (protein)
 and SEQ ID NO:100 (DNA)

15 MTGVFDSLVDLMHSTQIAASSTYHQHQPPSGGGAGPGGNSSSSSLHKPQESPTLPVSTATDSS
 YYTNQQHPAGGGGGGSPYAHMGSYQYQASGLNNVPYSAKSSYDLGYTAAYTSYAPYGTSSSPAN
 NEPEKEDLEPEIRIVNGKPKKVRKPRTIYSSFQLAALQRRFQKTQYLALPERAELAASLGLTQTQ
 VKIWFQNRRSKFKMKWKSGEIPSEQHPGASAPPCASPPVSAPASWDFGVPQRMAGGGGPGSGGS
 GAGSSGSSPSSAASFLGNYPWYHQTSGSASHLQATAPLLHPTQTPQPHHHHHHGGGGAPVSAG
 20 TIF (SEQ ID NO:36)

atgactggagtcctttgacagtcctagtggctgatatgcactcgaccagatcgccgcctccagcac
 gtaccaccagcaccagcagccccgagcggcgggcgccgcccgggtggcaacagcagcagca
 gcagcagcctccacaagccccaggagtcgcccacccttccgggtgtccaccgccaccgacagcagc
 tactacaccaaccagcagcaccggcgggcgggcgggcgggcgggggtcgccctacgcgcacat
 25 ggggttcctaccagtaccaagccagcggcctcaacaacgtcccttactccgccaagagcagctatg
 acctgggctacaccgcgcctacacctcctacgctccctatggaaccagttcgtccccagccaac
 aacgagcctgagaaggaggaccttgagcctgaaattcgatagtgaaagggaagccaaagaaagt
 ccggaaacccccgcaccatctactccagtttccagctggcggtccttcagcggcggtttccaaaaga
 ctcaatacttggccttgccggagcagcagccgagctggcggcctctctgggcctcaccagactcag
 30 gtcaaaatctggttccagaaccgcgggtccaagttcaagaagatgtggaagagtggtgagatccc
 ctcgagcagcaccctggggccagcgttctccaacctgtgcttcgcccgcagctctcagcgccgg
 cctcctgggactttggtgtgccgcagcggatggcgggcgggcggtggtccgggcagtggcggcagc
 ggcgccggcagctcgggctccagcccagcagcgcggcctcggttttctgggcaactaccctg
 gtaccaccagacctcgggatccgcctcacacctgcaggccacggcgccgctgctgcacccccactc
 35 agacccccgcagccgcacaccaccaccaccatcacggcgggggggcgccccgggtgagcgcgggg
 acgattttctaa (SEQ ID NO:100)

[00389] Homo sapiens runt-related transcription factor 2 (RUNX2)(NM_001024630) SEQ ID
 NO:37 (protein) and SEQ ID NO:101 (DNA)

40 MASNSLFSTVTPCQQNFFWDPSTSRRFSPSSSLQPGKMSDVSPVAAQQQQQQQQQQQQQQQQ
 QQQQQQEAAAAAAAAAAAAAAAAAAVPRLRPPHDNRTMVEIIADHPAELVRTDSPNFLCSVLPSHW
 RCNKTLPVAFKVVALGEVPDGTVVTVMAGNDENYSAELRNASAVMKNQVARFNDLRFVGRSGRGK
 SFTLTITVFTNPPQVATYHRAIKVTVDGPREPRRHRQKLDDSKPSLFSDRLSDLGRIPHPSMRVG
 VPPQNPRPSLNSAPSPFNPGQSQITDPRQAQSSPPWSYDQSYPSYLSQMTSPSIHSTTPLSSTR
 GTGLPAITDVPERRISDDDTATSDFCLWPSTLSKKSQAGASELGPFS DPRQFPSSISLTESRFSNP
 45 RMHYPATFTYTPPVTSGLMSLGMSATTHYHTYLPFPYPGSSQSQSGPFQTSSTPYLYYGTSSGSYQ
 FPMVPGGDRSPSRMLPPCTTTSNGSTLLNPNLPNQNDGVDADGSHSSSPTVLNSSGRMDES VWRP
 Y (SEQ ID NO:37)

atggcatcaaacagcctcttcagcacagtgacaccatgtcagcaaaacttcttttgggatccgag
 caccagccggcgcttcagccccccctccagcagcctgcagcccggcaaaatgagcgacgtgagcc
 50 cgggtggtggctgcgcaacagcagcagcaacagcagcagcagcaacagcagcagcagcagcagcaa

cagcagcagcagcagcaggaggcgggcgggcgggcggtgcgggcgggcgggcgggcggtgcgggcgggcgggc
 agctgcagtgccccggttgcgggcgccccacgacaaccgcaccatgggtggagatcatcgccgacc
 acccgccgaactcgtcgcaccgacagccccaacttctgtgctcggtgctgccctcgactgg
 5 cgctgcaacaagaccctgcccggtggccttcaagggtggtagccctcgagaggtaccagatgggac
 tgtggttactgtcatggcggttaacgatgaaaattattctgctgagctccggaatgcctctgctg
 ttatgaaaaaccaagtagcaaggttcaacgatctgagatttgtgggcccggagtggacgaggcaag
 agtttcaccttgaccataaccgtcttcacaaatcctcccaagtagctacctatcacagagcaat
 taaagttacagtagatggacctcggaacccagaaggcacagacagaagcttgatgactctaaac
 ctagtgttctctgaccgcctcagtgatttagggcgcatcctcatcccagtatgagagtaggt
 10 gtcccgccctcagaacccacggccctccctgaactctgcaccaagtcttttaataccacaaggaca
 gagtacagattacagacccaggcaggcacagtccttccccgcggtggtcctatgaccagtcttacc
 cctcctacctgagccagatgacgtccccgtccatccactctaccaccccgctgtcttccacacgg
 ggcaactgggcttctgccatcacgatgtgccttagggcgcatcttcagatgatgacactgccacctc
 tgacttctgcctctggccttccactctcagtaagaagagccaggcaggtgcttcagaactgggccc
 15 ctttttcagaccccaggcagttcccaagcatttcatccctcactgagagccgcttctccaaccca
 cgaatgcactatccagccacctttacttacccccgccagtcacctcaggcatgtccctcggtat
 gtccgccaccactcactaccacacctacctgccaccacctaccccggtcttcccaaaggccaga
 gtggacccttccagaccagcagcactccatatctctactatggcacttcgtcaggatcctatcag
 tttcccatgggtgccggggggagaccgggtctccttccagaatgcttccgccatgcaccaccacctc
 20 gaatggcagcagctattaaatccaaatttgccataaccagaatgatgggtgttgacgtgatggaa
 gccacagcagttccccaactgttttgaattctagtggcagaatggatgaatctgtttggcgacca
 tattga (SEQ ID NO:101)

[00390] Homo sapiens SRY (sex determining region Y)-box 5 (SOX5) (NM_006940) SEQ ID NO:38 (protein) and SEQ ID NO:102 (DNA)

25 MLTDPDLPQEFERMSSKRPA SPYGEADGEVAMVTSRQKVEEEESDGLPAFHLPLHVSFPNKP HSE
 EFQPVSLLTQETCGHRTPT SQHNTMEVDGNKVMSSFAPHNSSTSPQKAEEGGRQSGESLSSTALG
 TPERRKGLSLADVDTL KQRKMEELIKNEPEETPSIEKLLSKDWKDKLLAMGSGNFGEIKGTPE SL
 AEKERQLMGMINQLTSL REQLLAHDEQKKLAASQIEKQRQOMELAKQQQE QIARQQQQLLQQQH
 30 KINLLQQQIQVQGLPPLMIPVFPDQRTLAAAQQGFLLPFGFSYKAGCSDPYVPQLIPTTMAA
 AAAATPGLGLQLQLQLYAAQLAAMQVSPGGKLP GIPQGNLGA AVSPTS IHTDKSTNSPPPKSKDE
 VAQPLNLSAKPKTSDGKSPTSPTSPHPALRINS GAGPLKASVPAALASPSARVSTIGYLNHDHA
 VTKAIQEARQMKEQLRREQQLD GKVAVVNSLGLN NCRTEKEKTTLESLTQQLAVKQNEEGKF SH
 AMMDFNLSGDSGDSAGVSE SRIYRESRGRGSNEPHIKRPMNAFMVWAKDERRKILQAFPMHNSN
 35 ISKILGSRWKAMTNLEKQPYEEQARLSKQHLEKYPDYKYKPRPKRTCLVDGKKLRIGEYKAIMR
 NRRQEMRQYFNVGQQAQIPIATAGVVYPGAIAMAGMPSPHLPSEHSSVSSSPEPGMPVIQSTYGV
 KGEEPHIKEEIQAE DINGEIYDEYDEEEDDPD VDYGSDSENHIAGQAN (SEQ ID NO:38)
 atgcttactgaccctgatttacctcaggagtttgaaaggatgtcttccaagcgaccagcctctcc
 gtatggggaagcagatggagaggtagccatgggtgacaagcagacagaaagtggagaagaggaga
 40 gtgacggggtcccagcctttcaccttcccttgcatgtgagttttcccaacaagcctcactctgag
 gaatttcagccagtttctctgctgacgcaagagacttgtggccataggactcccacttctcagca
 caatacaatggaagttgatggcaataaagttatgtcttcatttgccccacacaactcatctacct
 cacctcagaaggcagaagaaggtgggacgacagagtggcgagtccttgtctagtagcgcctggga
 actcctgaacggcgcaagggcagtttagctgatgttggttgacaccttgaagcagaggaaaatgga
 agagctcatcaaaaacgagccggaagaaacccccag tattgaaaaactactctcaaggactgga
 45 aagacaagcttcttgcaatgggatcggggaactttggcgaaataaaagggactcccgagagctta
 gctgagaaagaaaggcaactcatgggtatgatcaaccagctgaccagcctccgagagcagctgtt
 ggctgccacgatgagcagaagaactagctgcctctcagattgagaaacagcgtcagcaaatgg
 agctggccaagcagcaacaagaacaaattgcaagacagcagcagcagcttctacagcaacaacac
 aaaatcaatttgctccagcaacagatccaggttcaaggtcagctgccgccattaatgattcccgt
 50 attccctcctgatcaacggacactggctgcagctgccagcaaggattcctcctccctccaggct

tcagctataaggctggatgtagtgacccttacctgttcagctgatcccaactaccatggcagct
 gctgccgcagcaacaccaggcttaggccactccaactgcagcagttatatgctgccagctagc
 tgcaatgcaggtatctccaggaggaagctgccaggcataccccaaggcaaccttggtgctgctg
 tatctcctaccagcattcacacagacaagagcacaaacagcccaccacccaaaagcaaggatgaa
 5 gtggcacagccactgaacctatcagctaaacccaagacctctgatggcaaataccccacatcacc
 cacctctccccatatgccagctctgagaataaacagtggggcaggccccctcaaagcctctgtcc
 cagcagcgtagtagtccttcagccagagtagcacaatagggtacttaaatgaccatgatgct
 gtcaccaaggcaatccaagaagctcggcaaatgaaggagcaactccgacgggaacaacagggtgct
 tgatgggaaggtggctggtgtgaatagtctgggtctcaataactgccgaacagaaaaaggaaaaa
 10 caaactggagagtctgactcagcaactggcagttaaacagaatgaagaaggaaaatttagccat
 gcaatgatggatttcaatctgagtggagattctgatggaagtgctggagtctcagagtcaagaat
 ttatagggaaatcccagggcggttagcaatgaacccccacataaagcgtccaatgaatgccttca
 tgggtgtgggctaaagatgaacggagaaagatccttcaagcctttcctgacatgcacaactccaac
 atcagcaagatatgggatctcgctggaaagctatgacaaacctagagaaaacagccatattatga
 15 ggagcaagcccgtctcagcaagcagcacctggagaagtaccctgactataagtacaagcccaggc
 caaagcgcacctgcctggtggatggcaaaaagctgcgcatgtgtgaatacaaggcaatcatgcgc
 aacaggcggcaggaaatgcggcagtagtctcaatgttgggcaacaagcacagatccccattgccac
 tgctggtgtgtgtaccctggagccatcgccatggctgggatgccctccccctcacctgccctcgg
 agcactcaagcgtgtctagcagcccagagcctgggatgcctgttatccagagcacttacggtgtg
 20 aaaggagaggagccacatatcaaagaagagatacaggccgaggacatcaatggagaaatttatga
 tgagtacgacgaggaagaggatgatccagatgtagattatgggagtgacagtgaaaaccatattg
 caggacaagccaactga (SEQ ID NO:102)

[00391] Homo sapiens SOX6 mRNA (AF309034) SEQ ID NO:39 (protein) and SEQ ID NO:103 (DNA)

MSSKQATSPFACAADGEDAMTQDLTSREKEEGSDQHVASHLPLHPIMHNKPHSEELPTLVSTIQQ
 DADWDSVLSSQQRMESENKLCSLYSFRNTSTSPHKPDEGSRDREIMTSVTFGTPERRKGLADV
 VDTLKQKKLEEMTRTEQEDSSCMEKLLSKDWKEKMERLNTSELLGEIKGTPESLAEKERQLSTMI
 TQLISLREQLLAHDEQKKLAASQIEKQRQQMDLARQQEQIARQQQQLLQQQHKKINLLQQQIQV
 QGHMPPLMIPIFPHDQRTLAAAAAAQQGFLFPPGITYKPGDNPVQFIPSTMAAAAASGLSPLQL
 30 QKGHASHPQINQRLKGLSDRFGRNLDTFEHGGGHSYNHKQIEQLYAAQLASMQVSPGAKMPSTPQ
 PPNTAGTVSPTGIKNEKRGTSPTVQVKDEAAAQPLNLSSRPKTAEPVKSPTSPTQNLFPASKTSP
 VNLPNKSSIPSPIGGSLGRGSSLDILSSLNSPALFGDQDTVMKAIQEARKMREIQIREQQQQQPH
 GVDGKLSSINNMGLNSCRNEKERTRFENLGPQLTGKSNEGDGKLGPGVIDLTRPEDAEGGATVAEA
 RVYRDARGRASSEPHIKRPMNAFMVWAKDERRKILQAFPMHNSNISKILGSRWKSMSNQEKQPY
 35 YEEQARLSKIHLEKYPNYKYKPRPKRTCIVDGKKLRIGEYKQLMRSRRQEMRQFFTVGQQPQIPI
 TTGTGVVYPGAITMATTTTSPQMTSDCSSTSASPEPSLPVIQSTYGMKTDGGSLAGNEMINGEDE
 MEMYDDYEDDPKSDYSSENEAPEAVSAN (SEQ ID NO:39)

atgtcttccaagcaagccacctctccatttgccctgtgcagctgatggagaggatgcaatgaccca
 ggatttaacctcaagggaaggaagaggcagtgatcaacatgtggcctcccatctgcctctgc
 40 accccataatgcacaacaaacctcactctgaggagctaccaacacttgctcagtaccattcaacaa
 gatgctgactgggacagcgttctgtcatctcagcaaagaatggaatcagagaataataagttatg
 ttccctatatcttccgaaatacctctacctcaccacataagcctgacgaaggagtcgggacc
 gtgagataatgaccagtgttacttttggaaacccagagcgccgcaaaggagtccttgccgatgtg
 gtggacacactgaaacagaagaagcttgaggaaatgactcggactgaacaagaggattcctcctg
 45 catggaaaaactactttcaaaagattggaaggaaaaaatggaaagactaaataccagtgaacttc
 ttggagaaattaaaggtacacctgagagcctggcagaaaaagaacggcagctctccaccatgatt
 acccagctgatcagtttacgggagcagctactggcagcgcatgatgaacagaaaaaactggcagc
 gtcacaaattgagaacaaacggcagcaaatggaccttgctcgccaacagcaagaacagattgcga
 gacaacagcagcaacttctgcaacagcagcacaaaattaatctcctgcagcaacagatccaggtt
 50 caggggtcacatgcctccgctcatgatcccaatttttccacatgaccagcggactctggcagcagc

tgctgctgcccacagggattcctcttccccctggaataacatacaaaaccaggtgataactacc
 ccgtacagttcattccatcaacaatggcagctgctgctgcttctggactcagccctttacagctc
 cagaaggggtcatgcctcccacccacaaattaaccaaaggctaaagggcctaagtgaacggttttgg
 caggaatttggacacctttgaacatggtggtggccactcttacaaccacaaacagattgagcagc
 5 tctatgccgctcagctggccagcatgcaggtgtcacctggagcaaagatgccatcaactccacag
 ccaccaaacacagcagggacggtctcacctactgggataaaaaatgaaaagagaggggaccagccc
 tgtaactcaagttaaggatgaagcagcagcagcagcctctgaatctctcatcccagaccaagacag
 cagagcctgtaaagtccccaacgtctcccaccagaacctcttcccagccagcaaaaccagccct
 gtcaatctgcccacaaaagcagcatccctagccccattggaggaagcctgggaagaggatcctc
 10 tttagatatcctatctagtctcaactcccctgccctttttggggatcaggatacagtgatgaaag
 ccattcaggagggcggaagatgcgagagcagatccagcgggagcaacagcagcaacagccacat
 ggtgttgacgggaaactgtcctccataaataatattggggctgaacagctgcaggaatgaaaagga
 aagaacgcgctttgagaatttggggccccagttaacgggaaagtcaaatgaagatggaaaactgg
 gccaggtgtcatcgaccttactcggccagaagatgcagagggaggtgccactgtgggtgaagca
 15 cgagtctacagggacgcccgcggccgtgccagcagcagccacacattaagcgaccaatgaatgc
 attcatgggttgggcaaaggatgagaggagaaaaatccttcaggccttccccgacatgcataact
 ccaacattagcaaaatcttaggatctcgctggaaatcaatgtccaaccaggagaagcaaccttat
 tatgaagagcaggccccggctaagcaagatccacttagagaagtacccaaactataaatacaaac
 ccgaccgaaacgcacctgcattgttgatggcaaaaagcttcggattggggagtataagcaactga
 20 tgaggtctcgggagacaggagatgaggcagttctttactgtggggcaacagcctcagattccaatc
 accacaggaacaggtgttggtgtatcctggtgctatcactatggcaactaccacaccatcgctca
 gatgacatctgactgctctagcacctcggccagccccggagcccagcctcccgggtcatccagagca
 cttatggtatgaagacagatggcggagccttagctggaaatgaaatgatcaatggagaggatgaa
 atggaaatgtatgatgactatgaagatgaccccaaatacagactatagcagtgaaaatgaagcccc
 25 ggaggtgtcagtgccaactga (SEQ ID NO:103)
[00392] Homo sapiens GATA binding protein 6 (GATA6) (NM_005257) SEQ ID NO:40
 (protein) and SEQ ID NO:104 (DNA)
 MALTDGGWCLPKRFGAAGADASDSRAFPAREPSTPPSPISSSSSSSCSRGGGERGPGGASNCGTPQL
 DTEAAAGPPARSLLLSSYASHPFPGAPHGPSAPGVAGPGGNLSSWEDLLLFTDLDQAATASKLLWS
 30 SRGAKLSPFAPEQPEEMYQTLAALSSQGPAAAYDGAPGGFVHSAAAAAAAAAAASSPVYVPTTRVG
 SMLPGLPYHLQSGSGSPANHAGGAGAHPGWPQASADSPPYGSGGGAAGGGAAGPGGAGSAAAHVS
 ARFPYSPSPPMANGAAREPGGYAAAGSGGAGGVSGGSSLAAMGGREPQYSSLSAARPLNGTYHH
 HHHHHHHHPSPYSPYVGAPLTPAWPAGPFETPVLHSLQSRAGAPLPVPRGPSADLLEDLSESR
 VNCGSIQTPLWRRDGTGHYLCNACGLYSKMNGLSRPLIKPQKRVPSRRLGLSCANCHTTTTTLW
 35 RRNAEGEPVCNACGLYMKLHGVPRPLAMKKEGIQTRKRKPKNINKSKTCSGNSNNSIPMTPTSTS
 SNSDDCSKNTSPTTQPTASGAGAPVMTGAGESTNPENSELKYSQDGLYIGVSLASPAEVTSSVR
 PDSWCALALA (SEQ ID NO:40)
 atggccttgactgacggcggtggtgcttgccgaagcgcttcggggccgcgggtgcggacgcccag
 cgactccagagcctttccagcgcgggagccctccacgcgccttcccccatctcttcctcgctcct
 40 cctcctgctcccggggcgagagcggggccccggcgcgccagcaactgcgggacgcctcagctc
 gacacggaggcgggcgggcggacccccggcccgtcgctgctgctcagttcctacgcttcgcaccc
 cttcgggggtccccacggaccttcggcgcttgggggtcgcgggccccgggggcaacctgtcgagct
 gggaggacttgctgctgttcactgacctcgaccaagccgcgacccgagcaagctgctgtggtcc
 agccgcggcgccaagctgagcccccttcgcacccgagcagccggaggagatgtaccagacctcgc
 45 cgctctctccagccaggggtccggccgcctacgacggcgcgccccggcggttcgtgcaactctgcgg
 ccgcggcggcagcagccgcggcgggccagctccccggtctacgtgccaccaccgcgctgggt
 tccatgctgcccggtaccgtaccacctgcaggggtcgggcagtgggccagccaaccacgcggg
 cggcgcgggcgcgcacccccggtggcctcaggcctcgggcgacagccctccatacggcagcgag
 gcggcgcgggtggcgggcgggcggggctggcggcgctggctcagccgcggcgacgctctcg
 50 gcgcgcttcccctactctcccagcccccatggccaacggcgccgcggggagccgggaggcta

10

15

20

25

30

35

40

50

DPYQILGPTSSRLANPGSGQIQLWQFLELLSDSANASCITWEGTNGEFKMTDPDEVARRWGERK
SKPNMNYDKLSRALRYYYDKNIMTKVHGKRYAYKFDFHGIAQALQPHPTESMYKYPSDISYMPS
YHAHQKQVNFVPPHPSSMPVTSSSFFGAASQYWTSTGGIYPNPVPRHPNTHVPSHLGSYY
(SEQ ID NO:42)

5 atggacgggactattaaggaggctctgtcgggtggtgagcgacgaccagtcctctttgactcagc
gtacggagcggcagcccatctccccaaggccgacatgactgcctcggggagtcctgactacgggc
agccccacaagatcaacccccctcccaccacagcaggagtggatcaatcagccagtgagggtcaac
gtcaagcgggagtatgaccacatgaatggatccaggagtcctccgggtggactgcagcgttagcaa
atgcagcaagctggtgggcggaggcgagtcacaaccccatgaactacaacagctatatggacgaga
10 agaatggccccctcctcccaacatgaccaccaacgagaggagagtcctcgtccccgcagacccc
aactgtggacacaggagcatgtgaggcaatggctggagtggggccataaaggagtacagcttgat
ggagatcgacacatcctttttccagaacatggatggcaaggaaactgtgtaaaatgaacaaggagg
acttcctccgcgccaccaccctctacaacacggaagtgtgttgtcacacctcagttacctcagg
gaaagttcactgctggcctataatacaacctcccacaccgaccaatcctcacgattgagtgtcaa
15 agaagacccttcttatgactcagtcagaagaggagcttggggcaataacatgaattctggcctca
acaaaagtctcctcccttggaggggcacaaacgatcagtaagaatacagagcaacggccccagcca
gatccgtatcagatcctggggcccgaccagcagtcgcctagccaaccctggaagcgggcagatcca
gctgtggcaattcctcctggagctgctctccgacagcgccaacgccagctgtatcacctgggagg
ggaccaacggggagttcaaaatgaaggaccccgatgaggtggccaggcgctggggcgagcggaaa
20 agcaagcccaacatgaattacgacaagctgagccgggccctccgttattactatgataaaaacat
tatgaccaaagtgcacggcaaaagatatgcttacaaatttgacttccacggcattgcccaggctc
tgcagccacatccgaccgagtcgtccatgtacaagtaccccttctgacatctcctacatgccttcc
taccatgcccaccagcagaaggtgaactttgtccctccccatccatcctccatgcctgtcacttc
ctccagcttctttggagccgcatcacaatactggacctccccacggggggaatctaccccaacc
25 ccaacgtcccccgccatcctaacacccacgtgccttcacacttaggcagctactactag (SEQ
ID NO:106)

[00395] Human erythroid Kruppel-like factor EKLF gene (U37106) SEQ ID NO:43 (protein)
and SEQ ID NO:107 (DNA)

MATAETALPSISTLTALGFPDQTDDFLKWWRSEEAQDMGPGPPDPTEPPLHVKSEDQPGEEEDD
30 ERGADATWDLDLLLTNFSGPEPGGAPQTCALAPSEASGAQYPPPPETLGAYAGGPGVLVAGLLGSE
DHSGWVRPALRARAPDAFVGPALAPAPAPEPKALALQPVYPGPGAGSSGGYFPRTGLSVPAASGA
PYGLLSGYPAMYAPQYQGHFQLFRGLQGPAPGPATSPSFLSCLGPGTVGTGLGTTAEDPGVIAE
TAPSKRGRRSWARKRQAAHTCAHPGCGKSYTKSSHLKAHLRTHTEKPYACTWEGCGWRFARSDE
LTRHYRKHTGQRPFRQCPCPRAFSRSDHLALHMKRHL (SEQ ID NO:43)
35 atggccacagccgagaccgccttgccctccatcagcacactgaccgccttgggcccccttcccgga
cacacaggatgacttcctcaagtgggtggcgctccgaagaggcgaggacatgggccccgggtcctc
ctgacccacaggagccgccccctccacgtgaagtctgaggaccagcccggggagggaagaggacgat
gagaggggcgcgacgccacctgggacctggatctcctcctaccaacttctcggggcccgagcc
cgggtggcgcgccccagacctgcgctctggcgcccagcgaggcctccggggcgcaatatccgcccgc
40 cgcccgagactctgggcgcatatgctggcgcccggggctgggtggctgggcttttgggttcggag
gatcactcgggttggtgcgccctgccctgcgagcccgggctcccgaagccttcgtgggcccagc
cctggctccagccccggcccccgagcccaaggcgctggcgctgcaaccgggtgtaccggggggcccg
gcgccggctcctcgggtggctacttcccgcggaccgggctttcagtgccctgcggcgctcggggcgcc
ccctacgggctactgtccgggtacccccgcgatgtaccgggcgcctcagtaaccaagggaacttcca
45 gctcttccgcggggtccagggacccgcgcccgggtcccggccacgtccccctccttccctgagttgtt
tgggacccgggacgggtgggcactggactcggggggactgcagaggatccagggtgtgatagccgag
accgcgccatccaagcgaggccgacgttcgtgggcgcgcaagaggcaggcagcgcacacgtgcgc
gcacccgggttgcggaagagctacaccaagagctccacactgaaggcgcatctgcgcacgcaca
cagggggagaagccatacgcctgcacgtgggaaggctgcggctggagattcgcgcgctcggacgag

ctgacccgccactaccggaacacacggggcagcgcccccttcgctgccagctctgcccacgtgc
 tttttcgcgctctgaccacctggccttgacatgaagcgccacctttga (SEQ ID NO:107)

[00396] Human MyoD (X56677) SEQ ID NO:44 (protein) and SEQ ID NO:108 (DNA)

MELLSPPLRDVDLTAPDGSLSFATDDFYDDPCFDSPLRFFEDLDPRLMHVGALLKPEEHSF
 5 PAAVHPAPGAREDEHVRAPSGHHQAGRCLLWACKACKRKTTNADRRKAATMRERRRLSKVNEAFE
 TLKRCTSSNPQNQLPKVEILRNAIRYIEGLQALLRDQDAAPPGAAAFYAPGPLPPGRGGEHYS
 SDASSPRNSCDGMMDYSGPPSGARRRNCYEGAYYNEAPSEPRPGKSAAVSSLDYLSSIVERIST
 ESPAAPALLLADVPSESPRRQEAAAPSEGESSGDPTQSPDAAPQCPAGANPNPIYQVL (SEQ
 ID NO:44)

10 atggagctactgtcgccaccgctccgcgacgtagacctgacggcccccgacggctctctctgctc
 ctttgccacaacggagcacttctatgacgacctgtttcgactccccggacctgcgcttcttcg
 aagacctggacccgcgcctgatgcacgtgggcgcgctcctgaaacccgaagagcactcgcacttc
 cccgcggcggtgcacccggccccggcgcgacgtgaggacgagcatgtgcgcgcgcccagcgggca
 ccaccaggcgggcgctgcctactgtgggctgcaaggcgtgcaagcgcaagaccaccaacgccg
 15 accgcgcgaaggcgccaccatgcgcgagcggcgccgctgagcaaagtaaagtgaggcctttgag
 aactcaagcgtgcacgtcgagcaatcaaaccagcgggtgcccgaagtgagatcctgcgcaa
 cgccatccgctatatcgaggcctgcaggctctgctgcgcgaccaggacgcgcgcgccccctggcg
 cagccgccttctatgcgcggggcccgctgccccggggcgcgaggcagcactacagcggcgac
 tccgacgcgtccagccgcgctccaactgctccgacggcatgatggactacagcggccccccgag
 20 cggcgccccggcgcggaactgctacgaaggcgctactacaacgaggcgcccagcgaaccaggc
 ccgggaagagtgcggcggtgtcgagcctagactacctgtccagcatcgtggagcgcacatccacc
 gagagccctgcggcgcccgccctcctgctggcgagcgtgccttctgagtcgcctccgcgcaggca
 agaggctgccgccccagcgaggagagagcagcggcgacccccaccagtcaccggacgcgcgcc
 cgcagtgcctgcggtgcgaaccccaaccgatataaccaggtgctctga (SEQ ID
 NO:108)

**[00397] Homo sapiens NGFI-A binding protein 2 (EGR1 binding protein 2)(NAB2)
 (NM_005967) SEQ ID NO:45 (protein) and SEQ ID NO:109 (DNA)**

MHRAPSPTAEQPPGGGDSARRTLQPLKPSARAMALPRTLGLQLYRVLQRANLLSYETFIQQG
 GDDVQQLCEAGEEEFLEIMALVGMATKPLHVRRLQKALREWATNPGLFSQPVPAVPVSSIPLFKI
 30 SETAGTRKGSMSNGHSGPGEKAGSARSFSPKSPLELGEKLSPLPGGPGAGDPRIWPRSTPESDV
 GAGGEEEAGSPPFSPPAGGVPEGTGAGGLAAGGTGGGPDRLEPEMVRMVVESVERIFRSFPRGD
 AGEVTSLLKLNKKLARSVGHI FEMDDNDSQKEEIRKYSI IYGRFDSKRREGKQLSLHELTINEA
 AAQFCMRDNTLLLRVELFSLSRQVARESTYLSLKGSRHLHPEELGGPPLKKLKQEVGEQSHPEI
 QQPPPGPESYVPPYRPSLEEDSASLSGESLDGHLQAVGSCPRLTPPPADLPLALPAHGLWSRHIL
 35 QQTLMDDEGLRLARLVSHDRVGRSLSPCVPAKPPLAEFEGLLDRC PAPGPHPALVEGRSSVKVEA
 EASRQ (SEQ ID NO:45)

40 atgcacagagcgccttccccacagccgagcagcgcgcggcgaggggacagcgcgcccgccggac
 cctgcagcccagactcaagcccagtgcccagccatggcactgcctcggacgctgggggagctgc
 agctgtaccgggtcctgcagcgcgccaacctccttctactatgagaccttcacccagcaggga
 ggggacgacgtgcagcagctgtgtgaggcgggtgaggaggagtttctggagatcatggcacttgt
 gggcatggccaccaagccccctcatgtccggcgccctgcagaaggcactgagagagtgggccacca
 atccagggtcttcagtcaaccagtgccctgctgttcccgctctccagcatcccgtcttcaagatc
 tctgagactgcgggtacccggaaaggagcatgagcaatgggcatggcagcccaggggaaaaggc
 aggcagtgcgccgagttttagccccaagagcccccttgaaacttgagagaagctatcaccactgc
 45 ctgggggacctggggcaggggacccccgatctggccaggccggagcactccagagtcggacgtt
 ggggcaggaggagaagaggaggtggctgcgccccccttctccccccctgcagggggaggagtc
 tgaggggactggggctgggggctggcagcaggtgggactgggggtggtccagaccgactggagc
 cagagatggtacgcatggtggtggaaagtgtggagaggatcttccggagcttcccaaggggggat
 gctggggaggtcacatccctgctaaagctgaataagaagctggcacggagcgttgggcacatctt
 50 tgagatggatgataatgacagccagaaggaaggagatccgcaaatacagcatcatctatggcc

gtttcgactctaagcggcgggaggggcaagcagctcagcctgcacgagctcaccatcaacgaggct
 gctgcccagttctgcatgagggacaacacgctcttattacggagagtggagctcttctctttgtc
 ccgccaagtagccccgagagagcacctacttgtcctccttgaagggctccaggcttcaccctgaag
 aactgggaggccctccactgaagaagctgaaacaagaggttgagaaacagagtcaccctgaaatc
 5 cagcagcctccccaggccctgagtcctatgtacccccataccgccccagcctggaggaggacag
 cgccagcctgtctggggagagtctggatggacatttgcaggctgtggggctcatgtccaaggctga
 cgccgccccctgtgacctgcctctggcattgccagcccatgggctatggagccgacacatcctg
 cagcagacactgatggacgaggggctgcggctcgccgctcgtctcccacgaccgctgggccc
 cctcagccccctgtgtgcctgcgaagccacctctcgcagagttcgaggaagggctgctggacagat
 10 gtcctgccccaggacccccatcccgcgtggtggagggtcgcaggagcagcgtgaaagtggaggct
 gaggccagccggcagtga (SEQ ID NO:109)

[00398] Homo sapiens early growth response 1 (EGR1) (NM_001964) SEQ ID NO:46 (protein) and SEQ ID NO:110 (DNA)

MAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGAPQFLGAAGAPEGSGSNSSSS
 15 SSGGGGGGGGSGNSSSSSSTFNPQADTGEQPYEHLTAESFPDISLNNEKVLVETSYPSQTRLPP
 ITYTGRFSLEPAPNSGNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVP
 SNDSSPIYSAAPTFTPTNTDIFPEPQSQAFFGSAGTALQYPPPAYPAKGGFQVPMIPDYLFQQ
 QGDLGLGTPDQKPFQGLSRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQSQLIKPSRMKYP
 NRPSKTPPHERPYACPVESCDRRFSRSDDELTRHIRIHTGQKPFQCRICMRNFSRSDHLTTHIRTH
 20 TGEKPFACDICGRKFARSDEKRHTKIHLRQKDKKADKSVVASSATSSLSSYPSPVATSYSPSVT
 TSYPSPATTSYPSPVPTSFSPPGSSTYSPSVHSGFSPSVATTYSSVPPAFAQVSSFPSSAVTN
 SFSASTGLSDMTATFSPRTIEIC (SEQ ID NO:46)

atggccgcggccaaggccgagatgcagctgatgtcccgcctgcagatctctgaccggttcggatc
 ctttctcactcgcccaccatggacaactaccctaagctggaggagatgatgctgctgagcaacg
 25 gggctccccagttcctcggcgcgcggggggccccagagggcagcggcagcaacagcagcagcagc
 agcagcggggcggtggaggcggcgggggcggcagcaacagcagcagcagcagcaccttcaa
 ccctcaggcggacacgggcgagcagccctacgagcacctgaccgcagagtcttttctgacatct
 ctctgaacaacgagaaggtgctggtggagaccagttaccccagccaaaccactcgactgcccccc
 atcacctatactggcgcgttttccctggagcctgcacccaacagtggcaaaccttgtggcccgga
 30 gccctcttcagcttggctcagtggcctagtgcagcatgaccaaccaccggcctcctcgtcctcag
 caccatctccagcggcctcctccgcctccgcctccagagcccaccctgagctgcgcagtgcc
 tccaacgacagcagtcctcatttactcagcggcacccaccttccccacgcgaacactgacatttt
 ccctgagccacaaagccaggccttcccgggctcggcagggacagcgtccagtaccgcctcctg
 cctaccctgccgccaagggtggcttccaggttcccatgatccccgactacctgtttccacagcag
 35 cagggggatctgggcctgggcaccccagaccagaagcccttccagggcctggagagccgcaccca
 gcagccttcgctaaccctctgtctactattaaggcctttgccactcagtcgggctcccaggacc
 tgaaggccctcaataaccagctaccagtcctcagctcatcaaacccagccgatgcgcaagtacccc
 aaccggcccagcaagacgccccccacgaacgccttacgcttgcccagtgagtcctgtgatcg
 ccgcttctcccgtccgacgagctcaccgcgcacatccgcacacacaggccagaagcccttcc
 40 agtgccgcacatctgcatgcgaacttcagccgcagcgaccacctcaccacccacatccgcacccac
 acaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtgttgccaggagcagatgaacgcaa
 gaggcataccaagatccacttgccgcagaaggacaagaaagcagacaaaagtgttggtggcctctt
 cggccacctcctctctctcttccctacccgtccccgggttgctacctcttaccctgccccgggttact
 acctcttatccatccccggccaccacctcatacccatccccctgtgcccacctccttctcctctcc
 45 cggctcctcgacctacccatccccctgtgcacagtggttccccctccccgtcgggtggccaccacgt
 actcctctgttccccctgctttcccgccagggtcagcagcttcccttccctcagctgtcaccaac
 tccttcagcgcctccacagggctttcgacatgacagcaaccttttctcccaggacaattgaaat
 ttgctaa (SEQ ID NO:110)

[00399] Homo sapiens growth factor independent 1 transcription repressor(GFI1) (NM_005263) SEQ ID NO:47 (protein) and SEQ ID NO:111 (DNA)

MPRSFLVKSKKAHSYHQPRSPGPDYSLRLNVPAPSRADSTSNAGGAKAEPRDRLSPESQLTEAP
 DRASASPDSCEGSVCERSSEFEDFWRPPSPSPASEKSMCPSLDEAQPFFPLPFKPYSWGLAGS
 DLRLHVQSYRPGALERGAGLGLFCEPAPEPGHPAALYGPKRAAGGAGAGAPGSCSAGAGATAGP
 GLGLYGDFGSAAAGLYERPTAAAGLLYPERGHGLHADKGAGVKVESELLCTRLLLGGGSYKCIKC
 5 SKVFSTPHGLEVHVRRSHSGTRPFACEMCGKTFGHAVSLEQHKAVHSQERSFDCKICGKSFKRSS
 TLSTHLLIHS DTRPYPCQYCGKRFHQKSDMKKHTFIHTGEKPHKCQVCGKAFSQSSNLITHSRKH
 TGFKPFPGCDL CGKGFQRKVDLRRHRETQHGLK (SEQ ID NO:47)

atgccgcgctcatttctcgtcaaaagcaagaaggctcacagctaccaccagccgcgctccccagg
 accagactattccctccgtttagagaatgtaccggcgccctagccgagcagacagcacttcaaag
 10 caggcggggcggaaggcggagccccgggaccgtttgtccccgaatcgcagctgaccgaagcccca
 gacagagcctccgcacccccagacagctgcaaggcagcgtctgcgaacggagctcggagtttga
 ggacttctggaggcccccgctcaccctccgcgtctccagcctcggagaagtcaatgtgcccatcgc
 tggacgaagcccagcccttccccctgcctttcaaaccgtactcatggagcggcctggcggttct
 gacctgcggcacctggtgcagagctaccgaccgtgtggggccctggagcgtggcgctggcctggg
 15 cctcttctgcgaaccgcggcgccggagcctggccaccgcggcgctgtacggcccgaagcgggctg
 ccggcgggcgcgggggcgggcgccaggagctgcagcgcagggggccggtgccaccgctggccct
 ggctagggctctacggcgacttcgggtctgcggcagccgggtgtatgagaggcccacggcagc
 ggcgggcttgctgtaccccgagcgtggccacgggctgcacgcagacaagggcgctggcgctcaagg
 tggagtcggagctgctgtgcacccgcctgctgctggcgggcggtcctacaagtgcacaaagtc
 20 agcaaggtgttctccacgcgcacgggctcgaggtgcacgtgcgcaggtcccacagcgggtaccag
 accctttgcctgcgagatgtgcggcaagaccttcgggcacgcggtgagcctggagcagcacaag
 ccgtgcactcgcaggaacggagctttgactgtaagatctgtgggaagagcttcaagaggtcatcc
 aactgtccacacacctgcttatccactcagacactcggccctaccctgtcagtactgtggcaa
 gaggttccaccagaagtgcagacatgaagaaacacactttcatccactggtgagaagcctcaca
 25 agtgccaggtgtgcggcaaggcattcagccagagctccaacctcatcaccacagccgcaaacac
 acaggcttcaagcccttcggctgcgacctctgtgggaagggtttccagaggaaggtggacctccg
 aaggcaccgggagacgcagcatggggtcaaatga (SEQ ID NO:111)

[00400] Homo sapiens paired box 5 (PAX5) (NM_016734) SEQ ID NO:48 (protein) and SEQ ID NO:112(DNA)

MDLEKNYPTPRTSRTGHGGVNQLGGVFNNGRPLPDVVRQRIVELAHQGVVRPCDISRQLRVSHGCV
 SKILGRYYETGSIKPGVIGGSKPKVATPKVVEKIAEYKRQNPTMFAWEIRDRLAERVCDNDTVP
 SVSSINRIIRTKVQQPPNPVPASSHSIVSTGVSQTQSSVSTDSAGSSYSISGILGITSPSADTN
 KRKRDEGIQESFPVNGHSLPGRDFLRKQMRGDLFTQQQLEVLDRVFERQHYSDFITTEPIKPEQ
 TTEYSAMASLAGGLDDMKANLASPTPADIGSSVPGPQSYPIVTGRDLASTTLPGYPHPVPPAGQG
 35 SYSAPTTLTGMVPGSEFSGSPYSHPOYSSYNDSWRFNPGLLGSPYYYSAAARGAAPPAATAYDR
 H (SEQ ID NO:48)

atggatttagagaaaaattatccgactcctcggaccagcaggacaggacatggaggagtgaatca
 gcttgggggggtttttgtgaatggacggccactcccggatgtagtccgccagaggatagtggaa
 ttgctcatcaaggtgtcaggccctgcgacatctccaggcagcttcgggtcagccatggttgtgtc
 40 agcaaaattcttggcaggtattatgagacaggaagcatcaagcctggggtaattggaggatccaa
 accaaaggtcgccacacccaaagtgggtgaaaaaatcgtgaatataaacgcaaaatcccacca
 tgtttgcctgggagatcagggaccggctgctggcagagcgggtgtgtgacaatgacaccgtgcct
 agcgtcagttccatcaacaggatcatccggacaaaagtacagcagccacccaaccaaccagtccc
 agcttccagtcacagcatagtgtccactggctccgtgacgcaggtgtcctcgggtgagcagcgatt
 45 cggccggctcgtcgtactccatcagcggcatcctgggcatcacgtccccagcgccgacaccaac
 aagcgcaagagagacgaaggtattcaggagtctccgggtgccgaacggccactcgttccgggag
 agacttctccggaagcagatgcggggagacttgttcacacagcagcagctggaggtgctggacc
 gcgtgtttgagaggcagcactactcagacatcttcaccaccacagagcccacaaagccgagcag
 accacagagtattcagccatggcctcgtggtgggtgggctggacgacatgaaggccaatctggc
 50 cagccccacccctgctgacatcgggagcagtggtgccaggcccgagtcctacccattgtgacag

gccgtgacttggcgagcacgaccctccccgggtaccctccacacgtcccccccgctggacagggc
agctactcagcaccgacgctgacagggatgggtgcctgggagtgagttttccgggagtccttacag
ccaccctcagtattcctcgtacaaagactcctggaggttcccccaaccggggctgcttggtccc
cctactattatagcgtgccgcccaggagccgccccacctgcagccgcccactgcctatgaccgt
5 cactga (SEQ ID NO:112)

[00401] Homo sapiens T-cell-specific T-box transcription factor T-bet mRNA (AF241243)
SEQ ID NO:49 (protein) and SEQ ID NO:113 (DNA)

MGIVEPGCGDMLTGTEPMPGSDEGRAPGADPQHRYFYPEPGAQDADERRGGGSLGSPYPGGALVP
APPSRFLGAYAYPPRPQAAGFPGAGESFPPPADAEQYQPGEGYAAPDPRAGLYPGPREDYALPAG
10 LEVSGKLRVALNNHLLWSKFNQHQTEMIITKQGRMFPPFLSFTVAGLEPTSHYRMFVDVVLVDQH
HWRYQSGKWVQCCKAEGSMPGNRLYVHPDS PNTGAHWMRQEVSFGLKLTNNKGASNNVTQMIVL
QSLHKYQPR LHIVEVNDGEPEAACNASNTHIFTFQETQFIAVTAYQNAEITQLKIDNNPFAGFR
ENFESMYTSVDTSIPSPGPNCQFLGGDHYSPLLENQYPVPSRFY PDLPGQAKDVVPQAYWLGAP
RDHSYAEFRAVSMKPAFLPSAPGPTMSYYRGQEV LAPGAGWPVAPQYPPKMGPASWFRPMRTL P
15 MEPGPGGSEGRGPEDQGPPLVWTEIAPIRPESSDSGLGEGDSKR RRVSPYPSSGSSSPAGAPSP
FDKEAEGQFYNYFPN (SEQ ID NO:49)

atgggcatcgtggagccgggttgccggagacatgctgacgggcaccgagccgatgccggggagcga
cgagggccgggcgccctggcgccgacccgcagcaccgctacttctaccggagccgggcgcgagg
acgcggacgagcgtcgcgggggcgccgagcctgggggtctccctaccggggggcgcccttggtgcc
20 gccccgcgagccgcttccttgaggcctacgcctaccgcgcgaccccgaggcgccggctccc
cggcgcggggcgagtccttcccgcgcggcgccgagccgagggctaccagccgggaggggctacg
ccgccccggacccgcgcgcgggctctaccggggcgccgctgaggactacgcgctaccgcggga
ctggaggtgtcggggaaactgaggggtcgcgctcaacaaccacctgttggtccaagttaaactca
gcaccagacagagatgatcatcaccagcagggacggcgagatgttcccattcctgtcatttactg
25 tggccgggctggagcccaccagccactacaggatgtttgtggacgtggtcttggtggaccagcac
cactggcggtaccagagcggaagtgggtgcagtgtggaaaggccgagggcagcatgccaggaaa
ccgctgtacgtccaccggactcccccaacacaggagcgcactggatgcgccaggaaagtttcat
ttgggaaactaaagctcacaacaacaagggggcgctccaacaatgtgacccagatgattgtgctc
cagtcctccataagtaccagccccggctgcatatcggttgagggtgaacgacggagagccagaggc
30 agcctgcaacgcttccaacacgcatatctttactttccaagaaaccagttcattgcccgtgactg
cctaccagaatgccgagattactcagctgaaaattgataataaccctttgccaaaggattccgg
gagaactttgagtcctgtacacatctgttgacaccagcatccctccccgcctggacccaactg
tcaattccttgggggagatcactactctcctctcctacccaaccagtatcctgttcccagccgt
tctaccccgaccttccctggccaggcgaaggatgtggttccccaggttactggctgggggcccc
35 cgggaccacagctatgaggctgagtttcgagcagtcagcatgaagcctgcattcttgccctctgc
ccctggggccaccatgtcctactaccgaggccaggaggtcctggcacctggagctggctggcctg
tggcaccaccagctaccctcccaagatgggcccggccagctgggtccgcccctatgcggactctgcc
atggaaccggccctggaggctcagagggacggggaccagaggaccaggggtcccccttggtgtg
gactgagattgcccccatccggccgaatccagtgattcaggactgggcgaaggagactctaaga
40 ggaggcgcgtgtccccctatccttccagtgggtgacagctcctcccctgctggggcccccttctcct
tttgataaggaagctgaaggacagttttataactattttcccaactga (SEQ ID NO:113)

[00402] Homo sapiens GATA binding protein 3 (GATA3) (NM_001002295) SEQ ID NO:50
(protein) and SEQ ID NO:114 (DNA)

MEVTADQPRVSHHHPAVLNGQHPDTHHPGLSHSYMDAAQYPLPEEVDVLFNIDGQGNHVPPYYG
45 NSVRATVQRYPPTHHGSQVCRPPLLHGSLPWL DGGKALGSHHTASPNWLS PFSKTSIHGSPGPL
SVYPPASSSSLSGGHASPHLFTFPPTPPKDVSPDPSLSTPGSAGSARQDEKECLKYQVPLPDSMK
LESSHSRGSMTALGGASSSTHHPITTYPPYVPEYSSGLFPPSSLLGGSPTGFGCKSRPKARSSTE
GRECVNCGATSTPLWRRDGTGHYLCNACGLYHKMNGQNRPLIKPKRRLSAARRAGTSCANCQTTT
TTLWRRNANGDPVCNACGLYYKLHNINRPLTMKKEGIQTRNRKMSSSKSKCKKVHDSLEDFPKNS

SFNPAALSRHMSSLISHISPFSSHMLTTPTPMHPPSSLSFGPHHPSSMVTAMG (SEQ ID NO: 50)

atggaggtgacggcggaccagccgcgctgggtgagccaccaccaccccgccgtgctcaacgggca
gcacccggacacgcaccacccgggcctcagccactcctacatggacgcggcgcagtaccgcgtgc
5 cggaggaggtggatgtgctttttaacatcgacgggtcaaggcaaccacgtcccgccctactacgga
aactcggtcagggccacgggtgcagaggtaccctccgacccaccacgggagccaggtgtgccgccc
gcctctgcttcatggatccctaccctggctggacggcggcaaagccctgggcagccaccacaccg
cctccccctggaatctcagccccttctccaagacgtccatccaccacgggtccccggggcccctc
tccgtctaccccccggcctcgtcctcctccttgtcggggggccacgccagcccgccactcttcac
10 cttcccgccccaccccgccgaaggacgtctccccggaccatcgctgtccaccccagggtcggccg
gctcggcccggcaggacgagaaagagtgctcaagtaccaggtgcccctgcccagacagcatgaag
ctggagtcgtcccactcccgtggcagcatgaccgcccctgggtggagcctcctcgtcgacccacca
ccccatcaccacctacccgcccctacgtgcccaggtacagctccggactcttcccccccagcagcc
tgctggggcggtccccccaccggcttcggatgcaagtccaggcccaaggcccggtccagcacagaa
15 ggcaggagtggtgtgaactgtggggcaacctcgaccccactgtggcgggcgagatggcacgggaca
ctacctgtgcaacgcctgcgggctctatcacaaaatgaacggacagaaccggccccctcattaagc
ccaagcgaaggctgtctgcagccaggagagcagggacgtcctgtgcaactgtcagaccaccaca
accacactctggaggaggaatgccaatggggaccctgtctgcaatgctgtgggctctactacaa
gcttcacaatatataacagacccctgactatgaagaaggaaggcatccagaccagaaaccgaaaaa
20 tgtctagcaaatccaaaaagtgcaaaaaagtgcattgactcactggaggacttccccaaagaacagc
tcgtttaacccggccgcccctctccagacacatgtcctccctgagccacatctcgcccttcagcca
ctccagccacatgctgaccacgcccacgcccgtgcacccgccatccagcctgtcctttggaccac
accacccctccagcatgggtcaccgccatgggttag (SEQ ID NO: 114)

[00403] Homo sapiens FOXP3 (EF534714) SEQ ID NO:51 (protein) and SEQ ID NO:115 (DNA)

MPNPRPGKPSAPSLALGPSGASPSWRAAPKASDLLGARGPGGTFQGRDLRGGAHASSSSSLNMP
PSQLQLPTLPLVMVAPSGARLGPLPHLQALLQDRPHFMHQLSTVDAHARTPVLQVHPLESPAMIS
LTPPTTATGVFSLKARPLPPGINVASLEWVSREPALCTFPNPSAPRKDSTLSAVPQSSYPLLA
NGVCKWPGCEKVFEEDFLKHCQADHLLDEKGRAQCQLQREMVSLEQQLVLEKEKLSAMQAH
30 AGKMALTKASSVASSDKGSCCIVAAGSQGPVPAWSGPREAPDSLFAVRRHLWGSHTNSTFPEFL
HNMDYFKFHNMRPFTYATLIRWAILAEPEKQRTLNEIYHWFTRMF AFFRNHPATWKNAIRHNS
LHKCFVRVESEKGAVWTVDELEFRKKRSQRPSRCSNPTPGP (SEQ ID NO: 51)

atgcccaacccagggcctggcaagccctcgccccccttcttggcccttggcccatccccaggagc
ctcgcccagctggagggtgcacccaaagcctcagacctgctggggggcccgggggccagggggaa
35 ccttccaggggcgagatcttcgaggcgggggcccatgctcctcttcttcttcttgaaccccatgcc
ccatcgagctgcagctgccacactgccctagtcattggtggcaccctccggggcacggctggg
ccccttggcccaacttacaggcactcctccaggacaggccacatttcatgcaccagctctcaacgg
tggatgccacgcccggaccctgtgctgcaggtgcacccctggagagcccagccatgatcagc
ctcacaccacccaccacggcactggggcttctcctcctcaaggcccgccctggcctcccacctgg
40 gatcaacgtggccagcctggaatgggtgtccaggagccggcactgctctgcaccttcccaaatc
ccagtgcacccaggaaggacagcacccttccggctgtgccccagagctcctaccactgctggca
aatggtgtctgcaagtggccggatgtgagaaggctctcgaagagccagaggacttccctcaagca
ctgccaggcgccacatcttctggatgagaagggcagggcacaatgtctcctccagagagagatgg
tacagtctctggagcagcagctgggtgctggagaaggagaagctgagtgccatgcaggccacctg
45 gctgggaaaatggcactgaccaaggcttcatctgtggcatcatccgacaagggtcctgctgcat
cgtagctgctggcagccaaggccctgtcgtcccagcctgggtctggcccccgaggggccctgaca
gcctgtttgctgtccggaggcacctgtggggtagccatggaaacagcacattcccagagttcctc
cacaacatggactacttcaagttccacaacatgcgacccctttcacctacgccacgctcatccg
ctgggccatcctggaggctccagagaagcagcggacactcaatgagatctaccactgggttcacac
50 gcatgtttgccttcttcagaaaccatcctgccacctggaagaacgccatccgccacaacctgagt

ctgcacaagtgctttgtgcggtggagagcgagaagggggctgtgtggaccgtggatgagctgga
gttccgcaagaaacggagccagaggccagcaggtgttccaaccctacacctggccctga
(SEQ ID NO:115)

[00404] Human orphan receptor ROR gamma (U16997) SEQ ID NO:52 (protein) and SEQ ID NO:116 (DNA)

MDRAPQRQHRASRELLAAKKTHTSQIEVIPCKICGDKSSGIHYGVITCEGCKGFFRRSQRCNAAY
SCTRQQNCPIDRTSRNRCQHCRLOKCLALGMSRDAVKFGRMSKKQRDSLHAEVQKQLQQRQQQQQ
EPVVKTPPAGAQQADTLTYTLGLPDGQLPLGSSPDLEASACPPGLLKASGSGPSYSNNLAKAGL
NGASCHLEYSPEERGKAEGRESFYSTGSQLTDPDRCGLRFEEHRHPGLGELGQGPDSYGSFSFRSTP
EAPYASLTEIEHLVQSVCKSYRETCQLRLEDLLRQSRNIFSRREEVTGYQRKSMWEMWERCAHHLT
EAIQYVVEFAKRLSGFMELCQNDQIVLLKAGAMEVVLVRLMCRAYNADNRTVFFEGKYGGMELFRA
LGCSELISSIFDFSHSLSALHFSEDEIALYALVLINAHRPGLQEKRKVEQLQYNLELAFHHHLC
KTHRQSILAKLPPKGKLRSLCSQHVERLQIFQHLHPIVVQAAFPPLYKELFSTETESPVGCPSDL
EEGLLASPYGLLATSLDPVPPSPFSPFMPNPGGWSPPALWK (SEQ ID NO:52)

atggacagggcccccacagagacagcaccgagcctcacgggagctgctggctgcaaagaagaccca
cacctcaciaaattgaagtgatcccttgcaaaatctgtggggacaagtctgtctgggatccactacg
gggttatcacctgtgaggggtgcaagggtcttctccgccggagccagcgtgtaacgcggcctac
tcctgcacccgtcagcagaactgccccatcgaccgcaccagccgaaaccgatgccagcactgccg
cctgcagaaatgcttggtgctggggatgtcccagatgctgtcaagttcggccgcatgtccaaga
agcagagggacagcctgcatgcagaagtgcagaaacagctgcagcagcggcaacagcagcaacag
gaaccagtgggtcaagacccctccagcaggggccccaggagcagataccctcacctacaccttggg
gctcccagacgggagctgcccctgggctcctcgctgacctgctgaggttctgctgtcccc
ctggcctcctgaaagcctcaggtctctgggcctcatattccaacaacttgccaaggcagggctc
aatggggcctcatgccacctgaatacagccctgagcggggcaaggctgagggcagagagagctt
ctatagcacaggcagccagctgacccctgaccgatgtggacttcgttttgaggaacacaggcatc
ctgggcttggggaactgggacagggccagacagctacggcagccccagtttccgcagcacaccg
gaggcaccctatgctctccctgacagagatagagcacctgggtgcagagcgtctgcaagtcctacag
ggagacatgccagctgcggctggaggacctgctgcggcagcgtccaacatcttctcccgggagg
aagtgactgggtaccagaggaagtccatgtgggagatgtgggaacgggtgtgcccaccacctcacc
gaggccattcagtagctgggtggagttcgccaagaggctctcaggctttatggagctctgccagaa
tgaccagattgtgcttctcaaagcaggagcaatggaaagtgggtgctgggttaggatgtgccgggct
acaatgctgacaaccgcacggtcttttttgaggcaatacgggtggcatggagctgttccgagcc
ttgggctgcagcagctcatcagctccatctttgacttctccactccctaagtgccttgcaactt
ttccgaggatgagattgccctctacacagcccttggttctcatcaatgcccatcgggcagggtcc
aagagaaaaggaaagtagaacagctgcagtacaatctggagctggcctttcatcatcatctctgc
aagactcatcgccaaagcatcctggcaagctgccacccaagggaagcttcggagcctgtgtag
ccagcatgtggaaaggctgcagatcttcagcacctccaccccatcgtggtccaagccgctttcc
ctccactctacaaggagctcttcagcactgaaaccgagtcacctgtgggctgtccaagtgacctg
gaagaggggactccttgctctccctatggcctgctggccacctccctggaccccggttccacctc
acccttttcccttcccatgaaccctggagggtggtccccaccagctctttggaagtga (SEQ
ID NO:116)

[00405] Homo sapiens hepatocyte nuclear factor 4, alpha (HNF4A), (NM_178849) SEQ ID NO:53 (protein) and SEQ ID NO:117 (DNA)

MRLSKTLVDMMDADYSAALDPAYTTLEFENVQVLTMGNDTSPSEGTLNAPNSLGVSALCAICGD
RATGKHYGASSCDGCKGFFRRSVRKNHMYSCRFSRQCVVDKDKRNQCRYCRLKKCFRAGMKKEAV
QNERDRISTRSSYEDSSLPSINALLQAEVLSRQITSPVSGINGDIRAKKIASIADVCESMKEQL
LVLVEWAKYIPAFCELPDDQVALLRAHAGEHLLLGATKRSMVFKDVLLLGNDYIVPRHCPPELAE
MSRVSIIRILDELVLFPQELQIDDNEYAYLKAIIFDPDAKGLSDPGKIKRLRSQVQVSLDYIND
RQYDSRGRFGELLLLLPTLQSIWQMIEQIQFIKLFMAKIDNLLQEMLLGSPSDAPHAHPLH

PHLMQEHMGTNVIVANTMPHTLSNGQMSTPETPQPSPPGGSGSEPYKLLPGAVATIVKPLSAIPQ
PTITKQEV (SEQ ID NO:53)

atgcgactctccaaaaccctcgctcgacatggacatggccgactacagtgcactggaccagc
ctacaccaccctggaatttgagaatgtgcaggtgttgacgatgggcaatgacacgtccccatcag
5 aaggcaccaacctcaacgcgcccacagcctgggtgtcagcgccctgtgtgccatctgcggggac
cgggccacgggcaaacactacggtgcctcgagctgtgacggctgcaagggcttcttcggaggag
cgtgcggaagaaccacatgtactcctgcagatttagccggcagtgctggtggacaaagacaaga
ggaaccagtgccgctactgcaggctcaagaaatgcttcgggctggcatgaagaaggaagccgtc
cagaatgagcgggaccggatcagcactcgaaggtcaagctatgaggacagcagcctgccctccat
10 caatgcgctcctgcaggcggaggtcctgtcccgcagatcacctccccctgtcctgggatcaacg
gcgacattcgggcgaagaagattgccagcatcgcagatgtgtgtgagtcctgaaggagcagctg
ctggttctcgttgagtgggccaagtacatcccagcttcttgcgagctccccctggacgaccaggt
ggcctgtcagagcccatgctggcgagcacctgctgctcggagccaccaagagatccatggtgt
tcaaggacgtgctgctcctaggaatgactacattgtccctcggcactgcccggagctggcggag
15 atgagccgggtgtccatacgcaccttgacgagctggtgctgcccttcaggagctgcagatcga
tgacaatgagtatgcctacctcaaagccatcatcttctttgaccagatgccaaggggctgagcg
atccagggaagatcaagcggctgcgttcccaggtgcaggtgagcttgaggactacatcaacgac
cgccagtatgactcgcgtggccgctttggagagctgctgctgctgctgccaccttgcagagcat
cacctggcagatgatcgagcagatccagttcatcaagctcttcggcatggccaagattgacaacc
20 tgttgaggagatgctgctgggaggggtccccagcgatgcaccccatgccaccacccccctgcac
cctcacctgatgcaggaacatatgggaaccaacgtcatcggttgccaacacaatgccactcacct
cagcaacggacagatgtccacccctgagacccacagccctcaccgccaggtggctcagggctctg
agccctataagctcctgccgggagccgtcgccacaatcgtaagccctctctgcatccccag
ccgaccatcaccaagcaggaagttatctag (SEQ ID NO:117)

[00406] Human signal transducer and activator of transcription Stat5A (U43185) SEQ ID
NO:54 (protein) and SEQ ID NO:118 (DNA)

MAGWIAQQQLQGDALRQMQLVLYGQHFPIEVRHYLAQWIESQPWDAIDLDPQDRAQATQLLEGLV
QELQKKAHQVGEDGFLKIKLRHYATQLQKTYDRCPLELVRCIRHILYNEQRLVREANNCSSPA
GILVDAMSQKHLQINQTFEELRLVTQDTENELKKLQQTQYFYIIQYQESLRIQAQFAQLAQLSPQ
30 ERLSRETALQQKQVSLEAWLQREAQTLQQYRVELAEKHQKTLQLLRKQQTII LDDELIQWKRRQQ
LAGNGGPPEGSLDLVLSWCEKLAELI WQNRQQIRRAEHLCOQLPIPGPVEEMLAEVNATITDIIS
ALVTSTFII EKQPPQVLKTQTKFAATVRLLVGGKLNVMNPPQVKATII SEQQAKSLLKNENTRN
ECSGEILNNCCVMEYHQATGTL SAHFRNMSLKRIKRADRRGAESVTEEKFTVL FESQFSVGSNEL
VFQVKTL SLPVVIVHGSQDHNATATVLWDNAFAEPGRVPFAVPDKVLWPQLCEALNMKFKA EVQ
35 SNRGLTKENLVFLAQKLFNNSSSHLEDYSGLSVSWSQFNRENLPGWNYTFWQWFDGVM EVLKKHH
KPHWNDGAILGFVNKQQAHDLLINKPDGTFLLRFS DSEIGGITIAWKFDSPERNLWNLKPF TTRD
FSIRSLADRLGDL SYLIYVFPDRPKDEVFSKY YTPVLAKAVDGYVKPQIKQVVPEFVNASADAGG
SSATYMDQAPSPAVCPQAPYNMYPQNPDHVL DQDGEFDLDETMDVARHVEELLRRPMDSLDSRLS
PPAGLFTSARGSL (SEQ ID NO:54)

atggcgggctggatccaggcccagcagctgcaggagacgcgctgcgccagatgcaggtgctgta
cggccagcacttccccatcgaggtccggcactacttgcccagtggttgagagccagccatggg
atgccattgacttggaacatccccaggacagagcccacccagctcctggagggcctggtg
caggagctgcagaagaaggcggagcaccaggtgggggaagatgggtttttactgaagatcaagct
gaggcactacgccacgcagctccagaaaacatatgaccgctgccccctggagctgggtccgctgca
40 tccggcacattctgtacaatgaacagaggctgggtccgagaagccaacaattgcagctctccggct
gggatcctgggtgacgccatgtcccagaagcaccttcagatcaaccagacatttgaggagctgcg
actggtcacgcaggacacagagaatgagctgaagaaactgcagcagactcaggagtacttcatca
tccagtaccaggagagcctgaggatccaagctcagtttgcccagctggcccagctgagccccag
gagcgtctgagccgggagacggccctccagcagaagcaggtgtctctggaggcctggttgacgcg
45 tgaggcacagacactgcagcagtagccgctggagctggccgagaagcaccagaagaccctgcagc

tgctgcggaagcagcagaccatcatcctggatgacgagctgatccagtggaagcggcggcagcag
 ctggccgggaacggcggggcccccgagggcagcctggacgtgctacagtcctgggtgtgagaagtt
 ggccgagatcatctggcagaaccggcagcagatccgcagggctgagcacctctgccagcagctgc
 ccatccccggcccagtgaggagatgctggccgaggtcaacgccaccatcacggacattatctca
 5 gccctggtgaccagcacattcatcattgagaagcagcctcctcaggtcctgaagaccagaccaa
 gtttgacgccaccgtacgcctgctgggtggcggaagctgaacgtgcacatgaatccccccaggg
 tgaaggccaccatcatcagtgagcagcaggccaagtctctgcttaaaaatgagaacaccgcgaac
 gagtgcagtggtgagatcctgaacaactgctgcgtgatggagtaccaccaagccacgggcaccct
 cagtgcacacttcaggaacatgtcactgaagaggatcaagcgtgctgaccggcgggggtgcagagt
 10 ccgtgacagaggagaagttcacagtcctgtttgagttcagttcagttggtgagcaatgagctt
 gtgttccaggtgaagactctgtccctacctgtggttgctcatcgccacggcagccaggaccacaa
 tgccacggctactgtgctgtgggacaatgcctttgctgagccgggcaggggtgccatttgccgtgc
 ctgacaaaagtgtgtggccgcagctgtgtgaggcgctcaacatgaaattcaaggccgaagtgcag
 agcaaccggggcctgaccaaggagaacctcgtgttctggcgagaaactgttcaacaacagcag
 15 cagccacctggaggactacagtggcctgtccgtgtcctgggtcccagttcaacagggagaacttgc
 cgggctggaactacaccttctggcagtggtttgacgggggtgatggaggtgttgaagaagcaccac
 aagccccactggaatgatggggccatcctaggttttgtgaataagcaacaggcccacgacctgct
 catcaacaagcccagcgggaccttcttgttgcgctttagtgactcagaaatcgggggcatcacca
 tcgcttggaagtttgattccccggaacgcaacctgtggaacctgaaaccattcaccacgcgggat
 20 ttctccatcaggtccctggctgaccggctgggggacctgagctatctcatctatgtgtttcctga
 ccgccccaaaggatgaggtcttctccaagtactacactcctgtgctggctaaagtgttgatggat
 atgtgaaaccacagatcaagcaagtggcctcctgagtttgtgaatgcatctgcagatgctgggggc
 agcagcgccacgtacatggaccaggccccctccccagctgtgtgccccaggctccctataacat
 gtaccacagaacctgacctgtactcgatcaggatggagaattcgacctggatgagacctggt
 25 atgtggccaggcagctggaggaactcttacgcgcgaccaatggacagctcttgactcccgccctctcg
 cccctgcegggtcttttcacctctgccagaggtccctctcatga (SEQ ID NO:118)
[00407] Homo sapiens squalene epoxidase (ERG1) mRNA,(AF098865) SEQ ID NO:55
(protein) and SEQ ID NO:119 (DNA)
 MWTFGLGIATFTYFYKKFGDFITLANREVLVCVLVFLSLGLVLSYRCRHRNGLLGRQQSGSQFAL
 30 FSDILSGLPFIGFFWAKSPPESENKEQLEARRRRKGTNISETSLIGTAACTSTSSQNDPEVIVG
 AGVLGSALAAVLSRDGRKVTVIERDLKEPDRIVGEFLQPGGYHVLKDLGLGDTVEGLDAQVNGY
 MIHDQESKSEVQIPYPLSENNQVQSGRAFHGGRFIMSLRKAAMAEPNKFIIEGVVLQLEEDDVV
 MGVOYKDKETGDIKELHAPLTVVADGLFSKFRKSLVSNKVSSSHFGVFLMKNPQFKANHAELI
 LANPSPVLIYQISSSETRVLVDIRGEMPRNLREYMVEKIYPQIPDHLKEPFLEATDNSHLRSMIPA
 35 SFLPPSSVKKRGVLLGLDAYNMRHPLTGGGMTVAFKDIKLWRKLLKGIPDLYDDAAIFEAKKSFY
 WARKTSHSFVNNILAQALYELFSATDDSLHQLRKACFLYFKLGGECVAGPVGLLSVLSPNPLVLI
 GHFFAVAIYAVYFCFKSEPWITKPRALLSSGAVLYKACSVIFPLIYSEMKYMVH (SEQ ID
 NO:55)
 atgtggacttttctgggcattgccactttcacctatttttataagaagttcggggacttcatcac
 40 tttggccaacagggaggtcctgttgtgctgctgggtgttccctctcgtggtggcctgggtgctcct
 accgctgtcgccaccgaaacgggggtctcctcgggcgcagcagagcggtcccagttcgccctc
 ttctcgatattctctcaggcctgcctttcattggcttcttctgggccaatccccccctgaatc
 agaaaataaggagcagctcgaggccaggaggcgcagaaaaggaaccaatatttcagaaacaagct
 taataggaacagctgcctgtacatcaacatcttctcagaatgacccagaagttatcatcgtggga
 45 gctggcgtgcttggctctgctttggcagctgtgctttccagagatggaagaaaggtgacagtcac
 tgagagagacttaaaagagcctgacagaatagttggagaattcctgcagccgggtggttatcatg
 ttctcaaagaccttggcttggagatacagtggaaaggtcttgatgccaggttgtaaattggttac
 atgattcatgatcaggaaagcaaatcagaggttcagattccttaccctctgtcagaaaacaatca
 agtgacagagtggaaagagctttccatcacggaagattcatcatgagctctcggaaagcagctatgg
 50 cagagcccaatgcaaagtttattgaaggtgttgtgttacagttattagaggaagatgatgttgtg

atgggagttcagtacaaggataaagagactggagatatcaaggaactccatgctccactgactgt
 tgttgcatgagtggtcttttctccaagttcaggaaaagcctggtctccaataaagtttctgtatcat
 ctcattttgttggtctttcttatgaagaatgcaccacagtttaagcaaactcatgctgaacttatt
 ttagctaaccgagtgccagttctcatctaccagatttcatccagtgaaactcgagtacttgttga
 5 cattagaggagaaatgccaaaggaatttaagagaatacatgggtgaaaaaatttaccacaaaatac
 ctgatcacctgaaagaaccattcttagaagccactgacaattctcatctgaggtccatgccagca
 agcttccttcctccttcctcatcagtgaagaaacgaggtgttcttcttttgggagacgcatataat
 gaggcacccacttactggtggaggaatgactgttgcttttaagatataaaactatggagaaaac
 tgctaaagggatccctgacctttatgatgatgcagctattttcgaggccaaaaaatcattttac
 10 tgggcaagaaaaacatctcattcctttgtcgtgaatatccttgctcaggctctttatgaattatt
 ttctgccacagatgattccctgcatcaactaagaaaagcctgttttctttatttcaaacttggtg
 gcaatgtgttgcggtcctgttgggtgctttctgtattgtctcctaaccctctagttttaatt
 ggacacttctttgctgttgcaatctatgccgtgtatttttgccttaagtcagaaccttgattac
 aaaacctcgagcccttctcagtagtggtgctgtattgtacaaagcgtgttctgtaatatctcctc
 15 taatttactcagaaatgaagtatatggttcattaa (SEQ ID NO:119)

[00408] Homo sapiens ets variant 2 (ETV2/ER71), (NM_014209) SEQ ID NO:56 (protein) and SEQ ID NO:120 (DNA)

MDLWNWDEASPQEVPPGNKLAGLEGAKLGFCFPDLALQGDTPATAETCWKGTSSSLASFPQLDW
 GSALLHPEVPWGAEPDSQALPWSGDWTDMACTAWDSWSGASQTLGPAPLPGPIPAAGSEGAAGQ
 20 NCVPVAGEATSWSRAQAAGSNTSWDCSVGPDGDTYWGSLGGEPRDCTISWGGPAGPDCTTSWN
 PGLHAGGTTSLKRYQSSALTVCSESPSQSDRASLARCPKTNHRGPQLWQFLLELLHDGARSSCI
 RWTGNSREFQLCDPKEVARLWGERKRKPGMNYEKLRSRGLRYYYRRDIVRKSGGRKYTYRFGGRVP
 SLAYPDCAGGGRGAETQ (SEQ ID NO:56)

atggacctgtggaactgggatgagggcatccccacaggaagtgcctccagggaacaagctggcagg
 25 gcttgaaggagccaaattaggcttctgtttccctgatctggcactccaaggggacacgccgacag
 cgacagcagagacatgctgaaaggtacaagctcatccctggcaagcttcccacagctggactgg
 ggctccgcgttactgcacccagaagttccatggggggcgagcccactctcaggctcttccgtg
 gtccggggactggacagacatggcgtgcacagcctgggactcttgagcgggcgccctcgcagaccc
 tggggccccgccccctctcgccccggggccccatccccgcccggctccgaaggcgccgcggggccag
 30 aactgcgtccccgtggcgaggagagggccacctcgtggtcgcgcgcccaggccgcccgggagcaacac
 cagctgggactgttctgtggggccccgacggcgatacctactggggcagtggcctggggcggggagc
 cgcgacaggactgtaccatttctgtggggcgggccccgaggcccgactgtaccacctcctggaac
 ccggggctgcatgcggggtggcaccacctctttgaagcgggtaccagagctcagctctcaccgtttg
 ctccgaaccgagccccgcagtcggaccgtgccagtttggtcgtatgccccaaaactaaccaccgag
 35 gtccattcagctgtggcagttcctcctggagctgctccacgacggggcgcgtagcagctgcac
 cgttggactggcaacagccgcgagttccagctgtgacccccaaagaggtggctcggtgtgggg
 cgagcgcaagagaaagccgggcatgaattacgagaagctgagccggggccttcgctactactatc
 gccgcgacatcgtgcgcaagagcggggggcgaaagtacacgtaccgcttcgggggcccgcgtgcc
 agcctagcctatccggactgtgcgggaggcgagcggggagcagagacacaataa (SEQ ID
 40 NO:120)

[00409] Homo sapiens GATA binding protein 2 (GATA2) (NM_001145661) SEQ ID NO:57 (protein) and SEQ ID NO:121 (DNA)

MEVAPEQPRWMAHPAVLNAQHPSHHPLAHNYMEPAQLLPDEVDFVFNHLD SQGNPYYPANPAH
 ARARVSYPARHRLTGGMCRPHLLHSPGLPWLDGGKAALSA AAAHHHNPWTVSPFSKTPHPSA
 45 AGGPGGPLSVYPGAGGSGGGSGSSVASLTPTAAHSGSHLFGFPPTPPKEVSPDPSTTGAASPAS
 SSAGGSAARGEDKDGVKYQVSLTESMKMESGSPLRPGLATMGTPATHHPIPTYPSYVPA AHDY
 SSGLFHPGGFLGGPASSFTPQQRSKARSCSEGRCVNCGATATPLWRRDGTGHYLCNACGLYHKM
 NGQNRPLIKPKRRLSARRAGTCCANCQTTTTTLWRRNANGDPVCNACGLYYKLHNVNRP LTMKK
 EGIQTRNRKMSNKS KSKGAECFEELSKCMQEKSSPFSAAALAGHMAPVGHLPFFSHSGHILPT
 50 PTPIHPSSSLSF GHPHPSSMVTAMG (SEQ ID NO:57)

atggaggtggcgcccgagcagccgcgctggatggcgacccggccgtgctgaatgcgagcacc
cgactcacaccacccgggctggcgcaactacatggaacccgcgcagctgctgcctccagacg
aggtggacgtcttcttcaatcacctcgactcgagggcaacccctactatgccaaccccgctcac
gcgcgggcgcgctctctacagccccgcgcagcccgctgaccggaggccagatgtgccgccc
5 acacttggtgcacagccccgggtttgccctggctggacgggggcaaagcagccctctctgccgctg
cggccccaccaccacaacccctggaccgtgagcccccttctccaagacgccactgcacccctcagct
gctggaggccctggaggccactctctgtgtaccaggggctgggggtgggagcgggggaggcag
cgggagctcagtggcctccctcaccctacagcagccactctggctcccaccttttcggcttcc
caccacgccacccaaagaagtgtctcctgaccctagcaccacgggggctgctctccagcctca
10 tcttccgcgggggtagtgagcccgaggagaggacaaggacggcgtcaagtaccaggtgtcact
gacggagagcatgaagatggaaagtggcagctccctgcgcccaggcctagctactatgggcaccc
agcctgctacacaccacccccatccccactacccctcctatgtgccggcggtgcccacgactac
agcagcggactcttccaccccgagggttccctggggggaccggcctccagcttcacccctaagca
gcgagcaaggctcgttccctgttcagaaggccgggagtggtgcaactgtggggccacagccaccc
15 ctctctggcgggcgggacggcaccggccactacctgtgcaatgcctgtggcctctaccacaagatg
aatgggcagaaccgaccactcatcaagcccaagcgaagactgtcgcccgccagaagagccggcac
ctggtgtgcaaattgtcagacgacaaccaccaccttatggcgccgaaacgccaacggggaccctg
tctgcaacgcctgtggcctctactacaagctgcacaatgttaacaggccactgaccatgaagaag
gaagggatccagactcggaaccggaagatgtccaacaagtccaagaagagcaagaaaggggcgga
20 gtgcttcgaggagctgtcaaagtgcagtgaggagaagtcatcccccttcagtgagctgccttg
ctggacacatggcacctgtggggccacctcccgcccttcagccactccggacacatcctgccact
ccgacgcccatccacccctcctccagcctctccttcggccacccccaccgctccagcatggtgac
cgccatgggctag (SEQ ID NO:121)

[00410] Homo sapiens hes-related family bHLH transcription factor with YRPW motif 1 (HEY1) (NM_012258) SEQ ID NO:58 (protein) and SEQ ID NO:122 (DNA)

MKRAHPEYSSSDSELDETIEVEKESADENGNLSSALGMSPTTSSQILARKRRRGII EKRRRDRI
NNSLSELRLVPSAFEKQGS AKLEKAEILQMTVDHLKMLHTAGGKGYFDAHALAMDYRSLGFREC
LAEVARYLSII EGLDASDPLRVRLVSHLN NYASQREAA SGAHAGLGHIPWGT VFGHHPHIAHPLL
LPQNGHGNAGTTASPT PHHQRLGSAHPEAPALRAPPSGSLGPVLPVVT SASKLSPPLLSSVAS
30 LSAFPFSFGSFHLLSPNALSPSAPTQAANLGKPYRPWGTEIGAF (SEQ ID NO:58)

atgaagcgagctcaccccgagtacagctcctcggacagcgagctggacgagaccatcgaggtgga
gaaggagagtgcggacgagaatggaaacttgagttcggctctaggttccatgtccccaactacat
cttcccagattttggccagaaaaagacggagaggaataattgagaagcgcgacgagaccggatc
aataacagtttgtctgagctgagaaggctggtacccagtgtcttttgagaagcagggatctgctaa
35 gctagaaaaagccgagatcctgcagatgaccgtggatcacctgaaaatgtgcatacggcaggag
ggaaaggttactttgacgcgcacgcccttgctatggactatcgaggtttgggatttcgggaatgc
ctggcagaagttgcggttatctgagcatcattgaaggactagatgcctctgacccgcttcgagt
tcgactggtttcgcatctcaacaactacgcttcccgagcgggaagccgcgagcggcgccacgcgg
gcctcggacacattccctgggggaccgtcttcggacatcaccgcacatcgcgaccccgctgttg
40 ctgccccagaacggccacgggaacgcgggcaccacggcctcaccacggaaccgcaccaccaggg
caggctgggctcggcacatccggaggcgctgtttgcgagcgccccctagcggcagcctcggac
cgggtgctccctgtggteacctccgectccaaactgtcgccgctctgctctcctcagtggcctcc
ctgtcggccttcccccttctcttccggctccttccacttactgtctcccaatgcactgagcccttc
agcaccacgcaggtgcaaaccttggcaagccctatagaccttgggggacggagatcgagctt
45 tttaa (SEQ ID NO:122)

[00411] Homo sapiens Hey2 (AB044755) SEQ ID NO:59 (protein) and SEQ ID NO:123 (DNA)

MKRPCEETTSESMDDETIDVGSENNYSQSTSSVIRLNSPTTTSQIMARKKRRGII EKRRRDRI
NSLSELRLVPTAF EKQGS AKLEKAEILQMTVDHLKMLQATGGKGYFDAHALAMDFMSIGFRECL
50 TEVARYLSSVEGLDSSDPLRVRLVSHLSTCATQREAAAMTSSMAHHHHPLPHPHWAAAFHHLPA

LLQPNGLHASESTPCRLSTTSEVPPAHGSALLTATFAHADSA LRMPSTGSVAPCVPLSTSLLSL
SATVHAAAAAATAAAHSFPLSFAGAFPM LPPNAAAAVAAATAISPPLSVSATSSPQQTSSGTNNK
PYRPWGTEVGAF (SEQ ID NO:59)

atgaagcgccccctgcgaggagacgacctccgagagcgacatggacgagaccatcgacgtggggag
5 cgagaacaattactcggggc aaagtactagctctgtgattagattgaattctccaacaacaacat
ctcagattatggcaagaaagaaaaggagagggattatagagaaaaggcgctcgggatcggataaat
aacagtttatctgagttgagaagacttgtgccaaactgcttttgaaaaacaaggatctgcaaagtt
agaaaaagctgaaatattgcaaatgacagtggatcatttgaagatgcttcaggcaacagggggta
aaggctactttgacgcacacgctcttgccatggacttcatgagcataggattccgagagtgccta
10 acagaagttgcgcgggtacctgagctccgtggaaggcctggactcctcggatccgctgcgggtgcg
gcttgtgtctcatctcagcacttgccaccacccagcgggagggcgccgcatgacatcctccatgg
cccaccaccatcatcgcgtccaccgcgcatcactggggcgccgccttccaccacctgcccgcagcc
ctgctccagcccaacggcctccatgcctcagagtc aaccccttgctgcctctccacaacttcaga
agtgcctcctgcccacggctctgctctcctcacggccacgcttgcccatgcggattcagccctcc
15 gaatgccatccacgggcagcgtcgccccctgctgcccactctctccacctctctctgtccctc
tctgccaccgtccacgcgcgagccgcagcagccaccgcggtgcacacagcttccctctgtcctt
cgcgggggcattccccatgcttcccccaaacgcagcagcagcagtgggcgcgccacagccatca
gcccgccttgtcagtatcagccacgtccagtcctcagcagaccagcagtggaacaacaataaa
ccttaccgaccctgggggacagaagttggagctttttaa (SEQ ID NO:123)

[00412] Homo sapiens forkhead box C1 (FOXC1) (NM_001453) SEQ ID NO:60 (protein) and
SEQ ID NO:124 (DNA)

MQARYSVSSPNSLGVVPYLGGEQSYRRAAAAAAGGGYTAMPAPMSVYSHPAHAEQYPGGMARAYG
PYTPQPQPKDMVKPPYSYIALITMAIQNAPDKKITLNGIYQFIMDRFPFYRDNKQGWQNSIRHNL
SLNECFVKVPRDDKKPGKGSYWTLDPDSYNMFENGSLRRRRRFRKKDAVKDKKEEKDRHLHLKEPP
25 PPGRQPPPAPPEQADGNAPGPQPPPVRIQDIKTENGTCPSPPQPLSPAAALGSGSAAAVPKIESP
DSSSSSLSSGSSPPGSLPSARPLSLDGADSAPPPAPSPAPPPHHSQGFSVDNIMTSLRGSPQSAA
AELSSGLLASAAASSRAGIAPPLALGAYSPGQSSLYSSPCSQTSSAGSSGGGGGGGAGAAGGAGGA
GTYHCNLQAMSLYAAGERGGHLQGAPGGAGGS AVDDPLPDYSLPVTSSSSSSSLSHGGGGGGGGG
GQEAGHHPAAHQRLTSWYLNQAGGDLGHLASAAAAAAGYPGQQQNFHSVREMFESQRIGLNN
30 SPVNGNSSCQMAFPSSQSLYRTSGAFVYDCSKF (SEQ ID NO:60)

atgcaggcgcgctactcgtgtccagccccaactccctgggagtggtgccctacctcggcgggcga
gcagagctactaccgcgcggcgggcgcgggcgccggggcggtacaccgccatgccggccccc
tgagcgtgtactcgcaacctgcgcacgcgcgagcagtagccggcggtatggccgcgcctacggg
ccctacacgcgcgcagccgcagcccaaggacatggtgaagcgcctatagctacatcgcgctcat
35 caccatggccatccagaacgccccgggacaagaagatcaccctgaacggcatctaccagttcatca
tggaccgcttcccccttctaccgggacaacaagcagggctggcagaacagcatccgccacaacctc
tcgctcaacgagtgccttcgtcaaggtgccgcgcgacgacaagaagccggggcaagggcagctactg
gacgctggaccggactcctacaacatgttcgagaacggcagcttctcgcgcgcgcgcgcgct
tcaagaagaaggacgcggtgaaggacaaggaggagaaggacaggtgcacctcaaggagccgccc
40 ccgcccggcgccagcccccgccgcgcgcggagcaggccgacggcaacgcgcgccggtccgca
gccgcccgcctgcgcatccaggacatcaagaccgagaacgggtacgtgcccctcgccgccccagc
ccctgtccccggcgccgcgcctgggcagcggcagcgccgcgcggtgcccagatcgagagcccc
gacagcagcagcagcagcctgtccagcgggagcagcccccgggcagcctgccgtcggcgcgggc
gctcagcctggacgggtgcggattccgcgcgcgcgcgcgcgcgcctccgcgccgcgcgcgcacc
45 atagccagggttcagcgtggacaacatcatgacgtcgctgcgggggtcgccgcagagcgcggcc
gcggagctcagctccggccttctggcctcgggcgccgcgtcctcgcgcgcggggatcgcaacccc
gctggcgctcggcgcctactcgccccggccagagctccctctacagctccccctgcagccagacct
ccagcgcgggcagctcgggcgggcgggcgggcgggcgggggcgcgggggcgcgggcgggcgcc
gggacctaccactgcaacctgcaagccatgagcctgtacgcggccggcgagcgcggggggccactt
50 gcaggcgcgccccggggcgcgggcgggctcggcgctggacgacccccctgcccgactactctctgc

10 APKDLVKPPYSYIALITMAIQNAPEKKITLNGIYQFIMDRFPFYRENKQGWQNSIRHNLSLNECF
 VKVPRDDKKPGKGSYWTLDPSYNYMFENGSLRRRRRFFKKKDVSKKEERAHLKEPPPAASKGAP
 ATPHLADAPKEAEKKVVIKSEAASPALPVITKVETLSPESALQGSRSAASTPAGSPDGLPEHH
 AAPNGLPGFSVENIMTLRTSPGGELSPGAGRAGLVVPLALPYAAAPPAAYGQPCAQGLEAGA
 AGGYQCSMRAMSLYTGAERPAHMCVPPALDEALSDHPSGPTSPLSALNLAAGQEGALAATGHHHQ
 15 HHGHHHPQAPPPPPAPQPQPTPQGAAAQAASWYLNHSGDLNHLPGHTFAAQQQTFPNVREMFN
 SHRLGIENSTLGESQVSGNASCQLPYRSTPPLYRHAAPYSYDCTKY (SEQ ID NO:61)

[00414] SEQ ID NO:62: Homo sapiens SRY (sex determining region Y)-box 7 (SOX7)
NM_031439 SEQ ID NO:62 (protein) and SEQ ID NO:126 (DNA)

MASLLGAYPWPEGLECPALDAELSDGQSPPAVPRPPGDKGSESRIRRPMNAFMVWAKDERKRLAV
QNPDLHNAELSKMLGKSWKALTLSQKRYPVDEAERLRLQHMQDYPNYKYRPRRKKQAKRLCKRVD
PGFLLSSLSRDQNALPEKRSGSRGALGEKEDRGEYSPGTALPSLRGCYHEGPAGGGGGGGTPSSVD
TYPYGLPTPPEMSPLDVLEPEQTFSSPCQEEHGHPRRIPHLPGHPYSPEYAPSPLHCSHPLGSL
ALGQSPGVSMSPVPGCPPSPAYYSPATYHPLHSNLQAHLGQLSPPEHPGFDALDQLSQVELLG
DMDRNEFDQYLNTPGHPDSATGAMALSGHVPVSQVTPGTPTETSLISVLADATATYYNSYSVS
(SEQ ID NO:62)

atggcttcgctgctgggagcctacccttgccccgaggggtctcgagtgcccgccctggacgccga
 gctgtcggatggacaatcgccgccggccgtcccccgccccgggggacaagggctccgagagcc
 gtatccggcgcccatgaacgccttcatggtttgggccaaggacgagaggaaacggctggcagtg
 cagaacccggacctgcacaacgccgagctcagcaagatgctgggaaagtctggaaggcgctgac
 5 gctgtcccagaagaggccgtacgtggacgagggcgagcggtgcgcctgcagcacatgcaggact
 accccaactacaagtaccggccgcgcaggaagaagcaggccaagcggtgtgcaagcgctggac
 ccgggcttccttctgagctccctctcccgggaccagaacgccctgccggagaagagaagcggcag
 ccggggggcgctgggggagaaggaggacaggggtgagtactccccggcactgccttgcacagcc
 tccggggctgctaccacgaggggcccggctggtggtggcgggcgccaccccgagcagtgtggac
 10 acgtaccgcgtacgggctgccacacctcctgaaatgtctcccctggacgtgctggagccggagca
 gaccttcttctcctccccctgccaggaggagcatggccatccccgccgcacccccacctgccag
 ggcacccgtactcacgggagtacgccccaaagccctctccactgtagccacccccctgggctccctg
 gcccttggccagtcccccggcgctctccatgatgtcccctgtaccgcggtgtcccccatctcctgc
 ctattactccccggccacctaccacccactccactccaacctccaagcccacctggggccagcttt
 15 cccgcctcctgagcacccctggcttcgacgccctggatcaactgagccaggtggaactcctgggg
 gacatggatcgcaatgaattcgaccagtatttgaacactcctggccaccagactccgccacagg
 ggccatggccctcagtgggcatgttcgggtctccaggtgacaccaacgggtcccacagagacca
 gcctcatctccgtcctggctgatgccacggccacgtactacaacagctacagtgtgtcatag
 (SEQ ID NO:126)
 20 **[00415]** Homo sapiens SOX18 (AB033888) SEQ ID NO:63 (protein) and SEQ ID NO:127
 (DNA)
 MQRSPPGYGAQDDPPARRDCAWAPGHGAAADTRGLAAGPAALAAPAAPASPPSPQRSPPRSPEPG
 RYGLSPAGRGERQAADSRIRPMNAFMVWAKDERKRLAQONPDHLNAVLSKMLGKAWKELNAAE
 25 KRPFVEEAERLRVQHRLDHPNYKYRPRRKKQARKARRLEPGLLLPGLAPPQPPPEPFPAASGSAR
 AFRELPLGAEFDGLGLPTPERSPLDGLEPGEAAFFPPPAAPEDCALRPFRAPIYAPTELSRDPGG
 CYGAPLAEALRTAPPAAPLAGLYYGTLGTPGPYPGPLSPPEAPPLESAEPLGPAADLWADVDLT
 EFDQYLNCSTRTPDAPGLPYHVALAKLGPRAMSCPEESSLISALSDASSAVVYSACISG (SEQ
 ID NO:63)
 30 atgcagagatcgccgcccggctacggcgccacaggacgacccgcccggcccgccgactgtgcatg
 ggccccgggacacggggccgcccgtgacacgcgcggcctcgccgcccggccccggccctcgccg
 cgcccgccgcgcccgcctcgccgcccagccgcagcgcagtccccgcgcagccccgagccgggg
 cgctatggcctcagcccggccggccgcggggaacgccaggcgagcagagtcgcgcacccggcg
 gcccatgaacgccttcatggtgtgggcaaaggacgagcgcaagcggtggctcagcagaacccgg
 35 acctgcacaacgcggtgctcagcaagatgctgggcaaagcgtggaaggagctgaacgcggcggag
 aagcgggcccttcgtggaggaagccgaacggctgcgcgtgcagcacttgcgcgaccaccccaacta
 caagtaccggccgcgcgcgaagaagcaggcgcgcaaggccccggcggtggagccccggcctcctgc
 tccccgggattagcgcccccgcagccaccgcccagcctttccccgcggcgctgtggctcggtcgc
 gccttcgcgcagctgcccccgctgggcgcgcaggttcgacggcctggggctgccacgcccagcgc
 ctgcctctggacggcctggagccccggcgaggtgccttcttccaccgcccgcggcgcccaggg
 40 actgcgcgtgcggcccttcgcgcgcctacgcgccaccgagttgtcgcgggacccccggcggt
 tgctacggggctcccctggcgaggcgctcaggaccgcgcccccccgcgcgcgctcgtggcct
 gtactacggcaccttgggacgcccggccgtaccccgcccgctgtcgccgcccggccgagggccc
 cgccgctggagagcgccgagccgctggggccccgcgcgcgatctgtgggcccagctggacctcacc
 gagttcgaccagtagctcaactgcagccggactcggcccagcgcggggctcccgtagcacgt
 45 ggcactggccaaactgggcccgcgcgcctatgtcctgcccagaggagagcagcctgatctccgcgc
 tgtcggacgccagcagcgcggtctattacagcgcggtgcatctccggctag (SEQ ID
 NO:127)

[00416] Any patents or publications mentioned in this specification are incorporated herein by reference to the same extent as if each individual publication is specifically and individually

indicated to be incorporated by reference. U.S. provisional patent application Serial No. 61/895,562, filed October 25, 2013 is incorporated herein by reference.

[00417] The compositions and methods described herein are presently representative of preferred embodiments, exemplary, and not intended as limitations on the scope of the invention.

5 Changes therein and other uses will occur to those skilled in the art. Such changes and other uses can be made without departing from the scope of the invention as set forth in the claims.

CLAIMS

1. A pharmaceutical composition comprising a protein transduction reagent-modified reprogramming protein for use in therapy, wherein the protein transduction reagent is non-covalently bound to the reprogramming protein and wherein the protein transduction reagent comprises a cation reagent and a lipid; wherein the cation reagent is a polyethylenimine (PEI); and wherein the lipid is selected from the group consisting of DOTMA, (N-1(-(2,3-dioleoyloxy)propyl-N,N,N-trimethyl-ammonium chloride; DOGS, (dioctadecylamido-glycylspermine); DOTAP, 1,2-dioleoyl-3-trimethylammonium-propane; DOPE, 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine; POPC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; DMPE 1,2-dimyristoyl-sn-glycero-3-phosphocholine and combinations thereof.

2. The composition of claim 1, wherein the polyethylenimine (PEI) is selected from the group consisting of PEI Mw 1,200 (PEI 1.2K), PEI Mw 2000 (PEI 2K), PEI Mw 4000 (PEI 4K) and PEI Mw 8000 (PEI 8K).

3. Use of a protein transduction reagent-modified reprogramming protein in the manufacture of a medicament for the treatment of a condition characterized by damaged, and/or defective cells in a subject, wherein the protein transduction reagent is non-covalently bound to the reprogramming protein and wherein the protein transduction reagent comprises a cation reagent and a lipid; wherein the cation reagent is a polyethylenimine (PEI); and wherein the lipid is selected from the group consisting of DOTMA, (N-1(-(2,3-dioleoyloxy)propyl-N,N,N-trimethyl-ammonium chloride; DOGS, (dioctadecylamido-glycylspermine); DOTAP, 1,2-dioleoyl-3-trimethylammonium-propane; DOPE, 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine; POPC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; DMPE 1,2-dimyristoyl-sn-glycero-3-phosphocholine and combinations thereof; and

wherein the condition is selected from the group consisting of cancer, heart disease, liver disease, atherosclerosis, neurological disease, neurodegenerative disease, blood disease, diabetes, pancreatic disease, obesity, muscle disease, breast disease and vascular disease.

4. The use of claim 3, wherein the condition is cancer.

5. The use of claims 3 or 4, wherein the condition is cancer and the protein transduction reagent-modified reprogramming protein is selected from the group consisting of: protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog.

6. The use of claims 3 or 4, wherein the condition is cancer and the protein transduction reagent-modified reprogramming protein comprises protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog.

7. The use of any one of claims 3-5, wherein the cancer is a brain tumor or breast cancer or pancreatic cancer and the protein transduction reagent-modified reprogramming protein is selected from the group consisting of:

5 protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog.

8. The use of claims 3, 4 or 6, wherein the cancer is a brain tumor or breast cancer or pancreatic cancer and the protein transduction reagent-modified reprogramming protein comprises
10 protein transduction reagent-modified Sox2, protein transduction reagent-modified Oct4 and protein transduction reagent-modified Nanog.

9. The use of claim 3, wherein the condition is selected from the group consisting of:
15 heart disease, liver disease, atherosclerosis, neurological disease, neurodegenerative disease, blood disease, diabetes, pancreatic disease, obesity, muscle disease, breast disease and vascular disease.

10. The use of claim 9, wherein the heart disease is selected from the group consisting of: atherosclerosis, myocardial infarction, and chronic heart failure.

11. The use of claim 9, wherein the liver disease is selected from the group consisting of: familial hyper-cholesterolaemia (FH), Crigler-Najjar syndrome and hereditary tryosinemia I.

5 12. The use of claim 9, wherein the neurological or neurodegenerative disease is selected from the group consisting of: spinal cord injury, traumatic brain injury, amyotrophic lateral sclerosis, spinal muscular atrophy and Parkinson's.

10 13. The use of claim 3, wherein the condition is a heart disease or heart damage and the protein transduction reagent-modified reprogramming protein comprises one or more of: protein transduction reagent-modified Gata4, protein transduction reagent-modified Hand2, protein transduction reagent-modified MEF2c and protein transduction reagent-modified Tbox5; wherein the condition is a liver disease or liver damage and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Hnf1a, protein transduction reagent-modified
15 Foxa1, protein transduction reagent-modified Foxa2 and protein transduction reagent-modified Foxa3; wherein the condition is a liver disease or liver damage and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Gata4, protein transduction reagent-modified Hnf1a and protein transduction reagent-modified Foxa3; wherein the condition is atherosclerosis and the pharmaceutical composition comprises one or more of:
20 protein transduction reagent-modified CEBP α , protein transduction reagent-modified CEBP β and protein transduction reagent-modified PU.1; wherein the condition is a neurodegenerative disease or neuronal tissue damage and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Sox2 and protein transduction reagent-modified Foxg1; one or more of: protein transduction reagent-modified Asc1,
25 protein transduction reagent-modified Lmx1a, protein transduction reagent-modified Nurr1, protein transduction reagent-modified Brn2 and protein transduction reagent-modified Myt11; one or more of: protein transduction reagent-modified Asc1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Myt11 and protein transduction reagent-modified NeuroD1; protein transduction reagent-modified Ngn2; one or both of: protein transduction reagent-modified Sox2 and protein transduction reagent-modified NeuroD1; one or more of:
30 protein transduction reagent-modified Brn2, protein transduction reagent-modified Asc11, protein transduction reagent-modified Myt11, protein transduction reagent-modified Lhx3 and protein transduction reagent-modified Hb9; one or more of: protein transduction reagent-modified Brn2, protein transduction reagent-modified Asc11, protein transduction reagent-modified Myt11, protein
35 transduction reagent-modified Lhx3, protein transduction reagent-modified Hb9, protein transduction reagent-modified

Lsl1 and protein transduction reagent-modified Ngn2; protein transduction reagent-modified Dlx2; protein transduction reagent-modified Dlx2 and protein transduction reagent-modified Ascl1; one or more of: protein transduction reagent-modified Ascl1, protein transduction reagent-modified Brn2, protein transduction reagent-modified Myt1l, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Foxa2; or one or more of: protein transduction reagent-modified Ascl1, protein transduction reagent-modified Lmx1a and protein transduction reagent-modified Nurr1; wherein the condition is a disease or disorder of the blood and the pharmaceutical composition comprises: protein transduction reagent-modified Oct4; wherein the condition is diabetes, a pancreatic disease or pancreatic tissue damage and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Ngn3, protein transduction reagent-modified Pdx1 and protein transduction reagent-modified Pax4 or one or more of: protein transduction reagent-modified Ngn3, protein transduction reagent-modified Pdx1 and protein transduction reagent-modified MafA; wherein the condition is obesity and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Prdm16 and protein transduction reagent-modified C/EBPb; wherein the condition is a muscle disease or muscle damage and the pharmaceutical composition comprises protein transduction reagent-modified MyoD; wherein the condition is arthritis or joint disease or injury and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Runx2, protein transduction reagent-modified Sox5 and protein transduction reagent-modified Sox6; wherein the condition is a breast disease or breast tissue damage and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Sox9, protein transduction reagent-modified Slug, protein transduction reagent-modified Stat5a, protein transduction reagent-modified Gata3 and protein transduction reagent-modified Brca-11; wherein the condition is a vascular disease or blood vessel damage and the pharmaceutical composition comprises one or more of: protein transduction reagent-modified Erg1, protein transduction reagent-modified Er71, protein transduction reagent-modified Fli1 and protein transduction reagent-modified Gata2 or one or more of: protein transduction reagent-modified Hey1, protein transduction reagent-modified Hey2, protein transduction reagent-modified FoxC1 and protein transduction reagent-modified FoxC2 or one or both of: protein transduction reagent-modified Sox7 and protein transduction reagent-modified Sox18.

14. The use of any one of claims 3 to 13, wherein the cation reagent is a polyethylenimine (PEI); and

wherein, the lipid is selected from the group consisting of DOTMA, (N-1-(2,3-dioleoyloxy)propyl)-N,N,N-trimethyl-ammonium chloride; DOGS, (dioctadecylamido-glycylspermine); DOTAP, 1,2-dioleoyl-3-trimethylammonium-propane; DOPE, 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine; POPC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; DMPE 1,2-dimyristoyl-sn-glycero-3-phosphocholine and combinations thereof.

5 15. The use of any one of claims 3 to 14, wherein the medicament is to be administered
by parenteral, oral, rectal, nasal, pulmonary, epidural, ocular, otic, intraarterial, intracardiac,
intracerebroventricular, intracranial, intradermal, intravenous, intramuscular, intraperitoneal,
intraosseous, intrathecal, intratumoral, intravesical, subcutaneous, topical, transdermal, or
10 transmucosal administration.

- 5 16. The use of any one of claims 3-15 wherein the protein transduction reagent-modified reprogramming protein is a reprogramming protein treated with a protein transduction reagent selected from the group consisting of: QQ1a, QQ2a, QQ3a, QQ4a, QQ5a, QQ6a, QQ7a, QQ8a and QQ9a as defined in the table below

Protein transduction reagent (QQ reagent)	PEI 1.2K	PEI 2K	PEI 4K	PEI 8K	DOTAP or DOTMA	DOPE or DOGS	POPC	DMPE
QQ1a	10-200µl	-	-	-	25-100µl	25-100µl	-	-
QQ2a	10-200µl	-	10-100µl	-	25-100µl	25-100µl	-	-
QQ3a	10-200µl	-	10-100µl	10-100µl	25-100µl	25-100µl	-	-
QQ4a	10-200µl	10-100µl	-	10-100µl	25-100µl	25-100µl	-	-
QQ5a	10-200µl	10-100µl	10-100µl	10-100µl	25-100µl	25-100µl	-	-
QQ6a	10-200µl	10-100µl	10-100µl	10-100µl	-	--	25-100µl	25-100µl
QQ7a	10-200µl	10-100µl	10-100µl	10-100µl	25-100µl	25-100µl	25-100µl	25-100µl
QQ8a	10-200µl	-	-	10-100µl	25-100µl	25-100µl	25-100µl	25-100µl
QQ9a	10-200µl	-	10-100µl	10-100µl	-	-	25-100µl	25-100µl

* PEI: polyethylenimine

DOTMA: (N-1-(2,3-dioleoyloxy)propyl-N,N,N-trimethyl-ammonium chloride;

DOGS: (dioctadecylamido-glycylspermine);

5 DOTAP: 1,2-dioleoyl-3-trimethylammonium-propane;

DOPE: 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine;

POPC: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine;

DMPE: 1,2-dimyristoyl-sn-glycero-3-phosphocholine.