METHOD FOR INDUCING PLURIPOTENCY IN CELLS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims the benefit of U.S. Provisional Patent Application No. 61/104,231, filed October 9, 2008, the contents of which are incorporated by reference in their entirety.

BACKGROUND OF THE INVENTION

[0002] **1. Field of the Invention:**

[0003] The present application relates to the field of inducing pluripotency in cells.

[0004] **2. General Background and State of the Art:**

[0005] It has been demonstrated, in mouse and human, that somatic cells can be reprogrammed by ectopic expression of transcription factors (Lowry et al., 2008; Maherali et al., 2007; Nakagawa et al., 2008; Okita et al., 2007; Park et al., 2008; Takahashi et al., 2006; Takahashi and Yamanaka, 2006; Wernig et al., 2007; Yu et al., 2006) to become pluripotent. The generation of induced pluripotent stem (iPS) cells holds great promise for the realization of truly personalized regenerative medicine (Yamanaka, 2007; Jaenish and Young, 2008) because stem cells derived from a patient's own skin cell can be used to generate cells and tissues to repair damage caused by disease or aging. Forced expression of combinations of the transcription factors, Oct4, Sox2, Klf4 and c-Myc or Oct4, Sox2, Nanog and Lin28 have been shown to cause mature cells to revert to the pluripotent state.

[0006] In earlier studies, the transcription factors were expressed using multiple viral vectors (Takahashi and Yamanaka, 2006; Okita et al., 2007; Maherali et al., 2007; Wernig et al., 2007; Takahashi et al., 2006;Yu et al., 2006; Park et al., 2008). The use of multiple vectors presented a problem because of multiple integration events, which could lead to increased risk of oncogenicity (Takahashi and Yamanaka, 2006; Aoi et al., 2008). Researchers have tried to overcome this problem by using single vector systems (Sommer et al., 2009), excisable vectors (Kaji et al., 2009; Soldner et al., 2009; Woltjen et al., 2009), non-integrating vectors (Stadtfeld et al., 2009; Yu et al., 2009) and transient transfections (Okita et al., 2009). However, these methods are extremely inefficient at achieving epigenetic reprogramming.

[0007] Methods for inducing pluripotency include transfection of the oncogene c-Myc, which is undesirable because of its potential to cause cancer. iPS cells can be generated without transfecting c-Myc (Nakagawa et al., 2008; Wernig et al., 2008). However, the efficiency of reprogramming was greatly decreased. Similarly, Klf4 can induce dysplasia (Foster et al., 2005).

[0008] Because of the problems associated with multiple viral vector integration and undesirable side effects of some of the genes that induce pluripotency, there is a need to replace the use of some or all of the pluripotency-inducing genes with the protein gene products and proteins that regulate their expression or whose expression is regulated by the pluripotency-inducing genes or small molecules that regulate the expression of genes or proteins that induce pluripotency. To this end, it has been reported that the introduction of the gene products, rather than the genes, also induced pluripotency (Zhou et al., 2009). Recombinant Oct4, Sox2, Klf4 and c-Myc, tagged with poly-arginine to facilitate entry into the cell, reprogrammed mouse somatic cells. Others have used small molecules to replace the need for one of the genes of the core set. A small molecule that upregulated Nanog eliminated the need for the Klf4 gene, which also upregulates Nanog (Lyssiotis et al., 2009). In another study, a small molecule HDAC inhibitor removed the requirement for both Klf4 and c-Myc (Huangfu et al., 2008, a&b). These studies show that: 1) the protein gene products can replace the need for the genes; 2) small molecules that upregulate genes can replace the need for the genes; and 3) genes (or gene products) in the same regulatory pathway can substitute for one another.

[0009] Despite these achievements, a major problem that remains is that these methods suffer from low efficiency of reprogramming. Current rates of inducing pluripotency in somatic cells are so low that they make therapeutic uses of iPS cells impractical. Therefore, what is needed is to identify proteins and small molecules that either alone or in addition to those already identified, induce pluripotency or improve the efficiency of the induction of pluripotency in cells.

SUMMARY OF THE INVENTION

[0010] In one aspect, the present invention is directed to a method for inducing or maintaining pluripotency in a cell comprising contacting the cell with a biological or chemical species that increases MUC1* activity. The biological species may be a gene or a protein. The protein may be preferably a MUC1* ligand, preferably the ligand is NM23. The protein may also

be an antibody, preferably bi-valent antibody that specifically recognizes the PSMGFR sequence of MUC1*. In another aspect, the chemical species may be a small molecule. Preferably, the small molecule may enhance the transcription of MUC1, its cleavage enzyme, or NM23. Preferably, the cleavage enzyme may be MMP-16, MMP-14 or ADAM-17. Further, the small molecule may enhance cleavage of MUC1. The small molecule may be a phorbol ester.

[0011] In another aspect, the gene above may encode MUC1 or MUC1*, or an effective fragment thereof, or a ligand of MUC1*, such as but not limited to a MUC1* antibody or NM23. Preferably, the fragment may be the cytoplasmic tail of MUC1 or MUC1*

[0012] In further other aspects, the invention is directed to contacting the cell with an additional molecule that increases expression of gene products that induce pluripotency, such as but not limited to a molecule that increases expression of OCT4, SOX2, NANOG, KLF4 or LIN28, preferably of OCT4, and SOX2.

[0013] These and other objects of the invention will be more fully understood from the following description of the invention, the referenced drawings attached hereto and the claims appended hereto.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014] The present invention will become more fully understood from the detailed description given herein below, and the accompanying drawings which are given by way of illustration only, and thus are not limitative of the present invention, and wherein;

[0015] FIGURES 1A-1F show that MUC1* increases growth rate. A. Clonogenic assay shows that transfecting rat fibroblasts (3Y1) with MUC1* increases growth rate but MUC1 (full-length) does not; B. MUC1* activity increases survival. Breast cancer cells that have acquired resistance to TAXOL® do so by increasing MUC1* expression. Treatment with anti-MUC1* Fab reverses acquired resistance to TAXOL®-induced cell death; C. Ligand-induced dimerization of MUC1* extracellular domain stimulates growth. The addition of bivalent anti-MUC1* antibodies stimulates the growth of MUC1*-positive cells. Blocking with the anti-MUC1* (mv) Fab inhibits cell growth. Bell-shaped growth curve is characteristic of receptor dimerization. Growth of control MUC1-negative HEK 293 cells was not affected; D. Suppression of MUC1*, using specific siRNA, abolishes the growth stimulatory effects of adding a MUC1* dimerizing ligand; E. NM23 is the native MUC1* activating ligand. NM23 stimulates growth of MUC1*-

positive cancer cells and produces bell-shaped curve indicative of receptor dimerizatio. Effect is abolished by siRNA suppression of MUC1; F. Direct binding of NM23 to the MUC1* peptide is detected by SPR. 15nM NM23 binds to MUC1* extracellular domain peptide but not to irrelevant peptide. Measurements were done using SPR (surface plasmon resonance) and NTA-Ni-SAM coated Au chips.

[0016] FIGURES 2A-2F show that MUC1 is cleaved on undifferentiated hESCs but MUC1 is not cleaved on differentiated hESCs. Immunocytochemistry shows that undifferentiated (pluripotent) stem cells express MUC1* and not the full-length protein; OCT4 is the gold standard marker for pluripotency. All pluripotent stem cells are MUC1*-positive. However, as soon as differentiation initiates (loss of OCT4 expression), cleavage stops and only full-length MUC1 (MUC1-FL) is detected. Panels A-C are photos of the same undifferentiated stem cell colony stained with: A. anti-MUC1* antibody that recognizes the PSMGFR peptide; B. anti-OCT4; C. anti-MUC1 full-length VU4H5. Panels D-F are photos of the same newly differentiated stem cell colony stained with: D. anti-MUC1* antibody that recognizes the PSMGFR peptide; E. anti-OCT4; F. anti-MUC1 full-length VU4H5.

[0017] FIGURE 3A-3F shows that NM23 (MUC1* ligand) co-localizes with MUC1* and OCT4 on undifferentiated hESCs, but not on differentiated cells. Immunocytochemistry shows that undifferentiated (pluripotent) stem cells express MUC1* and its activating ligand NM23. However, when stem cells begin to differentiate (loss of OCT4 expression), then MUC1 is expressed as the full-length protein and NM23 is no longer secreted. Dotted lines indicate the border between the undifferentiated and the newly differentiating portions. Panels A-C are photos of the same undifferentiated stem cell colony stained with: A. an antibody that recognizes NM23; B. anti-MUC1* antibody that recognizes the PSMGFR peptide; C. an overlay of (A), (B) and the same cells stained with DAPI to stain nuclei. Panels D-F are photos of the same undifferentiated stem cell colony stained with: D. an antibody that recognizes NM23; E. an antibody that recognizes OCT4; F. an overlay of (D), (E) and the same cells stained with DAPI to stain nuclei.

[0018] FIGURES 4A-4B show that stimulation of MUC1* promotes growth and inhibits differentiation of hESCs in the absence of bFGF and conditioned media. Ligand-induced dimerization of MUC1* extracellular domain produced, using bivalent anti-MUC1* antibody, essentially 100% pluripotent colonies after 5 weeks growth in minimal media without adding

bFGF or conditioned media. Colonies were grown on matrigel. The same results were obtained when NM23 or NM23 S120G mutant was used to activate MUC1*. Panels A-D are photos of wells where cell growth medium was supplemented with conditioned medium from fibroblast feeder cells plus either anti-MUC1* or bFGF. Panels E-H are photos of wells where stem cells were cultured in minimal medium plus either anti-MUC1* or bFGF. Images are of cells stained with: A. antibody that recognizes OCT4; B. DAPI staining of cells of (A); C. anti-MUC1*; D. DAPI staining of cells of (C); E. antibody that recognizes OCT4; F. DAPI staining of cells of (E); G. anti-MUC1*; H. DAPI staining of cells of (G).

[0019] FIGURE 5 shows a bar graph indicating that MUC1* activity is required for pluripotent stem cell growth. Blocking MUC1* with anti-MUC1* Fab caused total stem cell death within 8-12 hours even though bFGF and conditioned media (CM) was present. Bivalent anti-MUC1* stimulated growth. Cells cultured 25 hrs; live cells were measured in a Calcein fluorescent assay.

[0020] FIGURES 6A-6B show photos evidencing that MUC1* translocates to the nucleus of cells. A MUC1* Fab was fluorescently labeled (red) then incubated with MUC1*-positive cells. The photos show that, initially, MUC1* is uniformly distributed on the cell surface. However, after 40 minutes, MUC1* is concentrated in the nucleus. For comparison cells were also stained with a fluorescently labeled antibody (green) that recognizes EEA1, which remains uniformly distributed in cytoplasm throughout the experiment. A. photo of cells taken at time zero; B. photo of cells taken 40 minutes after the addition of the Fab of anti-MUC1*.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0021] In the present application, "a" and "an" are used to refer to both single and a plurality of objects.

[0022] As used herein, "increasing MUC1* activity" refers to directly or indirectly increasing MUC1* signaling, and includes without limitation the dimerization of MUC1* receptor and also increased production of MUC1* by cleavage of the MUC1 receptor. MUC1* activity may be also increased by higher transcriptional expression of MUC1 receptor, which is further cleaved and dimerized. Therefore, in one aspect, MUC1* activity may be increased by a higher activity of the effector molecule that dimerizes MUC1*, or the higher activity of the cleavage molecule that cleaves MUC1 so that MUC1* is formed, or increased expression of the

MUC1. Therefore, any chemical or biological species that is able to increase the activity of the MUC1* dimerizing ligand, MUC1 cleavage enzyme to form MUC1*, or any transcriptional activator that enhances expression of MUC1, is encompassed as a species that "increases MUC1* activity".

[0023] As used herein, "MUC1 Growth Factor Receptor" (MGFR) is a functional definition meaning that portion of the MUC1 receptor that interacts with an activating ligand, such as a growth factor or a modifying enzyme such as a cleavage enzyme. The MGFR region of MUC1 is that extracellular portion that is closest to the cell surface and is defined by most or all of the PSMGFR, as defined below. The MGFR is inclusive of both unmodified peptides and peptides that have undergone enzyme modifications, such as, for example, phosphorylation, glycosylation and so forth.

[0024] As used herein, "Primary Sequence of the MUC1 Growth Factor Receptor" (PSMGFR) refers to peptide sequence that defines most or all of the MGFR in some cases, and functional variants and fragments of the peptide sequence. The PSMGFR is defined as SEQ ID NO:6, and all functional variants and fragments thereof having any integer value of amino acid substitutions up to 20 (i.e. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20) and/or any integer value of amino acid additions or deletions up to 20 at its N-terminus and/or C-terminus. A "functional variant or fragment" in the above context refers to such variant or fragment having the ability to specifically bind to, or otherways specifically interact with, ligands that specifically bind to, or otherwise specifically interact with, the peptide of SEQ ID NO:6, while not binding strongly to identical regions of other peptide molecules identical to themselves, such that the peptide molecules would have the ability to aggregate (i.e. self-aggregate) with other identical peptide molecules. One example of a PSMGFR that is a functional variant of the PSMGFR peptide of SEQ NO:6 is SEQ ID NO:8, which differs from SEQ ID NO:6 by including an -SPY- sequence instead of the -SRY-.

[0025] As used herein, "MUC1*" refers to the MUC1 protein with the N-terminus truncated such that the extracellular domain is essentially comprised of the PSMGFR (SEQ ID NO:5).

[0026] As used herein "MUC1* associated factors" refers to agents that modify, activate, modulate the activity of, or modulate the expression of MUC1*. MUC1* associated factors include, without limitation, agents that affect dimerization of MUC1* receptor, increased production of MUC1*, induce cleavage of the MUC1 receptor, agents that increase MUC1*

activity by higher transcriptional expression of MUC1 receptor, which is further cleaved and dimerized.

[0027] As used herein, "effective amount" is an amount sufficient to effect beneficial or desired clinical or biochemical results. An effective amount can be administered one or more times. For purposes of this invention, an effective amount of an inhibitor compound is an amount that is sufficient to induce or maintain pluripotency of a cell or activate MUC1*

[0028] As used herein, "fragments" or "functional derivatives" refers to biologically active amino acid sequence variants and fragments of the native ligands or receptors of the present invention, as well as covalent modifications, including derivatives obtained by reaction with organic derivatizing agents, post-translational modifications, derivatives with nonproteinaceous polymers, and immunoadhesins.

[0029] As used herein, "ligand" refers to any molecule or agent, or compound that specifically binds covalently or transiently to a molecule such as a polypeptide. When used in certain context, ligand may include antibody. In other context, "ligand" may refer to a molecule sought to be bound by another molecule with high affinity, such as but not limited to a natural or unnatural ligand for MUC1* or a cleaving enzyme binding to MUC1 or MUC1* or a dimerizing ligand for MUC1*.

[0030] As used herein, the term "specifically binds" refers to a non-random binding reaction between two molecules, for example between an antibody molecule immunoreacting with an antigen, or a non-antibody ligand reacting with another polypeptide, such as NM23 specifically binding with MUC1* or an antibody binding to MUC1* or a cleaving enzyme binding to MUC1 or MUC1*

[0031] As used herein, "vector", "polynucleotide vector", "construct" and "polynucleotide construct" are used interchangeably herein. A polynucleotide vector of this invention may be in any of several forms, including, but not limited to, RNA, DNA, RNA encapsulated in a retroviral coat, DNA encapsulated in an adenovirus coat, DNA packaged in another viral or viral-like form (such as herpes simplex, and adeno-associated virus (AAV)), DNA encapsulated in liposomes, DNA complexed with polylysine, complexed with synthetic polycationic molecules, complexed with compounds such as polyethylene glycol (PEG) to immunologically "mask" the molecule and/or increase half-life, or conjugated to a non-viral protein. Preferably, the polynucleotide is DNA. As used herein, "DNA" includes not only bases A, T, C, and G, but also includes any of

their analogs or modified forms of these bases, such as methylated nucleotides, internucleotide modifications such as uncharged linkages and thioates, use of sugar analogs, and modified and/or alternative backbone structures, such as polyamides.

[0032] Sequence Listing Free Text

[0033] As regards the use of nucleotide symbols other than a, g, c, t, they follow the convention set forth in WIPO Standard ST.25, Appendix 2, Table 1, wherein k represents t or g; n represents a, c, t or g; m represents a or c; r represents a or g; s represents c or g; w represents a or t and y represents c or t.

MTPGTQSPFF LLLLLTVLTV VTGSGHASST PGGEKETSAT QRSSVPSSTE KNAVSMTSSV
LSSHSPGSGS STTQGQDVTL APATEPASGS AATWGQDVTS VPVTRPALGS TTPPAHDVTS
APDNKPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS
APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS
APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS
APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS
APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS
APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS
APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS
APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS
APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS
APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS
APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS
APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS
APDTRPAPGS TAPPAHGVTS APDTRPAPGS TAPPAHGVTS APDNRPALGS TAPPVHNVTS
ASGSASGSAS TLVHNGTSAR ATTTPASKST PFSIPSHHSD TPTTLASHST KTDASSTHHS
SVPPLTSSNH STSPQLSTGV SFFFLSFHIS NLQFNSSLED PSTDYYQELQ RDISEMFLQI
YKQGGFLGLS NIKFRPGSVV VQLTLAFREG TINVHDVETQ FNQYKTEAAS RYNLTISDVS
VSDVPFPFSA QSGAGVPGWG IALLVLVCVL VALAIVYLIA LAVCQCRRKN YGQLDIFPAR
DTYHPMSEYP TYHTHGRYVP PSSTDRSPYE KVSAGNGGSS LSYTNPAVAA

ASANL (SEQ ID NO: 1) describes full-length MUC1 Receptor (Mucin 1 precursor, Genbank Accession number: P15941).

MTPGTQSPFFLLLLLTVLT (SEQ ID NO: 2)
MTPGTQSPFFLLLLLTVLT VVTA (SEQ ID NO: 3)
MTPGTQSPFFLLLLLTVLT VVTG (SEQ ID NO: 4)

SEQ ID NOS:2, 3 and 4 describe N-terminal MUC-1 signaling sequence for directing MUC1 receptor and truncated isoforms to cell membrane surface. Up to 3 amino acid residues may be absent at C-terminal end as indicated by variants in SEQ ID NOS:2, 3 and 4.

GTINVHDVETQFNQYKTEAASRYNLTISDVSVSDVPFPFSAQSGAGVPGWGIALLVLVCVLVALA
IVYLIALAVCQCRRKNYGQLDIFPARDTYHPMSEYPTYHTHGRYVPPSSTDRSPYEKVSAGNGGS
SLSYTNPAVAAASANL (SEQ ID NO:5) describes a truncated MUC1 receptor isoform having

nat-PSMGFR at its N-terminus and including the transmembrane and cytoplasmic sequences of a full-length MUC1 receptor.

GTINVHDVETQFNQYKTEAASRYNLTISDVSVSDVPFPFSAQSGA (SEQ ID NO:6) describes Native Primary Sequence of the MUC1 Growth Factor Receptor (nat-PSMGFR – an example of "PSMGFR"):

TINVHDVETQFNQYKTEAASRYNLTISDVSVSDVPFPFSAQSGA (SEQ ID NO:7) describes Native Primary Sequence of the MUC1 Growth Factor Receptor (nat-PSMGFR – An example of "PSMGFR"), having a single amino acid deletion at the N-terminus of SEQ ID NO:6).

GTINVHDVETQFNQYKTEAASPYNLTISDVSVSDVPFPFSAQSGA (SEQ ID NO:8) describes "SPY" functional variant of the native Primary Sequence of the MUC1 Growth Factor Receptor having enhanced stability (var-PSMGFR – An example of "PSMGFR").

TINVHDVETQFNQYKTEAASPYNLTISDVSVSDVPFPFSAQSGA (SEQ ID NO: 9) describes "SPY" functional variant of the native Primary Sequence of the MUC1 Growth Factor Receptor having enhanced stability (var-PSMGFR – An example of "PSMGFR"), having a single amino acid deletion at the C-terminus of SEQ ID NO:8).

tgtcagtgccgccgaaagaactacgggcagctggacatctttccagcccgggatacctaccatcc
tatgagcgagtaccccacctaccacacccatgggcgctatgtgccccctagcagtaccgatcgta
gcccctatgagaaggtttctgcaggtaacggtggcagcagcctctcttacacaaacccagcagtg
gcagccgcttctgccaacttg (SEQ ID NO:10) describes MUC1 cytoplasmic domain nucleotide sequence.

CQCRRKNYGQLDIFPARDTYHPMSEYPTYHTHGRYVPPSSTDRSPYEKVSAGNGGSSLSYTNPAV
AAASANL (SEQ ID NO:11) describes MUC1 cytoplasmic domain amino acid sequence.

gagatcctgagacaatgaatcatagtgaaagattcgttttcattgcagagtggtatgatccaaat
gcttcacttcttcgacgttatgagcttttatttacccaggggatggatctgttgaaatgcatga
tgtaaagaatcatcgcacctttttaaagcggaccaaatatgataacctgcacttggaagatttat
ttataggcaacaaagtgaatgtcttttctcgacaactggtattaattgactatggggatcaatat
acagctcgccagctgggcagtaggaaagaaaaaacgctagccctaattaaaccagatgcaatatc
aaaggctggagaaataattgaaataataaacaaagctggatttactataaccaaactcaaaatga
tgatgctttcaaggaaagaagcattggattttcatgtagatcaccagtcaagacccttttcaat
gagctgatccagtttattacaactggtcctattattgccatggagattttaagagatgatgctat
atgtgaatggaaaagactgctgggacctgcaaactctggagtggcacgcacagatgcttctgaaa
gcattagagccctctttggaacagatggcataagaaatgcagcgcatggccctgattcttttgct
tctgcggccagagaaatggagttgtttttccttcaagtggaggttgtgggccggcaaacactgc
taaatttactaattgtacctgttgcattgttaaaccccatgctgtcagtgaaggtatgttgaata
cactatattcagtacattttgttaataggagagcaatgtttatttttcttgatgtactttatgtat
agaaaataa (SEQ ID NO:12) describes NME7 nucleotide sequence (NME7: GENBANK ACCESSION AB209049).

DPETMNHSERFVFIAEWYDPNASLLRRYELLFYPGDGSVEMHDVKNHRTFLKRTKYDNLHLEDLF
IGNKVNVFSRQLVLIDYGDQYTARQLGSRKEKTLALIKPDAISKAGEIIEIINKAGFTITKLKMM
MLSRKEALDFHVDHQSRPFFNELIQFITTGPIIAMEILRDDAICEWKRLLGPANSGVARTDASES
IRALFGTDGIRNAAHGPDSFASAAREMELFFPSSGGCGPANTAKFTNCTCCIVKPHAVSEGMLNT
LYSVHFVNRRAMFIFLMYFMYRK (SEQ ID NO:13) describes NME7 amino acid sequence (NME7: GENBANK ACCESSION AB209049).

atggtgctactgtctactttagggatcgtctttcaaggcgaggggcctcctatctcaagctgtga
tacaggaaccatggccaactgtgagcgtaccttcattgcgatcaaaccagatggggtccagcggg
gtcttgtgggagagattatcaagcgttttgagcagaaaggattccgccttgttggtctgaaattc
atgcaagcttccgaagatcttctcaaggaacactacgttgacctgaaggaccgtccattctttgc
cggcctggtgaaatacatgcactcagggccggtagttgccatggtctgggaggggctgaatgtgg
tgaagacgggccgagtcatgctcggggagaccaaccctgcagactccaagcctgggaccatccgt
ggagacttctgcatacaagttggcaggaacattatacatggcagtgattctgtggagagtgcaga
gaaggagatcggcttgtggtttcaccctgaggaactggtagattacacgagctgtgctcagaact
ggatctatgaatga (SEQ ID NO:14) describes NM23-H1 nucleotide sequence (NM23-H1: GENBANK ACCESSION AF487339).

MVLLSTLGIVFQGEGPPISSCDTGTMANCERTFIAIKPDGVQRGLVGEIIKRFEQKGFRLVGLKF
MQASEDLLKEHYVDLKDRPFFAGLVKYMHSGPVVAMVWEGLNVVKTGRVMLGETNPADSKPGTIR
GDFCIQVGRNIIHGSDSVESAEKEIGLWFHPEELVDYTSCAQNWIYE (SEQ ID NO:15) NM23-H1 describes amino acid sequence (NM23-H1: GENBANK ACCESSION AF487339).

atggtgctactgtctactttagggatcgtctttcaaggcgaggggcctcctatctcaagctgtga
tacaggaaccatggccaactgtgagcgtaccttcattgcgatcaaaccagatggggtccagcggg
gtcttgtgggagagattatcaagcgttttgagcagaaaggattccgccttgttggtctgaaattc
atgcaagcttccgaagatcttctcaaggaacactacgttgacctgaaggaccgtccattctttgc
cggcctggtgaaatacatgcactcagggccggtagttgccatggtctgggaggggctgaatgtgg
tgaagacgggccgagtcatgctcggggagaccaaccctgcagactccaagcctgggaccatccgt
ggagacttctgcatacaagttggcaggaacattatacatggcggtgattctgtggagagtgcaga
gaaggagatcggcttgtggtttcaccctgaggaactggtagattacacgagctgtgctcagaact
ggatctatgaatga (SEQ ID NO:16) describes NM23-H1 S120G mutant nucleotide sequence (NM23-H1: GENBANK ACCESSION AF487339).

MVLLSTLGIVFQGEGPPISSCDTGTMANCERTFIAIKPDGVQRGLVGEIIKRFEQKGFRLVGLKF
MQASEDLLKEHYVDLKDRPFFAGLVKYMHSGPVVAMVWEGLNVVKTGRVMLGETNPADSKPGTIR
GDFCIQVGRNIIHGGDSVESAEKEIGLWFHPEELVDYTSCAQNWIYE (SEQ ID NO:17) describes NM23-H1 S120G mutant amino acid sequence (NM23-H1: GENBANK ACCESSION AF487339).

atg tacaacatga tggagacgga gctgaagccg ccgggcccgc agcaaacttc
gggggcggc ggcggcaact ccaccgcggc ggcggccggc ggcaaccaga aaaacagccc
ggaccgcgtc aagcggccca tgaatgcctt catggtgtgg tcccgcgggc agcggcgcaa
gatggcccag gagaacccca agatgcacaa ctcggagatc agcaagcgcc tgggcgccga
gtggaaactt ttgtcggaga cggagaagcg gccgttcatc gacgaggcta agcggctgcg
agcgctgcac atgaaggagc acccggatta taaataccgg ccccggcgga aaaccaagac

gctcatgaag aaggataagt acacgctgcc cggcgggctg ctggcccccg gcggcaatag
catggcgagc ggggtcgggg tgggcgccgg cctgggcgcg ggcgtgaacc agcgcatgga
cagttacgcg cacatgaacg gctggagcaa cggcagctac agcatgatgc aggaccagct
gggctacccg cagcacccgg gcctcaatgc gcacggcgca gcgcagatgc agcccatgca
ccgctacgac gtgagcgccc tgcagtacaa ctccatgacc agctcgcaga cctacatgaa
cggctcgccc acctacagca tgtcctactc gcagcagggc acccctggca tggctcttgg
ctccatgggt tcggtggtca agtccgaggc cagctccagc ccccctgtgg ttacctcttc
ctcccactcc agggcgccct gccaggccgg ggacctccgg gacatgatca gcatgtatct
ccccggcgcc gaggtgccgg aacccgccgc ccccagcaga cttcacatgt cccagcacta
ccagagcggc ccggtgcccg gcacggccat taacggcaca ctgccctct cacacatgtg
a (SEQ ID NO:18) describes SEQ ID No:1 human SOX2 nucleotide sequence (SOX2:GENBANK ACCESSION NM_003106).

MYNMMETELKPPGPQQTSGGGGGNSTAAAAGGNQKNSPDRVKRPMNAFMVWSRGQRRKMAQENPK
MHNSEISKRLGAEWKLLSETEKRPFIDEAKRLRALHMKEHPDYKYRPRRKTKTLMKKDKYTLPGG
LLAPGGNSMASGVGVGAGLGAGVNQRMDSYAHMNGWSNGSYSMMQDQLGYPQHPGLNAHGAAQMQ
PMHRYDVSALQYNSMTSSQTYMNGSPTYSMSYSQQGTPGMALGSMGSVVKSEASSSPPVVTSSSH
SRAPCQAGDLRDMISMYLPGAEVPEPAAPSRLHMSQHYQSGPVPGTAINGTLPLSHM (SEQ ID NO:19) describes human SOX2 amino acid sequence (SOX2:GENBANK ACCESSION NM_003106).

atggcgggacacctggcttcagattttgccttctcgccccctccaggtggtggaggtgatgggcc
agggggccggagccgggctgggttgatcctcggacctggctaagcttccaaggccctcctggag
ggccaggaatcgggccgggggttgggccaggctctgaggtgtgggggattcccccatgccccccg
ccgtatgagttctgtgggggggatggcgtactgtgggccccaggttggagtggggctagtgcccca
aggcggcttggagacctctcagcctgagggcgaagcaggagtcggggtggagagcaactccgatg
gggcctccccggagccctgcaccgtcacccctggtgccgtgaagctggagaaggagaagctggag
caaaacccggaggagtcccaggacatcaaagctctgcagaaagaactcgagcaatttgccaagct
cctgaagcagaagaggatcaccctgggatatacacaggccgatgtggggctcaccctgggggttc
tatttgggaaggtattcagccaaacgaccatctgccgctttgaggctctgcagcttagcttcaag
aacatgtgtaagctgcggcccttgctgcagaagtgggtggaggaagctgacaacaatgaaaatct
tcaggagatatgcaaagcagaaaccctcgtgcaggcccgaaagagaaagcgaaccagtatcgaga
accgagtgagaggcaacctggagaatttgttcctgcagtgcccgaaacccacactgcagcagatc
agccacatcgcccagcagcttgggctcgagaaggatgtggtccgagtgtggttctgtaaccggcg
ccagaagggcaagcgatcaagcagcgactatgcacaacgagaggattttgaggctgctgggtctc
ctttctcaggggaccagtgtcctttcctctggccccagggccccattttggtaccccaggctat
gggagccctcacttcactgcactgtactcctcggtccctttccctgaggggaagcctttccccc
tgtctctgtcaccactctgggctctcccatgcattcaaactga (SEQ ID NO:20) describes Human OCT4 nucleotide sequence.

MAGHLASDFAFSPPPGGGGDGPGGPEPGWVDPRTWLSFQGPPGGPGIGPGVGPGSEVWGIPPCPP
PYEFCGGMAYCGPQVGVGLVPQGGLETSQPEGEAGVGVESNSDGASPEPCTVTPGAVKLEKEKLE
QNPEESQDIKALQKELEQFAKLLKQKRITLGYTQADVGLTLGVLFGKVFSQTTICRFEALQLSFK
NMCKLRPLLQKWVEEADNNENLQEICKAETLVQARKRKRTSIENRVRGNLENLFLQCPKPTLQQI
SHIAQQLGLEKDVVRVWFCNRRQKGKRSSSDYAQREDFEAAGSPFSGGPVSFPLAPGPHFGTPGY
GSPHFTALYSSVPFPEGEAFPPVSVTTLGSPMHSN (SEQ ID NO:21) describes Human OCT4 amino acid sequence.

atggccaacctggagcgcaccttcatcgccatcaagccggacggcgtgcagcgcggcctggtggg
cgagatcatcaagcgcttcgagcagaagggattccgcctcgtggccatgaagttcctccgggcct
ctgaagaacacctgaagcagcactacattgacctgaaagaccgaccattcttccctgggctggtg
aagtacatgaactcagggccggttgtggccatggtctgggaggggctgaacgtggtgaagacagg
ccgagtgatgcttggggagaccaatccagcagattcaaagccaggcaccattcgtggggacttct
gcattcaggttggcaggaacatcattcatggcagtgattcagtaaaaagtgctgaaaaagaaatc
agcctatggtttaagcctgaagaactggttgactacaagtcttgtgctcatgactgggtctatga
ataa (SEQ ID NO:22) describes NM23-H2 nucleotide sequence (NM23-H2: GENBANK ACCESSION AK313448).

MANLERTFIAIKPDGVQRGLVGEIIKRFEQKGFRLVAMKFLRASEEHLKQHYIDLKDRPFFPGL
VKYMNSGPVVAMVWEGLNVVKTGRVMLGETNPADSKPGTIRGDFCIQVGRNIIHGSDSVKSAEK
EISLWFKPEELVDYKSCAHDWVYE (SEQ ID NO:23) describes NM23-H2 amino acid sequence (NM23-H2: GENBANK ACCESSION AK313448).

atggctgtcagcgacgcgctgctcccatctttctccacgttcgcgtctggcccggcgggaaggga
gaagacactgcgtcaagcaggtgccccgaataaccgctggcgggaggagctctcccacatgaagc
gacttcccccagtgcttcccgccggcccctatgacctggcggcggcgaccgtggccacagacctg
gagagcgccggagccggtgcggcttgcggcggtagcaacctggcgcccctacctcggagagagac
cgaggagttcaacgatctcctggacctggactttattctctccaattcgctgacccatcctccgg
agtcagtggccgccaccgtgtcctcgtcagcgtcagcctcctcttcgtcgtcgccgtcgagcagc
ggccctgccagcgcgccctccacctgcagcttcacctatccgatccgggccgggaacgacccggg
cgtggcgccgggcggcacgggcggaggcctcctctatggcagggagtccgctccccctccgacgg
ctcccttcaacctggcggacatcaacgacgtgagcccctcgggcggcttcgtggccgagctcctg
cggccagaattggacccggtgtacattccgccgcagcagccgcagccgccaggtggcgggctgat
gggcaagttcgtgctgaaggcgtcgctgagcgcccctggcagcgagtacggcagcccgtcggtca
tcagcgtcacgaaaggcagccctgacggcagccacccggtggtggtggcgccctacaacggcggg
ccgccgcgcacgtgccccaagatcaagcaggaggcggtctcttcgtgcacccacttgggcgctgg
acccccctctcagcaatggccaccggccggctgcacacgacttcccccctggggcggcagctcccca
gcaggactaccccgaccctgggtcttgaggaagtgctgagcagcagggactgtcaccctgccctg
ccgcttcctcccggcttccatccccacccggggcccaattacccatccttcctgcccgatcagat
gcagccgcaagtcccgccgctccattaccaagagctcatgccacccggttcctgcatgccagagg
agcccaagccaaagaggggaagacgatcgtggccccggaaaaggaccgccacccacacttgtgat
tacgcgggctgcggcaaaacctacacaaagagttcccatctcaaggcacacctgcgaacccacac
aggtgagaaaccttaccactgtgactgggacggctgtggatggaaattcgcccgctcagatgaac
tgaccaggcactaccgtaaacacacggggcaccgcccgttccagtgccaaaaatgcgaccgagca
ttttccaggtcggaccacctcgccttacacatgaagaggcatttt (SEQ ID NO:24) describes KLF4 nucleotide sequence (KLF4: GENBANK ACCESSION AF022184).

MAVSDALLPSFSTFASGPAGREKTLRQAGAPNNRWREELSHMKRLPPVLPAGPYDLAAATVATDL
ESAGAGAACGGSNLAPLPRRETEEFNDLLDLDFILSNSLTHPPESVAATVSSSASASSSSSPSSS
GPASAPSTCSFTYPIRAGNDPGVAPGGTGGGLLYGRESAPPPTAPFNLADINDVSPSGGFVAELL
RPELDPVYIPPQQPQPPGGGLMGKFVLKASLSAPGSEYGSPSVISVTKGSPDGSHPVVVAPYNGG
PPRTCPKIKQEAVSSCTHLGAGPPLSNGHRPAAHDFPLGRQLPSRTTPTLGLEEVLSSRDCHPAL
PLPPGFHPHPGPNYPSFLPDQMQPQVPPLHYQELMPPGSCMPEEPKPKRGRRSWPRKRTATHTCD
YAGCGKTYTKSSHLKAHLRTHTGEKPYHCDWDGCGWKFARSDELTRHYRKHTGHRPFQCQKCDRA

FSRSDHLALHMKRHF (SEQ ID NO:25) describes KLF4 amino acid sequence (KLF4: GENBANK ACCESSION AF022184).

atggatttttttcgggtagtggaaaaccagcagcctcccgcgacgatgcccctcaacgttagctt
caccaacaggaactatgacctcgactacgactcggtgcagccgtatttctactgcgacgaggagg
agaacttctaccagcagcagcagcagagcgagctgcagcccccggcgcccagcgaggatatctgg
aagaaattcgagctgctgcccaccccgcccctgtcccctagccgccgctccgggctctgctcgcc
ctcctacgttgcggtcacacccttctcccttcggggagacaacgacggcggtggcgggagcttct
ccacggccgaccagctggagatggtgaccgagctgctgggaggagacatggtgaaccagagtttc
atctgcgacccggacgacgagaccttcatcaaaaacatcatcatccaggactgtatgtggagcgg
cttctcggccgccgccaagctcgtctcagagaagctggcctcctaccaggctgcgcgcaaagaca
gcggcagcccgaaccccgcccgcggccacagcgtctgctccacctccagcttgtacctgcaggat
ctgagcgccgccgcctcagagtgcatcgacccctcggtggtcttcccctaccctctcaacgacag
cagctcgcccaagtcctgcgcctcgcaagactccagcgccttctctccgtcctcggattctctgc
tctcctcgacggagtcctccccgcagggcagccccgagcccctggtgctccatgaggagacaccg
cccaccaccagcagcgactctgaggaggaacaagaagatgaggaagaaatcgatgttgtttctgt
ggaaaagaggcaggctcctggcaaaaggtcagagtctggatcaccttctgctggaggccacagca
aacctcctcacagcccactggtcctcaagaggtgccacgtctccacacatcagcacaactacgca
gcgcctccctccactcggaaggactatcctgctgccaagagggtcaagttggacagtgtcagagt
cctgagacagatcagcaacaaccgaaaatgcaccagccccaggtcctcggacaccgaggagaatg
tcaagaggcgaacacacaacgtcttggagcgccagaggaggaacgagctaaaacggagctttttt
gccctgcgtgaccagatcccggagttggaaaacaatgaaaaggcccccaaggtagttatccttaa
aaaagccacagcatacatcctgtccgtccaagcagaggagcaaaagctcatttctgaagaggact
tgttgcggaaacgacgagaacagttgaaacacaaacttgaacagctacggaactcttgtgcg
(SEQ ID NO:26) describes c-Myc nucleotide sequence (c-Myc:GENBANK ACCESSION BC000917).

MDFFRVVENQQPPATMPLNVSFTNRNYDLDYDSVQPYFYCDEEENFYQQQQQSELQPPAPSEDIW
KKFELLPTPPLSPSRRSGLCSPSYVAVTPFSLRGDNDGGGGSFSTADQLEMVTELLGGDMVNQSF
ICDPDDETFIKNIIIQDCMWSGFSAAAKLVSEKLASYQAARKDSGSPNPARGHSVCSTSSLYLQD
LSAAASECIDPSVVFPYPLNDSSSPKSCASQDSSAFSPSSDSLLSSTESSPQGSPEPLVLHEETP
PTTSSDSEEEQEDEEEIDVVSVEKRQAPGKRSESGSPSAGGHSKPPHSPLVLKRCHVSTHQHNYA
APPSTRKDYPAAKRVKLDSVRVLRQISNNRKCTSPRSSDTEENVKRRTHNVLERQRRNELKRSFF
ALRDQIPELENNEKAPKVVILKKATAYILSVQAEEQKLISEEDLLRKRREQLKHKLEQLRNSCA
(SEQ ID NO:27) describes c-Myc amino acid sequence (c-Myc: GENBANK ACCESSION BC000917).

atgggctccgtgtccaaccagcagtttgcaggtggctgcgccaaggcggcagaagaggcgcccga
ggaggcgccggaggacgcggcccgggcggcggacgagcctcagctgctgcacggtgcgggcatct
gtaagtggttcaacgtgcgcatggggttcggcttcctgtccatgaccgcccgcgccggggtcgcg
ctcgacccccagtggatgtctttgtgcaccagagtaagctgcacatggaagggttccggagctt
gaaggagggtgaggcagtggagttcacctttaagaagtcagccaagggtctggaatccatccgtg
tcaccggacctggtggagtattctgtattgggagtgagaggcggccaaaaggaaagagcatgcag
aagcgcagatcaaaaggagacaggtgctacaactgtggaggtctagatcatcatgccaaggaatg
caagctgccacccagcccaagaagtgccacttctgccagagcatcagccatatggtagcctcat
gtccgctgaaggcccagcagggccctagtgcacagggaaagccaacctactttcgagaggaagaa

gaagaaatccacagccctaccctgctcccggaggcacagaat (SEQ ID NO:28) describes LIN28 nucleotide sequence (LIN28: GENBANK ACCESSION AF521099).

MGSVSNQQFAGGCAKAAEEAPEEAPEDAARAADEPQLLHGAGICKWFNVRMGFGFLSMTARAGVA
LDPPVDVFVHQSKLHMEGFRSLKEGEAVEFTFKKSAKGLESIRVTGPGGVFCIGSERRPKGKSMQ
KRRSKGDRCYNCGGLDHHAKECKLPPQPKKCHFCQSISHMVASCPLKAQQGPSAQGKPTYFREEE
EEIHSPTLLPEAQN (SEQ ID NO:29) describes LIN28 amino acid sequence (LIN28: GENBANK ACCESSION AF521099).

atgtctcccgccccaagaccctcccgttgtctcctgctccccctgctcacgctcggcaccgcgct
cgcctccctcggctcggcccaaagcagcagcttcagccccgaagcctggctacagcaatatggct
acctgcctcccggggacctacgtacccacacacagcgctcaccccagtcactctcagcggccatc
gctgccatgcagaagttttacggcttgcaagtaacaggcaaagctgatgcagacaccatgaaggc
catgaggcgcccccgatgtggtgttccagacaagtttggggctgagatcaaggccaatgttcgaa
ggaagcgctacgccatccagggtctcaaatggcaacataatgaaatcactttctgcatccagaat
tacacccccaaggtgggcgagtatgccacatacgaggccattcgcaaggcgttccgcgtgtggga
gagtgccacaccactgcgcttccgcgaggtgccctatgcctacatccgtgagggccatgagaagc
aggccgacatcatgatcttctttgccgagggcttccatggcgacagcacgcccttcgatggtgag
ggcggcttcctggcccatgcctacttcccaggccccaacattggaggagacacccactttgactc
tgccgagccttggactgtcaggaatgaggatctgaatggaaatgacatcttcctggtggctgtgc
acgagctgggccatgccctggggctcgagcattccagtgacccctcggccatcatggcacccttt
taccagtggatggacacggagaattttgtgctgcccgatgatgaccgccggggcatccagcaact
ttatgggggtgagtcagggttccccaccaagatgccccctcaacccaggactacctcccggcctt
ctgttcctgataaacccaaaaaccccacctatgggcccaacatctgtgacgggaactttgacacc
gtggccatgctccgaggggagatgtttgtcttcaaggagcgctggttctggcgggtgaggaataa
ccaagtgatggatggatacccaatgcccattggccagttctggcggggcctgcctgcgtccatca
acactgcctacgagaggaaggatggcaaattcgtcttcttcaaaggagacaagcattgggtgttt
gatgaggcgtccctggaacctggctaccccaagcacattaaggagctgggccgagggctgcctac
cgacaagattgatgctgctctcttctggatgcccaatggaaagacctacttcttccgtggaaaca
agtactaccgtttcaacgaagagctcagggcagtggatagcgagtaccccaagaacatcaaagtc
tgggaagggatccctgagtctcccagagggtcattcatgggcagcgatgaagtcttcacttactt
ctacaaggggaacaaatactggaaattcaacaaccagaagctgaaggtagaaccgggctacccca
agtcagccctgagggactggatgggctgcccatcgggaggccggccggatgaggggactgaggag
gagacggaggtgatcatcattgaggtggacgaggagggcggcggggcggtgagcgcggctgccgt
ggtgctgcccgtgctgctgctgctcctggtgctggcggtgggccttgcagtcttcttcttcagac
gccatgggacccccaggcgactgctctactgccagcgttccctgctggacaaggtc (SEQ ID NO:30) describes MMP14 nucleotide sequence (MMP14: GENBANK ACCESSION BC064803).

MSPAPRPSRCLLLPLLTLGTALASLGSAQSSSFSPEAWLQQYGYLPPGDLRTHTQRSPQSLSAAI
AAMQKFYGLQVTGKADADTMKAMRRPRCGVPDKFGAEIKANVRRKRYAIQGLKWQHNEITFCIQN
YTPKVGEYATYEAIRKAFRVWESATPLRFREVPYAYIREGHEKQADIMIFFAEGFHGDSTPFDGE
GGFLAHAYFPGPNIGGDTHFDSAEPWTVRNEDLNGNDIFLVAVHELGHALGLEHSSDPSAIMAPF
YQWMDTENFVLPDDDRRGIQQLYGGESGFPTKMPPQPRTTSRPSVPDKPKNPTYGPNICDGNFDT
VAMLRGEMFVFKERWFWRVRNNQVMDGYPMPIGQFWRGLPASINTAYERKDGKFVFFKGDKHWVF
DEASLEPGYPKHIKELGRGLPTDKIDAALFWMPNGKTYFFRGNKYYRFNEELRAVDSEYPKNIKV
WEGIPESPRGSFMGSDEVFTYFYKGNKYWKFNNQKLKVEPGYPKSALRDWMGCPSGGRPDEGTEE
ETEVIIIEVDEEGGGAVSAAAVVLPVLLLLLVLAVGLAVFFFRRHGTPRRLLYCQRSLLDKV

(SEQ ID NO:31) describes MMP14 amino acid sequence (MMP14: GENBANK ACCESSION BC064803).

atgatcttactcacattcagcactggaagacggttggatttcgtgcatcattcgggggtgttttt
cttgcaaaccttgctttggattttatgtgctacagtctgcggaacggagcagtatttcaatgtgg
aggtttggttacaaaagtacggctaccttccaccgactgaccccagaatgtcagtgctgcgctct
gcagagaccatgcagtctgccctagctgccatgcagcagttctatggcattaacatgacaggaaa
agtggacagaaacacaattgactggatgaagaagccccgatgcggtgtacctgaccagacaagag
gtagctccaaatttcatattcgtcgaaagcgatatgcattgacaggacagaaatggcagcacaag
cacatcacttacagtataaagaacgtaactccaaaagtaggagaccctgagactcgtaaagctat
tcgccgtgcctttgatgtgtggcagaatgtaactcctctgacatttgaagaagttccctacagtg
aattagaaaatggcaaacgtgatgtggatataaccattatttttgcatctggtttccatggggac
agctctccctttgatggagagggaggatttttggcacatgcctacttccctggaccaggaattgg
aggagatacccattttgactcagatgagccatggacactaggaaatcctaatcatgatggaaatg
acttatttcttgtagcagtccatgaactgggacatgctctgggattggagcattccaatgacccc
actgccatcatggctccattttaccagtacatggaaacagacaacttcaaactacctaatgatga
tttacagggcatccagaaaatatatggtccacctgacaagattcctccacctacaagacctctac
cgacagtgcccccacaccgctctattcctccggctgacccaaggaaaaatgacaggccaaaacct
cctcggcctccaaccggcagaccctcctatcccggagccaaacccaacatctgtgatgggaactt
taacactctagctattcttcgtcgtgagatgtttgttttcaaggaccagtggttttggcgagtga
gaaacaacagggtgatggatggatacccaatgcaaattacttacttctggcggggcttgcctcct
agtatcgatgcagtttatgaaaatagcgacgggaattttgtgttctttaaaggtaacaaatattg
ggtgttcaaggatacaactcttcaacctggttaccctcatgacttgataacccttggaagtggaa
ttcccccctcatggtattgattcagccatttggtgggaggacgtcgggaaaacctatttcttcaag
ggagacagatattggagatatagtgaagaaatgaaaacaatggaccctggctatcccaagccaat
cacagtctggaaagggatccctgaatctcctcagggagcatttgtacacaaagaaaatggcttta
cgtatttctacaaaggaaaggagtattggaaattcaacaaccagatactcaaggtagaacctgga
catccaagatccatcctcaaggattttatgggctgtgatggaccaacagacagagttaaagaagg
acacagcccaccagatgatgtagacattgtcatcaaactggacaacacagccagcactgtgaaag
ccatagctattgtcattccctgcatcttggccttatgcctccttgtattggtttacactgtgttc
cagttcaagaggaaaggaacaccccgccacatactgtactgtaaacgctctatgcaagagtgggt
g (SEQ ID NO:32) describes MMP16 nucleotide sequence (MMP16:GENBANK ACCESSION AB009303).

MILLTFSTGRRLDFVHHSGVFFLQTLLWILCATVCGTEQYFNVEVWLQKYGYLPPTDPRMSVLRS
AETMQSALAAMQQFYGINMTGKVDRNTIDWMKKPRCGVPDQTRGSSKFHIRRKRYALTGQKWQHK
HITYSIKNVTPKVGDPETRKAIRRAFDVWQNVTPLTFEEVPYSELENGKRDVDITIIFASGFHGD
SSPFDGEGGFLAHAYFPGPGIGGDTHFDSDEPWTLGNPNHDGNDLFLVAVHELGHALGLEHSNDP
TAIMAPFYQYMETDNFKLPNDDLQGIQKIYGPPDKIPPPTRPLPTVPPHRSIPPADPRKNDRPKP
PRPPTGRPSYPGAKPNICDGNFNTLAILRREMFVFKDQWFWRVRNNRVMDGYPMQITYFWRGLPP
SIDAVYENSDGNFVFFKGNKYWVFKDTTLQPGYPHDLITLGSGIPPHGIDSAIWWEDVGKTYFFK
GDRYWRYSEEMKTMDPGYPKPITVWKGIPESPQGAFVHKENGFTYFYKGKEYWKFNNQILKVEPG
HPRSILKDFMGCDGPTDRVKEGHSPPDDVDIVIKLDNTASTVKAIAIVIPCILALCLLVLVYTVF
QFKRKGTPRHILYCKRSMQEWV (SEQ ID NO:33) describes MMP 16 amino acid sequence (MMP16:GENBANK ACCESSION AB009303)

[0034] Induction of Pluripotency

[0035] It was recently discovered that somatic cells can be reprogrammed to revert to the pluripotent state. The genes that code for transcription factors OCT4, SOX2, KLF4, NANOG, c-MYC and LIN28 or the proteins themselves can be introduced into somatic cells and cause a reversion to the pluripotent state. Many of these pluripotency factors were previously thought of as oncogenes. C-Myc is a well known oncogene and similarly, Klf4 has been shown to induce dysplasia (Foster et al., 2005). OCT4 was identified as the gold standard for identifying pluripotent stem cells. The presence of OCT4 in the nucleus indicates that the cell is pluripotent and its absence indicates that the cell has entered the differentiation process and is no longer able to differentiate into *any* cell type. Recently, it became known that OCT4 is also present in the nucleus of many cancer cells, but not in normal mature cells. The present inventors recently discovered that a cleaved form of the MUC1 transmembrane protein (SEQ ID NO:1), MUC1*, is a powerful growth factor receptor that is expressed on an estimated 75% of solid tumor cancers (Raina et al., 2009) and that it is also expressed in this "tumorigenic" form on pluripotent stem cells (Hikita et al., 2008). The present invention relates to MUC1* and MUC1* associated factors as well as methods employing them for the induction or maintenance of pluripotency or to enhance the efficiency of inducing pluripotency.

[0036] The present invention encompasses using MUC1* associated factors, which include protein factors, genes that encode them, or small molecules that affect their expression, to induce or improve the efficiency of generating iPS cells. We have shown that a cleaved form of the MUC1 transmembrane protein - MUC1* - is a primal growth factor receptor that mediates the growth of both cancer and pluripotent stem cells. Interrupting the interaction between MUC1* extracellular domain and its activating ligand, NM23, is lethal to pluripotent stem cells (Hikita et al., 2008) which indicates that this pathway is critical for pluripotency. NM23 is a ligand that activates MUC1* (Mahanta et al., 2008) (SEQ ID NOS:12-17 and 22-23). In addition to its ability to stimulate pluripotent stem cell growth, while inhibiting differentiation, NM23 has been shown to induce transcription of c-Myc (Dexheimer at al., 2009). Therefore, adding NM23 to cells undergoing conversion to the pluripotent state will provide the benefits of transfecting the c-Myc oncogene, while mitigating the associated health risks. In addition, stimulation of MUC1*, by either NM23 or a bivalent anti-MUC1* antibody, activates the MAP kinase proliferation pathway, which increases cell survival (Mahanta et al., 2008). NANOG expression induces pluripotency; the tumor suppressor p53 suppresses Nanog expression (Lin et al., 2007).

Therefore, the need for NANOG for inducing pluripotency is reduced or eliminated by suppressing p53. An ectopically expressed 72-amino acid fragment of the MUC1 cytoplasmic tail (MUC10-CT) has been shown to be present in the nucleus of cancer cells where it binds to the p53 promoter (Wei et al., 2007). The approximately 72 amino acid fragment of MUC1-CD such as shown in SEQ ID NO:11 can be used in combination with other pluripotency-inducing factors to induce or enhance iPS cell generation. However, this peptide does not correspond to a naturally occurring MUC1 species, and therefore may produce undesired effects. The present inventors disclose that MUC1* translocates to the nucleus (Examples 1 and 9, and Figure 6) and therefore is used alone or in combination with other pluripotency-inducing factors to induce or enhance iPS cell generation. In support of this approach, it has been reported that several genes from the core set of pluripotency genes regulate transcription of MUC1, its cleavage enzyme and/or its activating ligand NM23 (Boyer et al., 2005). OCT4 and SOX2 bind to the MUC1 promoter and also to the promoter of its cleavage enzyme, MMP-14. SOX2 and NANOG bind to the NM23 promoter. Given that MUC1* is critical for maintenance of hESCs and is the target of the key pluripotency genes, we disclose that the introduction of MUC1*, or agents that increase cleavage of MUC1 to the MUC1* form, along with its activating ligand, NM23 can be used to replace some or all of the previously identified pluripotency-inducing factors to induce or enhance the generation of iPS cells.

[0037] The present invention discloses novel reagents and methods, involving MUC1*, for inducing pluripotency in cells. These reagents and methods are used to induce pluripotency in somatic or mature cells. In another aspect of the invention, they are used to increase the efficiency of inducing pluripotency in mature cells. In yet another aspect of the invention they are used to maintain immature cells in an immature state. In another aspect of the invention they are used to inhibit differentiation. In another aspect of the invention, these reagents and methods are used for maintaining stem cells in the pluripotent state.

[0038] The invention involves reversing differentiation or maintaining stem-like characteristics by introducing to mature cells, or somewhat differentiated cells, genes or gene products that affect the expression of MUC1* and its associated factors. MUC1* is the cleaved form of the transmembrane protein MUC1. MUC1* associated factors include, but are not limited to, full-length MUC1, enzymes that cleave MUC1, MUC1* activating ligands and also transcription factors that affect the expression of MUC1 or MUC1*. The invention is also drawn

to the introduction of the genes or gene products for MUC1* or MUC1* associated factors to mature cells or somewhat differentiated cells will induce pluripotency or stem-ness in those cells or their progeny. The present application describes their use for maintaining pluripotency in stem cells. Agents that affect expression of MUC1* or MUC1* associated factors can be added in combination with, or to replace one or more genes or gene products that are already known to induce pluripotency including OCT4, SOX2, KLF4, NANOG, c-myc and LIN28.

[0039] Forced expression of combinations of the transcription factors, Oct4, Sox2, Klf4 and c-Myc or Oct4, Sox2, Nanog and Lin28 have been shown to cause mature cells to revert to the pluripotent state (Takahashi and Yamanaka, 2006). Each of the transcription factors that induce pluripotency regulates the transcription of about a dozen genes. Among these were several that the inventor has identified as being MUC1-associated factors. OCT4 and SOX2 bind to the MUC1 promoter itself. SOX2 and NANOG bind to the NM23 (NME7) promoter. NM23 (also known as NME) was previously identified, by the present inventor, as the activating ligand of MUC1* (Mahanta et al., 2008). NME7 is an activating ligand of MUC1*. OCT4 and SOX2 both bind to the promoter for MMP16 which we disclose herein is a cleavage enzyme of MUC1. OCT4, SOX2 or NANOG also bind to promoter sites for cleavage enzymes MMP2, MMP9, MMP10, ADAM TSL-1, ADAM TS-4, ADAM-17 (a MUC1 cleavage enzyme), ADAM-TS16, ADAM-19 and ADAM-28. Some or all of these cleavage enzyme may be upregulated to enhance the cleavage of MUC1 to the MUC1* form to induce pluripotency or maintain it (Boyer et al, 2005).

[0040] Our previous work with embryonic stem cells, which only express the cleaved form of MUC1, MUC1*, showed that dimerization of its extracellular domain stimulate growth and inhibit differentiation (Hikita et al., 2008). These effects were achieved by dimerizing the MUC1* extracellular domain using either a bivalent Anti-MUC1* antibody, recombinant NM23, or a mutant NM23 (S120G) that preferentially forms dimers (Kim et al., 2003). Inhibition of MUC1* extracellular domain using the monovalent Anti-MUC1* Fab was lethal within hours. These findings indicate that MUC1* is a significant "stemness" factor. In addition, OCT4 and SOX2 bind to the MUC1 gene promoter and also to the promoter of its cleavage enzymes. SOX2 and NANOG bind to the NM23 (NME7) promoter. Since blocking the extracellular domain of MUC1* are lethal to hESCs, it follows that the pluripotency genes, OCT4, SOX2, and NANOG, are induce expression of MUC1, its cleavage enzyme and its activating ligand. One or more of

the genes or gene products that have already been shown to induce pluripotency can be replaced by transfecting the gene or introducing the gene product, for MUC1* alone or in addition to its cleavage enzymes and/or activating ligands, NME7, NME-H1, NME-H2 or an antibody that dimerizes the PSMGFR epitope of MUC1 or MUC1*.

[0041] As those who are skilled in the art are familiar, signal sequences that direct the localization of the transfected gene product may be added to the gene. Examples of signal sequences are given as SEQ ID NOS:2-4. The invention contemplates that the N-terminal domain of MUC1* may be truncated or extended by up to nine (9) amino acids without substantially altering the effect of these genes or gene products. MUC1* exemplified as SEQ ID NO:5 and variants whose extracellular domain is essentially comprised of the sequences given in SEQ ID NOS: 6, 7, 8 and 9 are preferred.

[0042] MUC1, MUC1*, or associated factors, including those listed above, can substitute for one or more of the genes or gene products that induce pluripotency and are used to induce pluripotency or to maintain it.

[0043] In one case, somatic cells such as fibroblasts and dermal fibroblasts are transfected with a gene that encodes the MUC1 protein, which aids in inducing stem cell-like features and in some cases induces progeny to become pluripotent stem cells. In another aspect of the invention, a gene for MUC1* is transfected into cells to induce a return to a stem cell-like state and in some cases induce actual pluripotent stem cells. Each of the MUC1 or MUC1* genes may be introduced to the cell alone or in combination with other genes that aid in inducing pluripotency or stem cell-like characteristics. For example, DNA encoding MUC1 or preferentially MUC1* is introduced to the cell along with one or more of the genes that encode OCT4, SOX2, NANOG, LIN28, KLF4, and/or c-Myc. DNA encoding a truncated form of MUC1, preferentially MUC1*, is transfected into fibroblasts along with genes encoding OCT4, SOX2, NANOG, and LIN28 (Yu et al., 2007). In another embodiment, DNA encoding a truncated form of MUC1, preferentially MUC1*, is transfected into somatic cells, fibroblasts, or other cells, along with genes encoding OCT4, SOX2, KLF4, and c-Myc (Takahashi et al., 2007). Similarly, DNA encoding MUC1* and/or its activating ligand, NM23 or the S120G mutant of NM23, are transfected into cells to induce pluripotency. MUC1* and/or NM23 may be transfected along with other genes such as OCT4, SOX2, NANOG, LIN28, KLF4, and/or c-Myc to induce pluripotency or stem cell-like characteristics. DNA encoding antibodies that recognize MUC1*

or MUC1 may also be transfected into cells alone or with other genes to induce stem cell characteristics in the cells or their progeny. If secreted, anti-MUC1* antibodies will dimerize and thus activate the MUC1* receptor and its functions that promote or maintain stem-like characteristics.

[0044] Similarly, factors such as nucleic acids, proteins, modified proteins or small molecules that affect the expression of MUC1, MUC1* or their associated factors are introduced to cells to induce characteristics of stem cells or to induce a return to pluripotency. For example, genes or gene products for MUC1 cleavage enzymes, MMP14, MMp16, MMP2, MMP9, MMP10, ADAM TSL-1, ADAM TS-4 ADAM-17 (a MUC1 cleavage enzyme), ADAM-TS16, ADAM-19 and ADAM-28 are introduced to cells to induce pluripotency or similar characteristics.

[0045] In another embodiment, non-protein agents are added to cells to induce or enhance the induction of pluripotency. For example the phorbol ester phorbol 12-myristate 13-acetate (PMA) is a small molecule that increases the cleavage of MUC1 to MUC1* (Thathiah et al., 2003). In one aspect of the invention, phorbol ester is added to cells undergoing conversion to pluripotency to induce or increase the efficiency of iPS generation.

[0046] In another example, ligands that interact with MUC1 or MUC1* are added to somatic cells, dermal fibroblasts, fibroblasts, or somewhat differentiated cells to induce pluripotency either alone or in combination with other genes to induce or maintain pluripotency. For example, one or more of the genes encoding OCT4, SOX2, NANOG, LIN28, KLF4, and/or c-Myc are transfected into fibroblasts or other cells and then are cultured in the presence of ligands that activate MUC1 or MUC1*. Dimeric, protein ligands of MUC1* are preferred. In a preferred embodiment, a bivalent anti-MUC1* antibody is added to cells that have been transfected with genes that influence cells or their progeny to become pluripotent stem cells.

[0047] In a preferred embodiment, NM23 (NM23-H1, NM23-H2, or NME7) is introduced to cells, as the gene that encodes it, as the protein itself or as a protein bearing a leader sequence such as a poly-arginine tract, to facilitate entry into the cell, to aid in the induction or maintenance of pluripotency. The inventors recently showed that when NM23 is secreted by pluripotent stem cells (and cancer cells), it is an activating ligand of the cleaved form of MUC1 – MUC1* - and triggers the MAP kinase proliferation pathway. NM23 stimulation of MUC1* was shown to promote the growth of pluripotent hESCs and inhibited their differentiation (Hikita et

al., 2008). NM23 also induces the transcription of c-MYC (Dexheimer at al., 2009) and replaces the need for c-MYC. NM23 is added exogenously either in its native state to activate the MUC1* growth factor receptor or with a poly arginine tract to facilitate entry into the cell and nucleus where it induces C-MYC expression. NM23 (NME) may be added as the encoding nucleic acid, or as the expressed protein with or without a modification that facilitates entry into the cell. NME-H1, -H2 or -7 can be used in their native state or in mutant forms that favor the dimeric state, such as the S120G mutation.

[0048] In another aspect of the invention, a bivalent antibody that binds to the extracellular domain of MUC1* (PSMGFR) or a dimeric MUC1* ligand, such as NM23, or genes encoding them are added to MUC1*-expressing cells to induce pluripotency, increase the efficiency of the induction of pluripotency, to maintain pluripotency or to inhibit differentiation. The cells to which these MUC1 or MUC1* interacting proteins are added may be naturally occurring cells or those into which genes to induce stem cell-like characteristics have been added, or have already entered the differentiation process or may be stem cells.

[0049] Genes for inducing pluripotency may be introduced on the same or different plasmids, which may be lenti viral vector driven or adenovirus vectors or any integrating or non-integrating viral or non-viral vector, or any other system that facilitates introduction of these genes into the desired cells.

[0050] In many cases, it is preferential to achieve the effects of pluripotency-inducing proteins by introducing the proteins themselves rather than the nucleic acids or genes that encode them. The invention encompasses genes disclosed here for the induction of stem-like characteristics or pluripotency that can be replaced by the gene products, the proteins, either in their native state or modified with leader sequences such as poly-arginine tracts to allow entry into the cells. The products of these genes, i.e. proteins, or other proteins which interact with one or more of the products of the transfected genes are introduced to cells to induce or maintain pluripotency or other stem-cell like characteristics.

[0051] In other cases, it may be beneficial to introduce synthetic agents, such as small molecules, to induce stem-ness in mature or differentiated cells (Lyssiotis et al. 2009). In one aspect of the invention, small molecules are added to cells that either directly or indirectly induce the transcription of genes that induce pluripotency. In other cases, small molecules that directly or indirectly increase the production of MUC1* are added. In one instance, these small

molecules increase cleavage of MUC1 to the MUC1* form, which is a characteristic of stem cells. Phorbol ester, for example, is a small molecule that increases cleavage of MUC1 to MUC1*, so when added to cells, it promotes induction or maintenance of pluripotent state by generating MUC1*.

[0052] P53, which is also known as a tumor suppressor, promotes apoptosis. It would therefore be advantageous to inhibit p53 when culturing stem cells or inducing pluripotency in somatic or other cells. The present invention anticipates using p53 suppressors along with other reagents and methods of the invention to maintain stem-ness or induce stem-like or pluripotent characteristics. P53 can be suppressed by a number of methods. Small molecules such as Pifithrin-μ inhibits the pro-apoptotic effects of p53 (Strom, et al., 2006 Sep; Komarov, et al., 1999) and thus are optionally added to cells to increase efficiency of induction of pluripotency or to maintain stem-ness. In a preferred embodiment, p53 inhibitors are used along with genes or gene products that up-regulate MUC1 or MUC1*, including but not limited to the MUC1 or MUC1* genes or gene products, their activating ligands and their cleavage enzymes.

[0053] Another method for suppressing p53 activity to increase the efficiency of inducing pluripotency or maintaining stem-ness is by the introduction of the MUC1* protein to cell cultures. The MUC1* protein can be modified by adding on a poly-arginine tract to facilitate entry into the cell. It has been reported that the overexpression of the cytoplasmic tail, alone, of MUC1 (MUC1-CD) resulted in its translocation to the nucleus where it was found to bind to the p53 promoter. These studies could not determine whether MUC1-CD down or up-regulated p53. The present invention is also drawn to the repression of p53 by the ectopic expression of MUC1*, to increase the efficiency of inducing pluripotency or other stem-like characteristics. MUC1* can be introduced by inserting the gene into the cell, by adding the protein itself exogenously or by adding the MUC1* protein that is optionally modified with a poly-arginine tract.

[0054] The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims. The following examples are offered by way of illustration of the present invention, and not by way of limitation.

EXAMPLES

[0055] Example 1. MUC1* promotes growth and cell death resistance.

[0056] MUC1* promotes clonogenic growth (colony expansion) of fibroblasts. Single cell clones of 3Y1 cells transfected with either full-length MUC1 (SEQ ID NO:1), MUC1*$_{1110}$ (SEQ ID NO:5) or empty vector were plated at 1000 cells per 60mm dish in DMEM media containing 10% fetal bovine serum, penicillin/streptomycin and G418 (600 μg/ml). Cells were grown for 9 days and then fixed in 4% paraformaldehyde for 15 minutes at room temperature. Dishes were washed with water and then stained with 1% crystal violet in 70% methanol for 20 minutes at room temperature. Dishes were washed three times with water and allowed to dry overnight at room temperature and photographed. Figure 1A shows that the amount of crystal violet that is absorbed (an indicator of cell number) is much higher where MUC1* single cell clones #3 and #44 are growing. In contrast, cells that transfected with full-length MUC1 (single cell clones #8 and #17) showed no growth rate increase over cells transfected with the empty vector. This shows that the cleaved form, MUC1*, confers a growth and/or survival advantage and not the full-length protein.

[0057] Example 2. Anti-MUC1* Fab blocks resistance to cell death by TAXOL® in trastuzumab (HERCEPTIN®)-resistant cells (made resistant by culture in 1ug/ml HERCEPTIN®). Fessler et al., 2009 reported that HERCEPTIN® resistant cells are also resistant to TAXOL®, doxorubicin and cyclophosphamide. As reported, these drug resistant cancer cells achieve resistance by overexpressing MUC1*. The following experiment showed that blocking the PSMGFR portion of the MUC1* extracellular domain reversed acquired drug resistance in cancer cells. Parental (BT474) or resistant (BTRes1) cells were plated at a density of 10,000 cells/well in 96 well plates, 4 wells/condition. The following day, Anti-MUC1* Fab, control Fab, or no Fab were added to cells in the presence or absence of TAXOL® (Paclitaxel Sigma T7191). Two days later, cells were resuspended in 50 μl trypsin, and counted in the presence of trypan blue. Percent cell death was calculated as percent trypan blue uptake. BT474 cells underwent cell death in response to TAXOL® under each condition, and BTRes1 cells only underwent cell death in the presence of MUC1* antibody (Figure 1B).

[0058] Example 3. MUC1* acts as a growth factor receptor, and is activated by dimerization of its extracellular domain using an artificial (anti-MUC1* antibody) or its natural ligand, NM23

(NME). MUC1*-positive ZR-75-30 cells, 6000/well, or control (MUC1-negative) HEK293 cells 4000/well, were plated in 96 well plates. The following day, zero hour cell counts were taken, and different concentrations of anti-MUC1* antibody or Fab were added in medium with low (0.1%) serum every 24 or 48 hours. After several days of incubation, cells were resuspended in trypsin and counted, and percent normalized growth was calculated. Stimulation of ZR-75-30 cells, shown as a bell-shaped curve, as is demonstrated for ligand-induced growth stimulation, but not HEK293 cells (Figure 1C). In a similar experiment, using MUC1*-positive T47D breast cancer cells stably transfected with siRNA targeting MUC1, or control siRNA, stimulation of growth only occurred with control-transfected cells, further demonstrating specificity of antibody (Figure 1D). Identical results were demonstrated for MUC1*'s natural ligand, NM23 (Figure 1E).

[0059] Example 4. NM23 binds specifically to the PSMGFR peptide which is comprised essentially of the extracellular domain of MUC1*. Binding was measured by Surface Plasmon Resonance, using a Biacore3000 instrument and BiaEvaluation software. Histidine-tagged $MUC1^*_{1110\text{-}ecd}$ (SEQ ID NO:5) or irrelevant peptide (HHHHHH-SSSSGSSSSGSSSSGGRGDSGRGDS – SEQ ID NO:34) were immobilized on separate flow channels of 5.7% NTA-Ni^{++} SAM-coated SPR chips, prepared in our lab as described in Mahanta et al. 2008. 35 μL plugs of NM23, purified bovine or recombinant human, were injected into a constant flow stream of 5uL/minute and sensograms were recorded. NM23 purified from bovine liver (Sigma N-2635) was diluted in PBS alone. Affinities were measured over a wide range of concentrations using a 1:1 Langmuir model. Actual affinities may vary as first order kinetics cannot adequately describe this system. (Figure 1F).

[0060] Example 5. MUC1* Growth Factor Receptor and its ligand NM23 are on undifferentiated hESC, but not differentiated hESC. Human embryonic stem cells in the undifferentiated (pluripotent) state or in the newly differentiating state were analyzed by immunocytochemistry (ICC). Human embryonic stem cells (hESCs) were manually dissected and plated in 8-well chamber slides (Nunc) that had been pre-coated with matrigel. For undifferentiated cells, cells were fixed 5-7 days after plating. For differentiated cells, bFGF was removed from the culture medium 5-7 days after plating and cells were allowed to differentiate for 14 days before fixation. Cells were washed with PBS prior to fixation with 4% paraformaldehyde in 0.1M cacodylate buffer for 15 minutes at 4°C. Cells were blocked for 1

hour with 1% BSA and 1% donkey or goat serum in PBS. 0.1% NP-40 was used with antibodies against intracellular antigens. Primary antibodies were diluted in block and incubated with cells for 1 hour at 4°C. The primary antibodies for the following proteins were used: OCT4 (Santa Cruz, Clone Clones H-134 and C-10, 1:100 dilution), full-length MUC1 (VU4H5, Santa Cruz Biotechnology, 1:50 dilution), MUC1* (Minerva, 1:250 dilution), or NM23 (Santa Cruz, Clone NM301, 1:100 dilution)). Cells were washed 3 times in PBS for 5 minutes prior to incubation for 30 minutes with secondary antibodies: AlexaFluor 488 Goat anti-rabbit IgG, AlexaFluor 555 Goat anti-mouse IgG, AlexaFluor 350 Goat anti-rabbit IgG (Invitrogen, 1:200); Goat anti-mouse IgM-TR (Santa Cruz, 1:100). Cells were washed 3 times in PBS for 5 minutes prior to coverslip mounting using an anti-fade mounting medium (Biomeda). Nuclei were visualized by DAPI staining (1μg/ml) for 5 minutes. Immunostained cells were visualized on an Olympus BX-51 epifluorescent microscope. Results of these experiments show that MUC1* is on the surface of undifferentiated cells (pluripotent stem cells) (Fig 2A, 3B, 3C) but is not on differentiated hESCs (Fig 2 D). Figure 3 shows that the ligand of MUC1*, NM23, co-localizes with MUC1* (Figs. 3 A-C). MUC1* and its ligand NM23 are only expressed on pluripotent stem cells (OCT4-positive cells) and not on those that have differentiated, Figs. 3C and 3F (DAPI stains nuclei of OCT4-negative cells).

[0061] Example 6. MUC1* mediates growth of pluripotent stem cells.

[0062] The following experiment was performed to determine the effect of stimulating MUC1*, using a bivalent anti-MUC1*, on pluripotent stem cells. The results show that adding a MUC1* dimerizing ligand stimulates pluripotent (OCT4-positive) stem cell growth and also enables their growth in the absence of feeder cells, their extracts or bFGF.

[0063] Long term growth of pluripotent (OCT4-positive) hESC is mediated by MUC1* stimulation. hESCs were trypsin-dissociated and seeded in 8-well chamber slides pre-coated with matrigel at 4 x 10^4 cells/well. Media was changed and antibodies added every other day at a final concentration of 1 μg/ml for bivalent anti-MUC1* until discrete colonies were visible. Culture conditions include 'minimal stem cell medium' (hESC media without feeder-conditioned medium) and Hs27-conditioned medium, with and without bFGF supplementation. For each condition, cells were grown in quadruplicate. Cells were washed with PBS and fixed, and OCT4 immunostaining was conducted as described above. Figure 4, panels A-D are photos of cells grown over matrigel and conditioned medium from fibroblast feeder cells added. Panels E-H are

photos of cells grown over matrigel in which no conditioned medium from fibroblast feeder cells was added. The addition of anti-MUC1* antibody to cell cultures (Fig. 4 C, D) resulted in more pluripotent stem cells than growth supplemented by bFGF (Fig. 4 A, B). The addition of anti-MUC1* antibody to cells cultured in the absence of conditioned medium from fibroblast feeder cells (Fig. 4 G, H) resulted in an abundance of pluripotent stem cells, in sharp contrast to cells grown by adding bFGF (Fig. 4 E, F), which resulted in no pluripotent cells (absence of OCT4).

[0064] Example 7. The effect of stimulating MUC1* to enhance the growth of pluripotent stem cells was directly measured in a quantitative Calcein assay. Human embryonic stem cells (hESCs) were manually dissected and grown on matrigel-coated wells of a 96 well plate at a density of 1.9x 10^4 cells/well. Culture media contained hESC media supplemented with 30% Hs27-conditioned medium and 4 ng/ml bFGF. Antibodies were added at a final concentration of 1 μg/ml for bivalent anti-MUC1* and 100 μg/ml for monovalent anti-MUC1*. Experiments were performed in triplicate. 41 hours-post antibody treatment, live and dead cells were quantified with the LIVE/DEAD viability/cytotoxicity kit (Molecular Probes), following manufacturer's instructions. Fluorescence was measured using a Victor3V plate reader (Perkin Elmer). The bar graph of Figure 5 shows that stimulation of MUC1* using a dimerizing ligand (anti-MUC1*) enhanced stem cell growth, while blocking the extracellular domain of MUC1*, with the anti-MUC1* Fab, resulting in total stem cell death.

[0065] Example 8. A long-term stem cell growth experiment was done to compare the effects of stimulating the growth of stem cells using a bivalent anti-MUC1* antibody, NM23, NM23-mutant, or bFGF. hESCs were dissociated with trypsin and seeded in 8-well chamber slides pre-coated with Matrigel at a cell density of 8.2x10^4 cells/well. Media was changed and antibodies or wild type or mutant NM23 proteins were added every other day at final concentrations of 80ng/ml for Anti-MUC1* antibody, 1nM for wild type recombinant NM23 or mutant (S120G) NM23, or recombinant bFGF at a final concentration of 4ng/ml in 'minimal stem cell medium' (hESC media without feeder-conditioned medium). Cells were also grown as a control in minimal stem cell medium with 30% conditioned medium from Hs27 fibroblasts and 4ng/ml recombinant bFGF (Peprotech #100-18B). Results of this experiment show that MUC1* ligands do a better job of stimulating growth in minimal media of pluripotent colonies than does conditioned media plus bFGF, the 'normal' growth medium of these cells on Matrigel. Table 1 details the results.

[0066] **Table 1. hESCs cultured in minimal media for 4 weeks**

Growth condition	Week 1st colony appeared	Number of colonies	Morphology
Minimal Stem Cell Growth Media			
NM23 1nM	Week 2	2 colonies	2 large undifferentiated colonies in 1 of 1 wells; centers of colonies appear to begin to differentiate during week 3; by end of week 4, most of each colony remains undifferentiated
NM23-S120G mutant 1nM	Week 2	7 colonies	7 large undifferentiated colonies in 1 of 1 wells; centers of colonies appear to begin to differentiate during week 3; by end of week 4, most of each colony remains undifferentiated
anti-MUC1* 80 ng/ml	Week 2	5 colonies	7 large undifferentiated colonies in 1 of 2 wells; centers of colonies appear to begin to differentiate during week 3; by end of week 4, most of each colony remains undifferentiated
bFGF 4 ng/ml	-	0	No colonies
nothing	Week 2	2 colonies	2 very small, differentiated colonies
***Control** - 30% Conditioned Media from Hs27 Fibroblast Feeder Cells*			
bFGF 4 ng/ml	Week 2	5	5 mostly differentiated colonies

[0067] Example 9. MUC1* translocates to nucleus of cells. Anti-MUC1* monoclonal Ab was labeled *in vitro* with Alexa 555 dye, and bound at 4C to HCT-116 cells (MUC1-negative) transfected with MUC1*, that had been washed in cold PBS, at 4C. After 20 min, cells were washed twice in cold PBS, and cells were either fixed in 4% paraformaldehyde, or incubated with pre-warmed growth medium. Cells were washed after 40 minutes, and fixed with 4% paraformaldehyde for 5minutes, then blocked and permeabilized with 2.5%, BSA 2.5% FBS and 0.1% NP-40 in PBS. Endosomes were stained using an anti-EEA1 antibody (Cell Signaling Technologies, 2411S) and Alexa 488 (Invitrogen 1:200) (Fig. 6).

[0068] Example 10. MUC1* translocates to the nucleus where it functions as a transcription factor or co-factor.

[0069] It was previously reported in the scientific literature that an artificially expressed peptide, corresponding to the cytoplasmic domain of MUC1 (MUC1-CD), translocated to the

nucleus (Huang, et al. 2005). Another study reported that the MUC1-CD bound to the p53 promoter site (Wei, et al. 2007). The Wei et al., 2007, report did not however determine whether in this capacity, the MUC1-CD up- or down-regulated transcription of p53. To induce or maintain stemness or pluripotency, it is desirable to suppress p53 (Maimets, et al., 2008). Further, the forced expression of the cytoplasmic domain of MUC1, alone, does not correspond to any naturally occurring state or cleavage state of MUC1. We therefore sought to determine whether or not it was actually MUC1*, which is a marker of pluripotency, that is translocated to the nucleus.

[0070] A polyclonal antibody, anti-MUC1*, that recognizes the PSMGFR epitope on MUC1* but not the same epitope when MUC1 is full-length (Mahanta et al 2008) was used to stain human embryonic stem cells according to the methods of Hikita et al (2008). Embryonic stem cells exclusively express MUC1* but not full-length MUC1. Immunocytochemistry showed that MUC1* was often detected in the nuclei of embryonic stem cells. A fluorescently tagged anti-MUC1* Fab fragment was incubated with cells that had been transfected with MUC1* or the empty vector. Excess Fab was removed, and cells were either washed at 4° C in PBS and fixed in paraformaldehyde, or incubated with prewarmed medium at 37° C for 10 min, 25 min, or 40 min to facilitate receptor internalization before paraformaldehyde fixation. Cells were moved to 4° C to stop receptor internalization at various timepoints. The Fab bound specifically to MUC1*-transfected cells, and not to vector-transfected cells (data not shown). Figure 6 shows that after 40 minutes, MUC1* is translocated to the nucleus of the cell.

[0071] Example 11. Inhibit p53 and its pro-apoptotic and growth inhibitory effects, using MUC1-associated factors, to increase the efficiency of generating iPS (induced pluripotent stem) cells.

[0072] The small molecule Nutlin increases activity of p53 by interfering with the interaction between p53 and its natural inhibitor hDM2 (Vassilev et al., 2004) . The induction of p53 activity by Nutlin drives the differentiation of human embryonic stem cells (Maimets, et al., 2008). Similarly, overexpression of p53 in embryonic stem cells inhibits hESC growth, most likely through the induction of apoptosis (Maimets, et al., 2008). In p53-null mice, the establishment of primary tumors was enhanced (Zhou et al., 2000). Thus, interfering with p53 activity increases the efficiency of establishing iPS cell lines.

[0073] During the establishment of novel iPS lines using a "core set" of pluripotency genes, OCT4, SOX2 and KLF4, induction of pluripotency is enhanced by inhibiting p53 first by the small molecule Nutlin and then by introducing MUC1*. MUC1* is introduced to cells undergoing induction of pluripotency by: 1) transfecting DNA that encodes MUC1*; and b) by adding a recombinant MUC1* protein that has been modified with a poly-arginine tract to efficiently enter the cell.

[0074] The core set of pluripotency genes was transfected into dermal fibroblasts. However, the exogenous expression of the direct or indirect gene products into dermal fibroblasts or other mature cells is also anticipated. The invention also anticipates that the addition of MUC1, MUC1* or associated factors can eliminate some, or all, of the core set of pluripotency factors.

[0075] Induction of pluripotency is also enhanced when a peptide corresponding to the cytoplasmic tail of MUC1 is exogenously added to cells undergoing conversion to pluripotency. Optionally, MUC1-CD is modified with a leader sequence such as a poly-Arginine tract that allows the peptide to enter the cell.

[0076] Example 12. The introduction of MUC1-related proteins enhances pluripotency by inducing transcription of c-myc and by activating MUC1*.

[0077] c-Myc has been shown to enhance the induction of pluripotency. NM23 induces transcription of c-myc (Dexheimer et al., 2009) and eliminates the need for c-myc. NM23 is introduced to cells undergoing conversion to iPS by transfection of the encoding nucleic acids or by exogenously adding the protein itself which may be modified with sequences, such as a poly-Arginine tract, to aid in cellular entry. NM23, wild type or mutant S120G that prefers dimer formation, is added to dermal fibroblasts that have been transfected with OCT4, SOX2 and KLF4. The efficiency of iPS cell generation is enhanced.

[0078] NM23 (NME-H1, NME-H2 or NME-7) enhances the induction or maintenance of pluripotency. NM23 is introduced along with previously identified pluripotency factors, including but not limited to OCT4, SOX2, KLF4, as well as others disclosed herein.

[0079] Example 13. Identification of a new core set of pluripotency factors that include MUC1* associated factors

[0080] MUC1* is introduced to cells to induce or maintain pluripotency or to improve the efficiency of iPSc formation or to replace one or more of the pluripotency gene set comprised of OCT4, SOX2, KLF4. A DNA construct containing nucleic acid encoding MUC1* is transfected

into dermal fibroblasts along with combinations of the above-mentioned set of pluripotency genes (or their gene products). The efficiency of iPS colony formation is determined by enumerating the number of stem cells generated. Further, cells are analyzed by immunofluorescent detection of pluripotency markers, such as OCT4, SSEA 1, 3, 4, TRA 1-60 TRA 1-81, TRA 2-49/6E (alkaline phosphatase), and NANOG. Resultant cells are evaluated for karyotype stability and the ability to differentiate along the three different germlines (mesoderm, endoderm and ectoderm). This is determined by immunofluorescent detection using antibodies against germline-specific markers, such as CD34 or smooth muscle actin for mesoderm detection, GATA-4 or cytokeratin 19 for endoderm detection, and Nestin or beta-III tubulin for ectoderm detection. A MUC1* activating ligand, preferably anti-MUC1* antibody or NM23 (NME), is optionally added to further enhance the induction or maintenance of pluripotency.

REFERENCES

[0081] Aoi, T. et al. Science 321, 699–702 (2008).

[0082] Boyer L. A. et al. Cell 122, 947-956 (2005)

[0083] Dexheimer at al. Mol Cancer Ther 8, 1363-1377 (2009)

[0084] Fessler S. et al. Breast Cancer Res Treat. May 5. (2009)

[0085] Foster, K. W. et al. Oncogene 24, 1491-1500 (2005)

[0086] Hikita et al. PLoS ONE 3, e3312 (2008)

[0087] Huang, L. et al. Cancer Res., 65(22):10413-22 (2005)

[0088] Huangfu D. et al. Nat Biotechnol 26, 795-797 (2008)(a)

[0089] Huangfu D. et al. Nat Biotechnol 26, 1269-1275 (2008)(b)

[0090] Jaenisch, R. and Young, R. Cell 132, 567-582 (2008)

[0091] Kaji K., et al. Nature 458:771–775 (2009)

[0092] Kawamura et al., Nature 460(7259),1140-4 (2009)

[0093] Kim, et al. Biochem Biophys Res Commun 307: 281–289 (2003)

[0094] Komarov, et al. Science 285(5434), 1733-7 (1999)

[0095] Lin T et al. Nat Cell Biol 7, 165-171 (2005)

[0096] Lowry, W.E. et al. Proc. Natl. Acad. Sci. USA 105, 2883–2888 (2008).

[0097] Lyssiotis et al. Proc Natl Acad Sci USA 106, 8912-8917 (2009)

[0098] Mahanta et al. PLoS ONE 3, e2054 (2008)

[0099] Maherali, N. et al. Cell Stem Cell 1, 55–70 (2007).

[00100] Maimtes T et al., Oncogene 27, 5277–5287 (2008)

[00101] Nakagawa, M. et al. Nature Biotechnol 26, 101–106 (2008)

[00102] Okita, K et al. Nature 448, 313–317 (2007).

[00103] Okita K et al. Science 322, 949-953 (2009)

[00104] Park, I.H. et al. Nature 451, 141–146 (2008).

[00105] Raina et al. Cancer Res 69, 5133-5141 (2009)

[00106] Soldner F., et al. Cell 136:964–977 (2009)

[00107] Sommer C A et al., Stem Cells, 27(3), 543-9 (2009)

[00108] Stadtfeld M., et al. Science 322, 945-949 (2009)

[00109] Strom, et al. Nat Chem Biol. 2(9):474-9 (2006)

[00110] Takahashi, K. & Yamanaka, S. Cell 126, 663–676 (2006).

[00111] Takahashi, K. et al. Cell 131, 861–872 (2007).

[00112] Thathiah, A et al. J Biol Chem 274, 3386-3394 (2003)

[00113] Vassilev L. T. et al., Science 303, 844–848. (2004)

[00114] Wei et al. Cancer Res 67, 1853-1858 (2007)

[00115] Wernig, M. et al. Nature 448, 318–324 (2007).

[00116] Wernig M. et al. Cell Stem Cell 2, 10-12 (2008)

[00117] Woltjen K., et al. Nature 458, 766–770 (2009)

[00118] Yamanaka, S. Cell stem Cells 1, 39-49 (2007)

[00119] Yu J. et al. Science 324, 797-801 (2009)

[00120] Yu, J. et al. Science 318, 1917–1920 (2007)

[00121] Zhou et al. MCB, 20, 628-633 (2000

[00122] Zhou et al. Cell Stem Cell 4, 381-384 (2009)

[00123] All of the references cited herein are incorporated by reference in their entirety.

* * * * *

[00124] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention

specifically described herein. Such equivalents are intended to be encompassed in the scope of the claims.

In the claims:

1. A method for inducing or maintaining pluripotency in a cell comprising contacting the cell with a biological or chemical species that increases MUC1* activity by increasing MUC1 expression, increasing expression of its cleavage enzymes or by dimerizing MUC1*.

2. A method as in Claim 1, wherein the biological species is a gene.

3. A method as in Claim 1, wherein the biological species is a protein.

4. A method as in Claim 3, wherein the protein is a ligand that dimerizes MUC1*.

5. A method as in Claim 4, wherein the ligand is NM23 in its dimeric form.

6. A method as in Claim 4, wherein the protein is an antibody that recognizes the PSMGFR sequence of MUC1*.

7. A method as in Claim 1, wherein the chemical species is a small molecule.

8. A method as in Claim 7, wherein the small molecule enhances the transcription of MUC1, its cleavage enzyme, or NM23.

9. A method as in Claim 8, wherein the cleavage enzyme is MMP-16, MMP-14 or ADAM-17.

10. A method as in Claim 7, wherein the small molecule enhances cleavage of MUC1.

11. A method as in Claim 10, wherein the small molecule is phorbol ester.

12. A method as in Claim 2, wherein the gene encodes MUC1.

13. A method as in Claim 2, wherein the gene encodes MUC1*.

14. A method as in Claim 2, wherein the gene encodes a ligand of MUC1*.

15. A method as in Claim 14, wherein the ligand is a MUC1* antibody or NM23.

16. A method as in Claim 3, wherein the protein is MUC1* or an effective fragment thereof.

17. A method as in Claim 16, wherein the fragment is the cytoplasmic tail.

18. The method according to claim 1, further comprising contacting the cell with a molecule that increases expression of gene products that induce pluripotency.

19. A method as in Claim 18, wherein the molecule increases expression of OCT4, SOX2, NANOG, KLF4 or LIN28.

20. The method as in Claim 19, wherein the molecule increases expression of OCT4, and SOX2.

21. The method according to claim 1, wherein the cell is somatic or mature cell.

22. The method according to claim 1, wherein the cell is immature cell.

23. The method according to claim 22, wherein the cell is somewhat differentiated cell.