Methods and products for biasing cellular development

Kiminobu Sugaya  
*University of Central Florida*

Angel Alvarez  
*University of Central Florida*

Amelia Marutle  
*University of Central Florida*

Find similar works at: https://stars.library.ucf.edu/patents

University of Central Florida Libraries http://library.ucf.edu

This Patent is brought to you for free and open access by the Technology Transfer at STARS. It has been accepted for inclusion in UCF Patents by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

**Recommended Citation**

Sugaya, Kiminobu; Alvarez, Angel; and Marutle, Amelia, "Methods and products for biasing cellular development" (2015). *UCF Patents*. 369.  
https://stars.library.ucf.edu/patents/369
(54) METHODS AND PRODUCTS FOR BIASING
CELLULAR DEVELOPMENT

(71) Applicants: Kaminobu Sugaya, Winter Park, FL (US); Angel Alvarez, Chicago, IL (US); Amelia Marutle, Stockholm (SE)

(72) Inventors: Kaminobu Sugaya, Winter Park, FL (US); Angel Alvarez, Chicago, IL (US); Amelia Marutle, Stockholm (SE)

(73) Assignee: University of Central Florida Research Foundation, Inc., Orlando, FL (US)

(45) Date of Patent: Sep. 8, 2015

7,635,467 B2 12/2009 Sugaya et al.

OTHER PUBLICATIONS

Molecular and Cellular Neuroscience 17, 671-682, 2001.*


* cited by examiner

Primary Examiner — Nancy T Vogel
Attorney, Agent, or Firm — Timothy H. Van Dyke; Beusse, Welter, Sunks & Maire, PLLC

(57) ABSTRACT

Methods are described that bias cells, such as potent and multipotent stem cells, by transfection with a nucleic acid sequence, to differentiate to a desired end-stage cell or a cell having characteristics of a desired end-stage cell. In particular embodiments, human neural stem cells are transfected with vectors comprising genes in the homebox family of transcription factor developmental control genes, and this results in a greater percentage of resultant transformed cells, or their progeny, differentiating into a desired end-stage cell or a cell having characteristics of a desired end-stage cell.

5 Claims, 14 Drawing Sheets
prior art

FIG. 8
Cells were double-immunofluorescence stained with:

a) β III-tubulin (red) and b) CHAT (green), markers for cholinergic neurons.

c) co-localization of β III-tubulin & CHAT.

Blue signal is a counter staining for nuclei by DAPI.

FIG. 12
Cells were double-immunofluorescence stained with:

a) βIII-tubulin (green) and b) ChAT (red), markers for cholinergic neurons;
c) co-localization of βIII-tubulin & ChAT. (20 x magnification) Blue signal is a counter staining for nuclei by DAPI.
d) Insert for non-specific staining for βIII-tubulin & ChAT, respectively.

FIG. 13
FIG. 14

a) βIII-tubulin (green) and b) ChAT (red), markers for cholinergic neurons. c) co-localization of βIII-tubulin & ChAT. Blue signal is a counter staining for nuclei by DAPI. (40 x magnification)
Cells were double-immunofluorescence stained with βIII-tubulin (green) and nuclear counter staining by DAPI (blue). No Chat-positive neurons were observed. (a-b) 10 x and (c) 20 x magnification, respectively.

FIG. 15
METHODS AND PRODUCTS FOR BIASING CELLULAR DEVELOPMENT

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Ser. No. 60/621,483 filed Oct. 22, 2004; and is a divisional application of U.S. application Ser. No. 13/330,957, filed Dec. 20, 2011, which is a divisional application of U.S. application Ser. No. 11/258,603, filed Oct. 24, 2005, which hereby incorporated in its entirety.

FIELD OF INVENTION

The present invention is directed to methods and systems directed to altering the differentiation of a cell, more particularly to biasing a multipotent stem cell by transferring the cell with a nucleic acid sequence comprising a desired gene, the gene being expressed so that the cell, or its progeny, differentiate to a desired end-stage cell.

BACKGROUND

Proper cellular function and differentiation depends on intrinsic signals and extracellular environmental cues. These signals and cues vary over time and location in a developing organism (i.e., during embryogenesis), and remain important in developing and differentiating cells during post-natal growth and in a mature adult organism. Thus, in a general sense, the interplay of the dynamically changing set of intracellular dynamics (such as manifested by intrinsic chemical signaling and control of gene expression) and environmental influences (such as signals from adjacent cells) determine cellular activity. The cellular activity so determined is known to include cell migration, cell differentiation, and the manner in which a cell interacts with surrounding cells.

The use of stem cells and stem-cell-like cells of various types for cell replacement therapies, and for other cell-introduction-based therapies, is being actively pursued by a number of researchers. Embryonic stem cells from a blastocyst stage are frequently touted for their pluripotency—that is, their ability to differentiate into all cell types of the developing organism. Later-stage embryonic stem cells, and certain cells from generative areas of an adult organism, are identified as more specialized, multipotent stem cells. These cells include cells that are able to give rise to a succession of a more limited subset of mature end-stage differentiated cells of particular types or categories, such as hematopoietic, mesenchymal, or neuroectodermal end-stage differentiated cells. For example, a multipotent neural stem cell may give rise to one or more neuron cell types (i.e., cholinergic neurons, dopaminergic neurons, GABAergic neurons), which includes their specific cell classes (i.e., basket cell or a chandelier cell for GABAergic neurons), and to non-neuron glial cells, such as astrocytes and dendrocytes.

Further along the path of differentiation are cells derived from multipotent stem cells. For example, derivatives of a localized, non-migrating neuroectodermal type stem cell may migrate but, compared to their multipotent parent, have more limited abilities to self-renew and to differentiate (See Stem Cell Biology, Marshak, Gardner & Gottlieb, Cold Spring Harbor Laboratory Press, 2001, particularly Chapter 18, p. 407). Some of these cells are referred to a neuron-restricted precursors ("NRPs"), based on their ability, under appropriate conditions, to differentiate into neurons. There is evidence that these NRPs have different subclasses, although this may reflect different characteristics of localized multipotent stem cells (Stem Cell Biology, Marshak et al., pp 418-419).

One advantage of use of multipotent and more committed cells further along in differentiation, compared to pluripotent embryonic stem cells, is the reduced possibility that some cells introduced into an organism from such source will form a tumor (Stem Cell Biology, Marshak et al., p. 407). However, a disadvantage of cells such as cell types developed from multipotent stem cells, for instance, embryonic progenitor cells, is that they are not amenable to ongoing cell culture. For instance, embryonic neural progenitor cells, which are able to differentiate into neurons and astrocytes, are reported to survive only one to two months in a cell culture.

Generally, it is known in the art that the lack of certain factors critical to differentiation will result in no or improper differentiation of a stem cell. Researchers also have demonstrated that certain factors may be added to a culture system comprising stem cells, such as embryonic stem cells, so that differentiation to a desired, stable end-stage differentiated cell proceeds. It also is known in the art to introduce and express a transcription factor gene, Nurr1, into embryonic stem cells, and then process the cells through a five-step differentiation method (Kim, Jong-Hoon et al., Dopamine Neurons Derived from Embryonic Stem Cells Function in an Animal Model of Parkinson’s Disease, Nature, 418:50-56 (2002)), resulting in differentiated cells having features of dopaminergic cells. However, the starting cell for this was an embryonic stem cell, and the differentiation process through to a cell having the features of a dopaminergic neuron, requires substantial effort that includes the addition and control of endogenous factors. In addition, because the starting cell is an early-stage embryonic stem cell having pluripotency, there is a relatively higher risk that some cells implanted from this source will become tumorigenic.

Also, without being bound to a particular theory, it is believed that to the extent a particular method of differentiation results in a greater percentage of cells that are dedicated or predetermined to differentiate to a desired functional cell type (i.e., a cholinergic neuron), this reduces the chance of tumor formation after introduction of cells derived from such method. As disclosed herein, embodiments of the present method that utilize multipotent stem cells as the starting material provide an increased percentage of cells predisposed (i.e., biased) to or differentiated to a desired cell type. This is believed to provide for reduced risk of tumor formation equivalent to or superior to the use of more differentiated cells such as NRPs.

There are many possible applications for methods, compositions, and systems that provide for improved differentiation of stem cells to a desired functional, differentiated cell. For example, not to be limiting, millions of people suffer from deafness and balance defects caused by damage to inner ear hair cells (IEHCs), the primary sensory receptor cells for the auditory and vestibular system after exposure to loud noises, antibiotics, or antitumor drugs. Since IEHCs rarely regenerate in mammals, any damage to these organs is almost irreversible, precludes any recovery from hearing loss, and results in potentially devastating consequences. Current therapies utilizing artificial cochlear implants or hearing aids may partially improve but not sufficiently restore hearing. Therefore, cell therapy to replace the damaged IEHC may be one of the most promising venues today. In the past, IEHC production from progenitor cells from the vestibular sensory epithelia of the bullfrog [Cristobal, 1998 #28] and possible existence of IEHC progenitors in mammalian cochlea sensory epithelia [Kojima, 2004 #29] has been reported. However limited quantity of IEHC progenitor prevents clinical
application of this type of cell to treat deafness. Thus novel technology to produce IEHCs from other cell sources is needed.

While stem cells are known to be the building blocks responsible for producing all of a body’s cells, the specific differentiation process towards IEHC lineage is not clear. Embryonic stem cells transplanted into the inner ear of adult mice or embryonic chickens did not differentiate into IEHCs {Sakamoto, 2004 #19}. Neural stem cells (NSCs) grafted into the modiolus of cisplatin-treated cochleae of mice only differentiated into glial or neuronal cells within the cochlea {Tamara, 2004 #18}. In order to produce IEHCs from these stem cells, modification or direction of the cell fate decision may be needed.

Another possible application for methods, compositions, and systems of the present invention is biasing Human Neural Stem Cells (“HNSCs”) to differentiate to cholinergic neurons, or to cells having characteristics of cholinergic neurons. Such biasing would provide for an improved percentage of such stem cells in a culture vessel to differentiate to this desired end-stage nerve cell. Improvements to the percentage of cells that are known to be biased to differentiate to this end-stage neuron cell, or to cells having characteristics of a cholinergic neuron, may lead to improvements both in research and treatment technologies for diseases and conditions that involve degeneration or loss of function of cholinergic neurons. Alzheimer’s disease is one example of a malady known to be associated with degeneration of the long-projecting axons of cholinergic neurons.

Thus, there is a need in the art to improve the compositions, methods and systems that provide biased and/or differentiated cells from stem cells or stem-cell-like cells. More particularly, a need exists to obtain a higher percentage of desired end-stage neuron cell, such as cholinergic neurons or inner ear hair cells. The present invention addresses these needs.

**BRIEF DESCRIPTION OF THE FIGURES**

FIGS. 1-16 are appended hereto, are part of the specification, and are described herein and/or on the figures themselves.

**FIG. 1** A Hath1 gene (SEQ ID NO:4) was amplified from the *homo sapiens* BAC clone RP11-680J17 by PCR and then cloned it into a mammalian expression directional cloning vector, pcDNAHismax TOPO TA.

**FIG. 2** The hair cell marker calretinin was identified on certain cells via immunocytochemistry.

**FIG. 3** Math1 full length mRNA was amplified by RT-PCR and digested with Apai, which cuts position 441 of Hath1ORF. Expecting fragment sizes are 441 by and 624 bp. M: 100 bp marker; 5, 6, without transfection (control); 7, 8: transfected with mammalian expression vector containing Hath2.

**FIG. 4** The presence of the actual protein calretinin on the cell surface was determined via Western Blot. Seven days for differentiation was allowed before any analysis of the cells. Protein was isolated from the cells and calretinin was identified in the cell isolate.

**FIG. 5** Cells transfected with Hath1 and grown to allow for differentiation were visualized using Transmission Electron Microscopy. A subset of the cells exhibited distinct hair-like projections. These were the actual hairs from the transfected HNSCs that differentiated into cells having this characteristic feature of IEHCs.
articles, scientific publications, and other references referenced herein are hereby incorporated by reference in this application in order to more fully describe the state of the art to which the present invention pertains.

Reference to particular buffers, media, reagents, cells, culture conditions and the like, or to some subclass of same, is not intended to be limiting, but should be read to include all such related materials that one of ordinary skill in the art would recognize as being of interest or value in the particular context in which that discussion is presented. For example, it is often possible to substitute one buffer system or culture medium for another, such that a different but known way is used to achieve the same goals as those to which the use of a suggested method, material or composition is directed.

It is important to an understanding of the present invention to note that all technical and scientific terms used herein, unless defined herein, are intended to have the same meaning as commonly understood by one of ordinary skill in the art. The techniques employed herein are also those that are known to one of ordinary skill in the art, unless stated otherwise. For purposes of more clearly facilitating an understanding the invention as disclosed and claimed herein, the following definitions are provided.

DEFINITIONS

Stem cells are undifferentiated cells that exist in many tissues of embryos and adult organisms. In embryos, blastocyst stem cells are the source of cells that differentiate to form the specialized tissues and organs of the developing fetus. In adults, specialized stem cells in individual tissues are the source of new cells, replacing cells lost through cell death due to natural attrition, disease, or injury. Stem cells may be used as substrates for producing healthy tissue where a disease, such as the epithelium or blood cells.


---

1. Proliferate: Stem cells are capable of dividing to produce daughter cells.
2. Exhibit self-maintenance or renewal over the lifetime of the organism: Stem cells are capable of reproducing by dividing symmetrically or asymmetrically to produce new stem cells. Symmetric division occurs when one stem cell divides into two daughter stem cells. Asymmetric division occurs when one stem cell forms one new stem cell and one progenitor cell. Symmetric division is a source of renewal of stem cells. This permits stem cells to maintain a consistent level of stem cells in an embryo or adult mammal.
3. Generate large number of progeny: Stem cells may produce a large number of progeny through the transient amplification of a population of progenitor cells.
4. Retain their multilineage potential over time: The various lines of stem cells collectively are the ultimate source of differentiated tissue cells, so they retain their ability to produce multiple types of progenitor cells, which in turn develop into specialized tissue cells.
5. Generate new cells in response to injury or disease: This is essential in tissues which have a high turnover rate or which are more likely to be subject to injury or disease, such as the epithelium or blood cells.

Thus, key features of stem cells include their capability of self-renewal, and their capability to differentiate into a range of end-stage differentiated tissue cells.

By “neural stem cell” (NSC) is meant a cell that (i) has the potential of differentiating into at least two cell types selected from a neuron, an astrocyte, and an oligodendrocyte, and (ii) exhibits self-renewal, meaning that at a cell division, at least one of the two daughter cells will also be a stem cell.

By “potential cell” is meant a stem cell that has the capability to differentiate into a number of different types of end-stage cell types, and to self-renew, and may include stem cells classified as pluripotent, multipotent, or cells more differentiated than multipotent (i.e., a dedicated progenitor) under different stem cell classification schemes.

By “a presumptive end-stage cell” is meant a cell that has acquired characteristics of a desired end-stage cell type, but which has not been conclusively identified as being the desired end-stage cell. A presumptive end-stage cell possesses at least two, and often more, morphological and/or molecular phenotypic properties of the desired end-stage cell.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See, for example, Molecular Cloning A Laboratory Manual, 3 Ed. ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 2001); DNA Cloning, Volumes I and II (D. N. Glover ed., 1985); Oligonucleotide Synthesis (M. J. Gait ed., 1984); Mullis et al., U.S. Pat. No. 4,683,195; Nucleic Acid Hybridization (B. D. Hames & S. J. Higgins eds. 1984); Transcription And Translation (B. D. Hames & S. J. Higgins eds. 1984); Culture Of Animal Cells (R. I. Freshney, Alan R. Liss, Inc., 1987); Immobilized Cells And Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide To Molecular Cloning (1984); the treatise, Methods In Enzymology (Academic Press, Inc., N.Y.); Gene Transfer Vectors For Mammalian Cells (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); Methods In Enzymology, Vols. 154 and 155 (Wu et al. eds.), Immunochemical Methods In Cell And Molecular Biology (Mayer and Walker, eds., Academic Press, London, 1987); Handbook Of Experimental
The present invention provides for more consistent differentiation of desired cell type under appropriate external conditions. This is done in some embodiments so that in a pre-implantation culture, the desired ratio so that a higher percentage of cells introduced into a particular cell medium, a tissue culture, or a living organism in a particular location differentiate into the end-stage differentiated cell type that is desired. Without being bound to a particular theory, this is believed to increase the probability of success and overall effectiveness, and to decrease the risks associated with implantation of cells obtained from embryonic stem cells or embryonic-cell-like cells.

While not meant to be limiting as to the type of nucleic acid sequence introduced, examples herein utilize introduction to a cell of a nucleic acid sequence comprising a homebox gene. This is a gene group that includes a number of known developmental control genes. A homeobox gene is a gene containing an approximately 180-base-pair segment (the “homebox”) that encodes a protein domain involved in binding to (and thus regulating the expression of) DNA. The homebox segment is remarkably similar in many genes with different functions. However, specific homeobox genes are known to operate at different stages, and in different tissue environments, to yield very different specific results. For example, in relatively early embryological development in the vertebrate embryo, expression of genes of the Hox family of homeobox genes appears to affect development of the brain based on position along the anterior/posterior axis. This is believed to control identity and phenotypic specializations of individual rhombomeres. (Fundamental Neuroscience, Zigmond et al., p. 435). Later in development, LIM homebox gene expression is associated with the projection pattern of developing primary motor neurons, and more generally, expression of a particular combination of LIM homeobox genes appears to be related to motor neuron subtype identity and to targeting specificity (Fundamental Neuroscience, Zigmond et al., p. 507). Also, some LIM homebox genes appear to affect developmental progression, rather than fate, of motor neurons, which suggests a role of cell-to-cell signaling in the embryo to fully effectuate the differentiation in vivo (Fundamental Neuroscience, Zigmond et al., pp. 443-444). These highly specialized and variable roles for homeobox genes in general, and for LIM family homebox genes more particularly, demonstrate the subtle, specific, and highly variable effects that these genes may have on cell and tissue development and differentiation.

Further with regard to function of homeobox genes, these genes encode transcriptional regulators that play critical roles in a variety of developmental processes. Although the genetic and developmental mechanisms that control the formation of forebrain cholinergic neurons are just beginning to be elucidated, it is known that the vast majority of forebrain cholinergic neurons derive from a region of the subcortical telencephalon that expresses the Ntx2-1 homeobox gene.

It has recently been reported that Ntx2-1 appears to specify the development of the basal telencephalon by positively regulating transcription factors such as the LIM-homeobox genes Lhx8 (also known as L3 or Lhx7) and Gbx1, which are associated with the development of cholinergic neurons in the basal forebrain (Zhao et al., 2003; Asbreuk et al., 2002, Waters et al., 2003).

In the spinal cord, Isl1, Lhx1, Lhx3 and Lhx4 have been shown to be important for the development of spinal cord cholinergic neurons (Pfaff et al., 1996; Sharma et al., 1998; Kania et al., 2000). Given that the spinal cord cholinergic neurons are reported to require multiple LIM-homeobox genes for their development, it is expected that Lhx8 is not the only LIM-homeobox gene that is required in generating telencephalic cholinergic neurons. Other candidates are Lhx6 and Isl1, which are also expressed in the basal telencephalon (Marin et al., 2000). Also, it is suggested that Dlx1/2 and...
Mash, though not directly regulating Lhx8, participate in controlling the number of cholinergic neurons that are formed in the telencephalon (Marin et al., 2000). Thus, at a minimum, developmental control genes that may be used in the present invention to transfect cells to bias those cells (or their progeny) to differentiate to a desired end-stage cell type, here that cell type being cholinergic neurons, include, but are not limited to Lhx8, Gbx1, Lhx6, Is1, Dlx1/2 and Mash.

The Human Neural Stem Cells (HNSCs), such as discussed in the examples below, are obtained from cultures that were started from clones obtained from human fetal brain tissue. One lineage was obtained by isolating individual cells from neurospheres of a fetal brain tissue sample obtained from Cambrex, and ultimately identifying one multipotent stem cell for clonal propagation. A second lineage was obtained by isolating a desired multipotent cell from a 9-week old fetal brain (Christopher L. Brannen and Kiminobu Sugaya, Neuroreport 11, 1123-8 (2000)). The HNSCs so obtained were maintained in serum-free medium, and have been demonstrated to have the capability to differentiate into neurons and glial cells such as astrocytes and dendrocytes.

The following examples are provided to further disclose the genesis, operation, scope and uses of embodiments of the present invention. These examples are meant to be instructive, and illustrative, and not to be limiting as to the scope of invention as claimed herein. These examples are to be considered with the referred to drawings.

**Example 1**

This example demonstrates that transfection of a human neural stem cell with Hath1 results in the transfected cell (or its progeny) differentiating into a cell having markers of an inner ear hair cell (IEHC). Hath1 (in humans) and Math1 (in mice) are basic helix-loop-helix transcription factors (and homologs of the Drosophila gene atonal) that are expressed in inner ear sensory epithelia. Since embryonic Math1-null mice failed to generate cochlear and vestibular hair cells, it appears to be required for the generation of inner ear hair cells (Bermingham N A, Hassan B A, Price S D, Vollrath M A, Ben-Arie N, Eato CK, Bellen H J, Lysakowski A, Zoghbi H Y. 1999. Math1: An essential gene for the generation of inner ear hair cells. Science 284 (June 11): 1837-1841). Fate determination of mammalian IEHC is generally completed by birth. However, overexpression of Math1 in postnatal rat cochlear explant cultures resulted in production of extra hair cells from columnar epithelial cells located outside the sensory epithelium, which normally give rise to inner sulcus cells. Math1 expression also facilitated conversion of postnatal utricular supporting cells into hair cells (Zheng, G L, Gao Wq. 2000). Overexpression of Hath1 induces robust production of extra hair cells in postnatal rat inner ears. Nat Neuroscience June; 3(6):580-6. In vivo, Math1 overexpression leads to the appearance of immature hair cells in the organ of Corti and new hair cells adjacent to the organ of Corti in the interdental cell, inner sulcus, and Hensen cell regions, indicating nonsensory cells in the mature cochlea retain the competence to generate new hair cells after over expression of Math1 (Kawamoto K, Ishimoto S, Minoda R, Brough D E, Raphael Y. 2003. Math1 gene transfer generates new cochlear hair cells in mature guinea pigs in vivo. J Neurosci June 1; 23(11):4395-400). Based on the above-summarized work, it was hypothesized that Hath1 may be necessary, and sufficient as a single introduced gene for expression in a multipotent neural stem cell, to positively affect differentiation to an IEHC, or to a cell having characteristics of an IEHC.

A Hath1 gene (SEQ ID NO:4) was amplified from the homo sapiens BAC clone RP11-680117 by PCR and then cloned into a mammalian expression directional cloning vector, pcDNAHismax TOPO TA (See FIG. 1; 6×His tag disclosed as SEQ ID NO: 16). Upon the insertion of the Hath1 gene expressible sequence into the directional cloning vector, the expressible sequence was operatively linked to the CMV promoter, and was also positioned upstream (with regard to reading) of a polyadenylation transcription termination site. The clone was confirmed by sequencing of the insert.

An established a non-serum HNSC culture system was utilized to investigate the differentiation of human neural stem cells (HNSCs) within a defined condition. (Christopher L. Brannen and Kiminobu Sugaya, Regeneration and Transplantation, 11:5, 1123-1128 (2000)). The serum-free supplemented growth medium consisted of HAMS-F12 (Gibco, BRL, Burlington, ON), antibiotic/antimycotic mixture (1:100, Gibco), B27 (1:50, Gibco), human recombinant FGF-2 and EGF (20 ng/ml each, R and D Systems, Minneapolis, Minn.), and heparin (5 µg/ml, Sigma, St. Louis, Mo.). Cells were maintained in 20 ml of this medium at 37° C. in a 5% CO2, humidified incubation chamber.

The mammalian expression vector containing Hath1 gene was transfected into HNSCs by using the Neuroporter Kit (Gene Therapy Systems, Inc. San Diego, Calif.) and Hath1 gene expression was confirmed by RT-PCR. These Hath1-transfected HNSCs were differentiated for 7 days by the depletion of mitotic factors (FGF-2, EGF) from the culture media. After differentiation the cells were fixed for immunocytochemistry and Electron Microscopy.

The immunocytochemistry revealed the existence of cells expressing calretinin, a hair cell marker, which were immunoreactive in this culture. These calretinin immunopositive cells resembled morphology of IEHC. The calretinin expression in the culture was also confirmed by Western blot, which showed single band specific to calretinin molecular weight (29 kD). Further electron microscopy analysis of the cells also showed a typical IEHC morphology. These results indicate that Hath1-transfected HNSCs were differentiated into a vector expressing a Hath1 gene differentiate into IEHCs or into cells having characteristics of IEHCs. Comparisons with non-transfected controls using Western blot and room temperature PCR showed the presence of Hath1 protein and Hath1 mRNA in cells transfected with Hath1, but not in the controls.

Thus, embodiments of the present method provide for improved approaches to obtain IEHCs, or cell having characteristics of IEHCs, that are derived from HNSCs. Embodiments of the present invention provide a higher percentage of a population of cells biased, or disposed, to differentiate to IEHCs, or to cells having characteristics of IEHCs. The HNSCs utilized in this example are readily and continuously cultured in serum-free culture medium. Without being limited, in vitro and in vivo studies and trials using cells so obtained from HNSCs may include electrophysiological assessment of the cells and investigation of functional recovery after transplantation of the cells into the animal model of deafness. Positive findings in such pre-clinical studies may advance the art farther toward treatment of deafness via cell transplantation therapy using IEHCs produced from HNSCs.

**Material and Methods**

**Hath1 Transfection**

The human Hath1 gene (SEQ ID NO:1) is amplified from the Homo sapiens BAC clone RP11-680117 by PCR, using a forward primer (5'-TCCGATCCCTGAGCGTCCGAGCCTT-3', SEQ ID NO: 14) and reverse primer (5'-GCTTCGTCCTACCTCTCCTAACTGGCC-3', SEQ ID NO: 15). The PCR amplification is conducted in 20 µl volumes containing the BAC
cloned into a directional pcDNA-Hismax TOPO TA vector and the clone is confirmed by sequencing of the insert.

The gene expression of Hathl is assessed by RT-PCR with the following condition: 95°C (30s), 50°C (30s), 72°C (60s) for 35 cycles, with an initial denaturation of 95°C (5') and final elongation of 72°C (15'). The PCR amplified fragment is cloned into Human Neural Stem Cells (HNSCs) using the Neuroporter Kit. The Neuroporter Kit utilizes a lipid-based transfection system for the use with cultured primary neurons, neuronal cell lines, and glial cells. DNA and Neuroporter are used in a ratio of 10 µg DNA/75 µl Neuroporter, utilizing 37.5 µl per well in a 6-well plate and with total volumes of 1.5 mL growth media per well. 10 µg of DNA is added to DNA Diluent to make a total volume of 125 µl; this is incubated for 5' at room temperature. 75 µl of the Neuroporter Reagent is added to serum-free media to make a final volume of 125 µl. These solutions are incubated for 10 minutes to allow Neuroporter/DNA complexes to form, and then added directly to the HNSCs in a 6-well plate. One day later, the media is replaced with fresh growth media; one day later, this is replaced with differentiation media (Basal Medium Eagle) to induce spontaneous differentiation. The cells are cultured for 1-2 weeks in a basal differentiation medium containing Eagle's salts and L-glutamine, which is not supplemented with FGF-2 or EGF, and is serum-free.

RT-PCR

TRizol reagent is used to extract RNA for RT-PCR and protein for a Western Blot. 6 µl of the template RNA is added to 1 x Reaction Mix, 1 µM of each Hath1-specific primer, and 1 µl of the RT-Platinum® Taq Mix. The total volume of the solution is 20 µl. The RT-PCR condition is 94°C (15'), 59°C (30s), 72°C (60s) for 40 cycles, with an initial denaturation of 95°C (30s) and 94°C (5').

Immunocytochemistry

The cells are fixed with 4% paraformaldehyde for 30' at room temperature, washed in phosphate-buffered saline (PBS, pH 7.2), then blocked with 3% normal goat serum in PBS containing 0.05% Triton-X100 for 1 hour. The cells are incubated with primary antibody calretinin overnight at 4°C. Following PBST washing, the cells are incubated with secondary antibody biotinylated anti-rabbit made in goat in PBS containing 0.05% Triton-X100 (PBST), with a dilution factor of 1:200. This incubation takes 1 hour. The cells are then washed with PBST, and incubated with ABC reagent for 1 hour. Following a PBS wash and staining with DAB for 5-8', the cells are washed with PBS and distilled water, then stained with methyl green (5'). The cells are washed with water, ethanol, and xylene, coverslipped with permount, and ready for viewing with microscopy.

Western Blot

A Western Blot is performed to assay protein expression. The protein is extracted with TRizol reagent. 15 µl of the protein is loaded with the size marker on a PVDF membrane and run at 200 V and 110 mA/gel for 50. The transfer is run overnight at 15V, 170 mA at 4°C. The membrane is then washed with PBST 2x10' while rotating, and blocked with 3% milk for 60'. This is washed 2x10' with PBST and blocked with the primary antibody calretinin (1:500) overnight at 4°C. After washing 3x5' with PBST the membrane is incubated with the secondary antibody (1:2000) and shaken for 1 hour.

For detection, 7.5 mL of ECL solution is warmed to room temperature and 187.5 µl of solution B is added to solution A. 7.5 mL is added to the membrane at RT for 5'. The membrane is then placed in an x-ray film cassette and exposed as needed for chemiluminescent detection.

Electron Microscopy

Cells were fixed with 3% glutaraldehyde with cacodylate buffer 0.1M, and dehydrated with a series of alcohols beginning with 50% up to 100% absolute ethanol followed by hexamethyldisilazene (HMDS). The cultured cells were allowed to air-dry at room temperature. The specimens were attached to aluminum stubs using double sided carbon coated tape, sputter-coated with Platinum and palladium using the Cressington 208 HR High Resolution Coater. Samples were viewed with a Jeol 6320F Field Emission Microscope (high resolution images) and recorded with a digital camera. Samples were also viewed with the Hitachi Variable pressure microscope in V-P mode (variable pressure mode) and digital images were captured.

Results

A non-serum HNSC culture system was utilized (Christopher L. Brannen and Kiminobu Sugaya, Regeneration and Transplantation, 11:5, 1123-1128 (2000)). This culture system provides for the differentiation and expansion of HNSCs in vitro in the absence of serum. This system provides for the observation of differentiation of HNSCs within a defined condition. These HNSCs have been cultured in a medium consisting of DMEM/F12, antibiotic-antimycotic mixture (1:100), B-27 supplement (1:50), human recombinant FGF-2 and FGF (20 ng/ml each), and heparin (5 µg/ml). These cells have been maintained at 37°C in a 5% CO2 humidified incubation chamber for more than 3 years in the lab. These cells are CD133—a stem cell marker, which is known to be expressed in stem cells—positive, and GFAP- and βIII-tubulin-negative before differentiation. Upon differentiation, various differentiated cells typically express glial fibrillary acidic protein (GFAP), or βIII-tubulin, which are glial and neuronal markers, respectively.

Preferential differentiation of HNSCs into IEHCs can be induced in vitro by the transfection of Hath1. The human Hath1 gene was amplified from the Homo sapiens BAC clone RP11-6801J7 by PCR and cloned into a directional pcDNA-Hismax TOPO TA vector. This was confirmed by sequencing of the insert. After confirming expression of the gene by RT-PCR, the Neuroporter kit was utilized to transfect HNSCs. These HNSCs were known to be viable and capable of differentiation, aggregating in neurospheres when multipotent. Once they began the process of differentiation, they left their neurospheres. After allowing 7 days for differentiation, these cells were either stained for hair cell specific markers or assayed for protein expression. Via immunocytochemistry, the hair cell marker calretinin was identified on certain cells (FIG. 2). Via RT-PCR, the expression of this protein was also verified (FIG. 3).

The presence of the actual protein calretinin on the cell surface was determined via Western Blot. Seven days for differentiation was allowed before any analysis of the cells. Protein was isolated from the cells and calretinin was identified in the cell isolate (FIG. 4).

Using Transmission Electron Microscopy, cells transfected with Hath1 and grown to allow for differentiation were visualized. A subset of the cells exhibited distinct hair-like projections. These were the actual hairs from the transfected HNSCs that differentiated into cells having this characteristic feature of IEHCs (FIGS. 5 and 6).
Discussion

In order to replace damaged IEHCs, a renewable source must be created. The HNSCs cultured in serum-free medium were shown to have the ability to become transfected by Hath1 and then differentiate in vitro into IEHCs, or cells having characteristics of IEHCs. In the present example, transfection with and expression of Hath1 appears to be an essential step in the genesis from HNSCs to IEHCs, or cells having characteristics of IEHCs.

Before transfection, HNSCs do not express Hath1. Following transfection with the Neuroporter Kit, they express this gene in their DNA as verified by RT-PCR. They also produce the hair cell specific marker calretinin as verified by immunocytochemistry and Western Blot. Furthermore, actual hairs from the transfected cells can be visualized through electron microscopy. Thus, characteristics of IEHCs are shown by these data, and it appears that these cells either are end-stage IEHCs or are presumptive IEHC cells in that they have at least two characteristics of IEHCs.

Cells expressing IEHC markers and differentiating into cells with hairlike extremities have been generated in this example. These methods, and the cells produced by the methods of the present invention, as shown in this example, advance the art of differentiating multipotent stem cells toward obtaining end-stage neuron-type cells.

Example 2

Introduction

A cholinergic deficit is one of the primary features of Alzheimer’s disease (AD), where there is a marked degeneration of long-projecting axons of cholinergic neurons in the basal forebrain and target areas in the hippocampus and cerebral cortex. Recent progress in stem cell technologies suggests the probability of using neuroreplacement strategies in AD therapy, although several hurdles are implicated: i) is it possible to generate large number of cholinergic neurons from stem cells; and ii) can long-projecting cholinergic neurons be replaced? Toward improving the ability to conduct research in the area of cell implantation and replacement therapies, and toward achieving desired results in later-developed therapies, embodiments of the present invention are directed to bias human neural stem cells (HNSCs) to differentiate to cells having characteristics of cholinergic neurons through genetic manipulation of endogenous neural precursors in situ.

The LIM-homeobox gene Lhx8 has been reported to be crucial for the proper development of basal forebrain cholinergic neurons in mouse (Zhao et al., 2003; Mori et al., 2004). Lhx8 is expressed in progenitor and postmitotic cells, suggesting that it may have an important role in specification of neural precursor cells and maintenance of phenotype in differentiating and mature neurons. Furthermore, previous studies using the human neuroblastoma cell line, LA-N-2, have demonstrated that treatment with retinoic acid (RA) further enhances cholinergic characteristics of these cells, thus providing a good in vitro model of cholinergic neurons (Crosland, 1996).

The present example utilizes an in vitro assay cell coculture model with plated RA-differentiated LA-N-2 cells and membrane inserts containing Lhx8-transfected HNSCs, to assess whether the Lhx8-transfected HNSCs adopt a cholinergic neuronal fate. The rationale behind this co-culture model is that HNSCs are influenced by intrinsic as well as extracellular factors in the microenvironment and therefore, able to respond by differentiating into specific cell types according to the environmental cues to which they are exposed. Culture of RA-differentiated LA-N-2 in basal media under a serum-free condition, results in the release of factors to the Lhx8-transfected HNSCs in co-culture. It should be noted that there is no cell-to-cell contact in this co-culture system. Thus it is reasonable to assume that any modification of the cell fate of the genetically modified HNSCs by the cholinergic-differentiated LA-N-2 cells would come from membrane permeable endogenous factor(s) released from the cholinergic-differentiated LA-N-2 cells.

Materials & Method

HNSCs culture: Human NSCs were originally purchased from BioWhittaker, Walkersville, Md. These cells have been expanded and passaged in a serum-free culture medium containing bFGF and EGF in our laboratory for over three years (Braunen & Sugaya, 2000). The HNSCs were cultured at a density of 50 spheres in 75 cm² culture flasks (Corning, Cambridge, Mass.) in 20 ml of a serum-free supplemented growth medium consisting of HAMS-F12 (Gibco, BRL, Burlington, ON), antibiotic-antimyotic mixture (1:100, Gibco), bFGF (1.50 µg/ml), human recombinant EGF (20 ng/ml each, R&D Systems, Minneapolis, Minn.) and heparin (5 µg/ml, Sigma, St. Louis, Mo.) incubated at 37°C in a 5% CO² humidified incubation chamber (Fisher, Pittsburgh, Pa.). To facilitate optimal growth conditions, HNSCs were sectioned into quarters every 2 weeks and fed by replacing 50% of the medium every 4-5 days.

LA-N-2 Human Neuroblastoma Culture:

LA-N-2 cells were obtained from Dr. Jan Blusztajn (Boston University, MA). The cells were cultured in Leibovitz L-15 medium (Gibco, BRL, Burlington, ON) containing 10% fetal calf serum and antibiotic-antimyotic mixture (Gibco) in a humidified incubator at 37°C without CO2. The medium was replaced every 3 days. For treatment with retinoic acid (RA), the cells were sub-plated at a density of 0.5-1X10⁶ cells/plate using 0.25% trypsin/1 mM EDTA (Gibco, BRL) and allowed to attach overnight. A fresh stock of 4 mM all-trans retinoic acid RA (Sigma, St. Louis, Mo.) was prepared in 100% ethanol under amber lighting. RA solution was diluted into culture media (final concentration, 10⁻⁶M) and we replaced the media in the cells with the RA-containing media. The media was changed every 48 h during the differentiation of the cells, which was complete after 7-14 days.

Lhx8 Subcloning:

The mouse cDNA clone for Lhx8 (SEQ ID NO: 7, a kind gift from Dr Westphal, NIH, Bethesda, Md.) was inserted into the EcoRI site of the pcDNA 3.1/Zeo mammalian expression vector (Invitrogen). Insertion was subsequently confirmed by restriction digestion and sequence analysis. This mouse Lhx8 (SEQ ID NO: 7) has high homology to the human sequence (70-80%).

Transfection:

HNSCs were placed in 6-well polystyrene coated plates and transfected with 4 µg pcDNA 3.1/Lhx8 plasmid using the Neuroporter transfection system (Gene Therapy Systems, see description in Example 1). Upon the insertion of the Lhx8 gene expressible sequence into the directional cloning vector, the expressible sequence was operatively linked to the CMV description in Example 1). Upon the insertion of the Lhx8 gene expressible sequence into the directional cloning vector, the expressible sequence was operatively linked to the CMV promoter, and was also positioned upstream (with regard to reading) of a polyadenylation transcription termination site. Lhx8 expression was confirmed after 48 hrs by RT-PCR using primers designed from the gene cDNA sequence, 5'TGCTG-GCAGTGCCGCTCTGTC'3 (SEQ ID NO: 12, upper primer) and 5'CGCTCTTGAGTATGACG'3 (SEQ ID NO: 13, lower primer). To initiate differentiation, HNSCs were placed in serum-free basal medium, and allowed to differentiate for 10-15 days in culture.
Co-Cultures of Transfected HNSCs and RA-Treated LA-N-2 Cells:

HNSCs (~5×10^4) transfected with pcDNA 3.1/Lhx8 and non-transfected HNSCs (controls) were transferred into cell culture inserts with an appropriate pore size and suspended in basal media (in the absence of FGF-2 and EGF and without the addition of other extrinsic differentiation factors) over differentiated LA-N-2 cells plated in 6-well plates. For immunochemical analyses of HNSCs, the culture insert was removed after 10-20 days of co-culture and the HNSCs were fixed with 4% paraformaldehyde overnight at 4°C. Also, transfected HNSCs were cultured without the presence of differentiated LA-N-2 cells to assess the need for and effectiveness of the co-culturing.

Immunocytochemistry:

Following fixation, HNSCs were briefly washed 3×5 min in Phosphate buffered saline (PBS), then blocked with 3% normal donkey serum in PBS containing 0.05% Tween 20 (PBS-T) and incubated with goat IgG polyclonal anti-human ChAT (1:500, Chemicon), mouse IgG2b monoclonal anti-human βIII-tubulin (1:1000, Sigma) or rabbit anti-human glial filament protein (GFAP) (1:1000, Sigma) overnight at 4°C. The corresponding secondary antibodies (donkey anti-goat, donkey anti-mouse, and donkey anti-rabbit, respectively) conjugated to rhodamine or FITC (Jackson IR Laboratories, Inc.) were added for a 2 hr incubation at RT in the dark. Cells were then washed with PBS (3×5 min) and mounted with Vectashield with DAPI (Vector Laboratories, CA) for fluorescent microscopic observation. LA-N-2 cells were similarly treated to prepare for microscopic observations. Results

LA-N-2 cells treated with RA expressed Lhx8, βIII-tubulin, and ChAT. This is demonstrated in FIGS. 7A-D. FIG. 7A shows LA-N-2 cells stained red indicating the presence of βIII-tubulin. FIG. 7B shows LA-N-2 cells stained green indicating the presence of ChAT. FIG. 7C shows LA-N-2 cells stained green indicating the presence of NGFβ (blue stain indicating counter-staining for nuclei by DAPI). FIG. 7D (insert) shows non-specific staining for ChAT.

In vitro, HNSCs expressing the LIM homeobox gene, Lhx8, differentiated into mainly βIII-tubulin and ChAT-positive cells, in co-culture with LA-N-2 cholinergic cells. For the transfected HNSCs cultured without the presence of differentiated LA-N-2 cells, there was no significant difference from the non-transfected HNSCs with regard to the number of cells differentiating to cells having characteristics of cholinergic cells. This demonstrated the need under these experimental conditions for the differentiated LA-N-2 cells (and the factors released by them).

Non-transfected HNSCs differentiated into mainly βIII-tubulin and GFAP-positive cells in co-culture with LA-N-2 cholinergic cells. With regard to percentage differences between non-transfected cells and transfected cells, in one trial less than two percent of non-transfected cells, and over 40 percent of transfected cells, were observed at the end of the trial to have characteristics of cholinergic neurons.

Conclusions and Comments

Expression of the LIM-homeobox gene Lhx8 triggers HNSCs to adopt a cholinergic neural lineage. Cells having the noted characteristics of cholinergic neurons either are cholinergic neurons or presumptive cholinergic neurons in that they have at least two characteristics of cholinergic neurons.

LA-N-2 cells in co-culture with HNSCs expressing Lhx8, suggest that the microenvironment is also important for the differentiation and survival of cholinergic neurons.

The present invention may provide utility by biasing human neural stem cells through genetically manipulation so that the cells so manipulated may be used in research, including as cells transplantable, such as in experiments, and therapies, including regarding replacing damaged cholinergic neurons.

As to the efficiency of biasing to a desired cell type, and to observing cells having characteristics of a desired end-stage cell type, without being bound to a particular theory, it is believed that the factors that increase the efficiency of biasing by transfection include: 1) inherent properties of the cell to be transfected; 2) inherent efficiency of the selected vector or method of transfection; 3) relative percentage of cells in which the introduced nucleic acid sequence enters the nucleus compared to remains in the cytoplasm; and 4) number of copies of the nucleic acid sequence that are available for expression in the cell. Methods of transfection are well-known in the art, and the use and modification of known approaches to transfection of a cell with a nucleic acid sequence to be expressed therein to improve the percentage of biasing are within the scope of the present invention.

Thus, it is appreciated that in some embodiments of the present invention, a multipotent stem cell is transfected with a desired development control gene, and the expression of the gene during in vitro culture biases the differentiation of that cell to a desired end-stage differentiated cell. In other embodiments, the multipotent stem cell may be transfected in vivo with a developmental control gene whose expression biases transfected cells to differentiate into a desired end-stage cell. In any of such embodiments, accessory cells may provide factors that are needed for, or that assist with, the differentiation of the transfected cell. These accessory cells, such as the co-cultured LA-N-2 cells in the above example, need not be in contact with the transfected cells, demonstrating here that the factors are membrane permeable. These factors may include the same factor that is expressed by the transfected gene, or may be other factors known in the art or later determined to be useful in achieving a desired differentiation.

Also, it is appreciated that multipotent stem cells may be cultured without an accessory cell, and may receive factors by direct addition of factors to the culture medium, or such factors may be released by cells at a site of implantation, or may be added to a site of implantation.

Example 3

Using the same vector formation and transfection methods as in Example 2, the Human Lhx8 gene (SEQ ID NO: 6) is transfected into HNSCs. Transfected HNSCs are mixed into a first treatment that includes LA-N-2 cells that are treated with RA and that express both Lhx8 and ChAT. A co-culture control comprises NIHSCs that are not transfected but that are in the same culture vessel as LA-N-2 that are treated with RA and that express both Lhx8 and ChAT. For the first treatment and the co-culture control, HNSCs cells are placed in cell culture inserts with an appropriate pore size and suspended in basal media (in the absence of FGF-2 and EGF and without the addition of other extrinsic differentiation factors) over differentiated LA-N-2 cells plated in 6-well plates.

Immunocytochemistry follows the same procedure as in Example 2 above.

Results indicate that HNSCs transfected with the Human Lhx8 gene (SEQ ID NO: 6) also are predisposed, or biased, to differentiate into cells that have characteristics of cholinergic neurons. Observable results include cells that are positive for βIII-tubulin and ChAT.

Example 4

An additional development control gene, Gbx1 sequence (SEQ ID NO: 9), is transfected into HNSCs and is evaluated
as to its capacity to bias HNSCs to differentiate to cholinergic cells, or to cells having characteristics of cholinergic cells. The Gbx1 cDNA sequence (SEQ ID NO: 9) is inserted into the enhanced green fluorescent protein (EGFP) vector pEGFP-C1 (BD Biosciences Clontech) at the EcoRI site within the vector’s multiple cloning site, which is 3 of a CMV promoter and the PGFP gene (See FIG. 8; sequence disclosed as SEQ ID NO: 17). Further, in that a question remains as to whether the percentage of biasing is related directly to the percentage of transfection of cells in population of cells exposed to a transfecting vector, the human Lhx8 cDNA (SEQ ID NO:6) independently also is inserted into a second pEGFP-C1 vector. This allows for visualization of both vectors, each bearing an expressible sequence for a different developmental control gene, in cells in respective cell populations into which these vectors are transfected. Culture methods of the HNSCs into which the Gbx1 and the Lhx8 genes are transfected are as described above in Example 2.

This experiment provides an estimate of the ratio of HNSCs that become cholinergic neurons based on percent transfection of the population. Compared to non-transfected control HNSCs, the transfected cells have characteristics of the desired end-stage differentiated cell type, that is, a cholinergic neuron.

This demonstrates that a number of development control genes, particularly transcription factor genes, may be introduced into a HNSC to bias that cell (or its progeny) to differentiate to a cell having the characteristics of a desired end-stage differentiated neural cell type.

Example 5

Cell sorting technology is combined with the above-described embodiments of the present invention, particularly the vectors of Example 4, to improve the yield and selection of desired cells having the bias to differentiate to a desired end-stage cell (or having already so differentiated). For example, not to be limiting, the introduction of genetic marking such as described above, using EFGP, and the use of Fluorescent Activated Cell Sorter (FACS) techniques is utilized to sort and select cells that have been transfected with the desired developmental control gene (which is linked to a marker on the vector). The FACS technology is well known in the art (See, for example, U.S. patent application number 2002/0127715 A1.)

Using FACS, HNSCs that are transfected with a vector bearing both EFGP and either Gbx1 or Lhx8 are sorted and thereby concentrated. This adds to the utility and effectiveness of the biasing by reducing the number and percentage of cells that are not transfected.

The above examples utilize specific sequences of genes incorporated into respective vectors and introduced into HNSCs. However, the present invention is not meant to be limited to the specifics of these examples. First, in addition to Math1, Hath1, Lhx8 and Gbx1, other developmental control genes of interest include Lshx6, Isl1, Dlx1/2 and Mash. Examples of cDNA sequences, and corresponding translated polypeptide and protein sequences, of these and other developmental control genes are readily obtainable from the GenBank online database (See ncbi.nlm.nih.gov/entrez/query.fcgi), and these are hereby incorporated by reference for that purpose.

Also, as to the nucleic acid sequences comprising the genes of interest, specific sequences of which are provided in the above examples and in the above paragraph, it is appreciated that substantial variation may exist in a nucleic acid sequence for a gene, yet a polypeptide or protein may nonetheless be produced in a cell from one of a number of such variant nucleic acid sequences, wherein such polypeptide or protein has a desired effect on the cell comparable to a polypeptide or protein produced from one of the nucleic acid sequences specified in the above examples. That is, variations may exist in a nucleic acid sequence for a gene yet the variations nonetheless function effectively when substituted for a nucleic acid sequence of a specified gene.

Accordingly, embodiments of the present invention also include and/or employ nucleic acid sequences that hybridize under stringent hybridization conditions (as defined herein) to all or a portion of a nucleic acid sequence represented by any of the SEQ ID Nos. 1-13, or their complements, or to sequences for Isl1, Dlx1/2, Mash, or their complements. The hybridizing portion of the hybridizing nucleic acid sequences is typically at least 15 (e.g., 20, 25, 30, or 50) nucleic acids in length. The hybridizing portion of the hybridizing nucleic acid sequence is at least 80%, e.g., at least 95%, or at least 98%, identical to the sequence of a portion or all of a nucleic acid sequence encoding one of genes identified by the noted Sequence ID numbers, or one of their complements. Hybridizing nucleic acids of the type described herein can be used, for example, as a cloning probe, a primer (e.g., a PCR primer), or a diagnostic probe, as well as for a gene transfected into a cell as described in the examples above.

Hybridization of the oligonucleic acid probe to a nucleic acid sample typically is performed under stringent conditions. Nucleic acid duplex or hybrid stability is expressed as the melting temperature or Tm, which is the temperature at which a probe dissociates from a target DNA. This melting temperature is used to define the required stringency conditions. If sequences are to be identified that are related and substantially identical to the probe, rather than identical, then it is useful to first establish the lowest temperature at which only homologous hybridization occurs with a particular concentration of salt (e.g., SSC or SSPE). Then, assuming that 1% mismatching results in a 1°C. decrease in the Tm, the temperature of the final wash in the hybridization reaction is reduced accordingly (for example, if sequences having >95% identity with the probe are sought, the final wash temperature is decreased by 5°C.). In practice, the change in Tm can be between 0.5°C. and 1.5°C. per 1% mismatch. Stringent conditions involve hybridizing at 68°C. in 5xSSC/5xDenhardt’s solution/1.0% SDS, and washing in 0.2xSSC/0.1% SDS at room temperature. Moderately stringent conditions include washing in 2xSSC at 42°C. The parameters of salt concentration and temperature can be varied to achieve the optimal level of identity between the probe and the target nucleic acid. Additional guidance regarding such conditions is readily available in the art, for example, by Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, N.Y.; and Ausubel et al. (eds.), 1995, Current Protocols in Molecular Biology, (John Wiley & Sons, N.Y.) at Unit 2.10.

The above-specified conditions are not meant to be limiting. For example, provided herein are additional identified sequences for Math1 (SEQ ID Nos:2 and 3), and Hath1 (SEQ ID NO:5). Numerous other similar sequences are known and searchable at GenBank. Also, the methods and compositions disclosed and claimed herein for other sequences may be practiced with Gbx1 (SEQ ID NO:9) and sequences similar to it.

Further, the sequences for introduced genes and polypeptides or proteins expressed by them may also be defined in terms of homology to one of the sequences provided in the above examples and discussion. In the context of the present
application, a nucleic acid sequence is “homologous” with the sequence according to the invention if at least 70%, preferably at least 80%, most preferably at least 90% of its base composition and base sequence corresponds to the sequence specified according to the invention. According to the invention, a “homologous protein” is to be understood to comprise proteins which contain an amino acid sequence at least 70% of which, preferably at least 80% of which, most preferably at least 90% of which, corresponds to the amino acid sequence disclosed in (Gish and States, 1993); wherein corresponds is to be understood to mean that the corresponding amino acids are either identical or are mutually homologous amino acids. The expression “homologous amino acids” denotes those which have corresponding properties, particularly with regard to their charge, hydrophobic character, steric properties, etc. Thus, a protein may be from 70% up to less than 100% homologous to any one of the proteins expressed by one of the disclosed introduced genes. Homology, sequence similarity or sequence identity of nucleic acid or amino acid sequences may be determined conventionally by using known software or computer programs such as the BestFit or Gap pairwise comparison programs (GGC Wisconsin Package, Genetics Computer Group, 575 Science Drive, Madison, Wis. 53711). BestFit uses the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2: 482-489 (1981), to find the best segment of identity or similarity between two sequences. Gap performs global alignments: all of one sequence with all of another similar sequence using the method of Needleman and Wunsch, J. Mol. Biol. 48:443-453 (1970). When using a sequence alignment program such as Bestfit, to determine the degree of sequence homology, similarity or identity, the default setting may be used, or an appropriate scoring matrix may be selected to optimize identity, similarity or homology scores. Similarly, when using a program such as Bestfit to determine sequence identity, similarity or homology between two different amino acid sequences, the default settings may be used, or an appropriate scoring matrix, such as blosum45 or blosum80, may be selected to optimize identity, similarity or homology scores.

Alternatively, as used herein, “percent homology” of two amino acid sequences or of two nucleic acids is determined using the algorithm of Karlin and Altschul (Proc. Natl. Acad. Sci. USA 87:2264-2268, 1990), modified as in Karlin and Altschul (Proc. Natl. Acad. Sci. USA 90:5873-5877, 1993). Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul et al. (J. Mol. Biol. 215:403-410, 1990). BLAST nucleic acid searches are performed with the NBLAST program, score=100, wordlength=12, to obtain nucleic acid sequences homologous to a nucleic acid molecule of the invention. BLAST protein searches are performed with the XBLAST program, score=50, wordlength=3, to obtain amino acid sequences homologous to a reference polypeptide. To obtain gapped alignments for comparison purposes, Gapped BLAST is utilized as described in Altschul et al. (Nucleic Acids Res. 25:3389-3402, 1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) are used. See ncbi.nlm.nih.gov.

Further, in addition to the homology, as indicated in certain claims (i.e., for some embodiments), is a requirement that the homologous or hybridizable nucleic acid sequence or polypeptide or protein functions analogously to the specified sequence of which it is homologous or with which it is hybridizable. That is, the homologous or hybridizable variant functions to achieve the same result, i.e., to increase the probability of a transfected cell, or the percentage of a number of cells, that are biased to differentiate to a cell, or cells, respectively, having characteristics of a desired end-stage differentiated cell.

While the transfection into HNSCs in the above examples uses the Neuroporter approach (Gene Therapy Systems, Inc. San Diego, Calif.), it is appreciated that any known or later-developed method of introduction of a nucleic acid sequence may be employed in the methods and systems, and to produce the compositions, of the present invention. For example, and not to be limiting, introduction of a nucleic acid sequence may be effectuated by stable or transient transfection, lipofection by methods other than Neuroporter, calcium phosphate treatment, electroporation, infection with a recombinant viral vector, and the use of vectors comprising a plasmid construct. Generally and collectively, these methods are considered to be included in the term “means to transfect,” in the term “step for transfecting.” Also, the use of the particular promoter and polyadenylation transcription termination site are not meant to be limiting, as many promoter and transcription termination sites are known and used routinely in the art.

As to the use of different means to transfect, and in view of the above discussion of the relative percentage of cells biased to cells having characteristics of a desired end-stage cell type, it is appreciated that types of transfection, cells that are transfected, and other factors, including post transfection conditions, affect the percentage of cells ultimately biased. In view of these factors, and considering the importance of the specific developmental control genes that are introduced to a cell in certain embodiments of the present invention, in some embodiments the percentage of transfected cells biased exceeds 40 percent, in other embodiments the percentage of transfected cells biased exceeds 50 percent, in other embodiments the percentage of transfected cells biased exceeds 65 percent, and in other embodiments the percentage of transfected cells biased exceeds 70 percent. However, it also is appreciated that determination of the percentage of cells that are in fact transfected in a given container of cells may be difficult to assess, the performance of the present invention in certain embodiments may be expressed in an alternative manner. That is, in some embodiments of the present invention in which a number of cells has been exposed to a selected method or means of transfection for the purpose of introducing a desired developmental control gene (such as Lhx8), the percentage of total cells that are biased to a desired end-stage cell type, or to a cell having characteristics of a desired end-stage cell type, is at least 35 percent, in other embodiments such percentage of total cells exceeds 50 percent, and in other embodiments such percentage of total cells exceeds 70 percent.

Further, it is appreciated that embodiments of the present invention are described as follows:

1. A neural stem cell, including a human neural stem cell, comprising an introduced nucleic acid sequence having an expressible developmental control gene, the expression of said gene being effective to increase the probability of differentiation of said cell to a desired neural cell type.

2. A neural stem cell, including a human neural stem cell, comprising an introduced nucleic acid sequence having an expressible developmental control gene, the expression of said gene being effective to increase the probability of differentiation of said cell to a cell having characteristics of a cholinergic neuron.

3. A neural stem cell, including a human neural stem cell, comprising an introduced nucleic acid sequence having an expressible developmental control gene, the expres-
transfecting homeobox genes may be practiced with progenitor cells; alternatively, embodiments of the present invention may be selected from the group consisting of Math1, Hath1, Lhx8, Gbx1, Lhx6, Isl1, Dlx1/2, Mash and Nurrl. The developmental control gene in the above third description of embodiments of the present invention may be selected from the group consisting of Lhx8, Gbx1, Lhx6, Isl1, Dlx1/2, and Mash. The developmental control gene in the above third description of embodiments of the present invention may be selected from the group consisting of Math1 and Hath1. Finally, the developmental control gene in the above fourth description of embodiments of the present invention may be Nurrl, Pitx3 (SEQ ID NO: 13) or other later-identified specific genes.

Also, it is appreciated that the present invention, particularly for the genes Math1, Hath1, Gbx1, Lhx6, Isl1, Dlx1/2, and Mash, may be utilized in progenitor cells, that is, in cells that are considered to fall within the definitions of pluripotent, of multipotent, and of progenitor cells (i.e., more differentiated than multipotent yet capable of limited self-renewal).

Based on the above examples and disclosure, in view of the knowledge and skill in the art, it also is appreciated that embodiments of the present invention also are used for any homeobox gene, so that a homeobox gene is transfected to a stem cell to effect a biasing of the stem cell to differentiate to a desired end-stage cell, or to a cell having characteristics of the end-stage cell. The stem cell may be a pluripotent or a multipotent stem cell; alternatively, invention embodiments transfecting homeobox genes may be practiced with progenitor cells as described herein. Cells so biased by these genes may be used to develop properly.

According to another embodiment, transfection with Nkx2-5 (SEQ ID NO: 12) biases the differentiation toward the development of cardiac cells. See FIG. 17. Red=human specific Troponin I, Green=Human cells. Following transfection, multipotent stem cells were cocultured with rat cardiomyocytes that provide environmental signals to allow the transfected cells to develop properly.

Further, and more generally, embodiments of the present invention may be practiced by transfecting a stem or a progenitor cell with a nucleic acid sequence comprising a developmental control gene, so that the transfecting is effective to bias the cell to differentiate to a desired end-stage cell, or to a cell having characteristics of the end-stage cell.

Also, it is appreciated that the methods of the present invention may be applied to the daughter cells of multipotent cells, which may have begun some stages of differentiation but are still capable of being biased by transfection of appropriate developmental control genes as described herein, but by virtue of initiating differentiation (or being less self-renewing) may be some opinions therefore not be considered to be multipotent cells. For the purposes of this invention, such daughter cells, which may be found in culture with the multipotent stem cells from which they arose, are termed “biasable progeny cells.”

It is appreciated that embodiments of the present invention also may be defined and claimed with regard to the polypeptide or protein sequences expressed as a result of the transfections disclosed and discussed above. For example, not to be limiting, the peptide sequences, disclosed as the translation sequences in the attached Sequence Listing pages, and their expression in a transfected cell, are used to identify and/or characterize a characteristic and/or result of embodiments of the present invention. Translation sequences are obtainable from the respective GenBank database data entries for cDNAs as described herein, and those database entries are incorporated by reference for such information.

While a number of embodiments of the present invention have been shown and described herein in the present context, such embodiments are provided by way of example only, and not of limitation. Numerous variations, changes and substitutions will occur to those of skill in the art without materially departing from the invention herein. For example, the present invention need not be limited to best mode disclosed herein, since other applications can equally benefit from the teachings of the present invention. Also, in the claims, means-plus-function and step-plus-function clauses are intended to cover the structures and acts, respectively, described herein as performing the recited function and not only structural equivalents or act equivalents, but also equivalent structures or equivalent acts, respectively. Accordingly, all such modifications are intended to be included within the scope of this invention as defined in the following claims, in accordance with relevant law as to their interpretation.
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGGGGGCTAT</td>
<td>continued</td>
</tr>
<tr>
<td>CCGGGGGAGG</td>
<td>continued</td>
</tr>
<tr>
<td>ATCGGTGTG</td>
<td>continued</td>
</tr>
<tr>
<td>CTGGGGTTC</td>
<td>continued</td>
</tr>
<tr>
<td>CGGGGAGG</td>
<td>continued</td>
</tr>
<tr>
<td>TGGGGGGC</td>
<td>continued</td>
</tr>
<tr>
<td>CGGGGCTT</td>
<td>continued</td>
</tr>
<tr>
<td>GCGGGCTT</td>
<td>continued</td>
</tr>
</tbody>
</table>

**<210> SEQ ID NO:** 2
**<211> LENGTH:** 2110
**<212> TYPE:** DNA
**<213> ORGANISM:** Mus musculus

**<400> SEQUENCE:**
```
cggggcttat ccccttcgtt gaactgggtt gccagcacct cctctaacac ggcacctccg agccattgca gtgcgatgtc ccgcctgctg catgcagaag agtgggctga ggtaaaagag ttgggggacc accatcgcca tccccagccg caccacgtcc cgccgctgac gccacagcca cctgctaccc tgcaggcgag agaccttccc gtctacccgg cagaactgtc cctcctggat agcaccgacc cacgcgcctg gctgactccc actttgcagg gcctctgcac ggcacgcgcc gcccagtatc tgctgcattc tcccgagctg ggtgcctccg aggccgcggc gccccgggac gaggctgaca gccagggtga gctggtaagg agaagcggct gtggcggcct cagcaagagc cccgggcccg tcaaagtacg ggaacagctg tgcaagctga agggtggggt tgtagtggac gagcttggct gcagccgcca gcgagcccct tccagcaaac aggtgaatgg ggtacagaag caaaggaggc tggcagcaaa cgcaagggaa cggcgcagga tgcacgggct gaaccacgcc ttcgaccagc tgcgcaacgt tatcccgtcc ttcaacaacg acaagaagct gtccaaatat gagaccctac agatggccca gatctacatc aacgctctgt cggagttgct gcagactcccaatgtcggag agcaaccgcc gccgcccaca gcttcctgca aaaatgacca ccatcacctt cgcaccgcct cctcctatga aggaggtgcg ggcgcctctg cggtagctgg ggctcagcca gccccgggag ggggcccgag acctaccccg cccgggcctt gccggactcg cttctcaggc ccagcttcct ctgggggtta ctcggtgcag ctggacgctt tgcacttccc agccttcgag gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctgt cgccctggct gcccggggcg acctggtgtg cgatctccga gtgagagggg gagggtcaga ggaggaagga aaaaaaatca gaccttgcag aagagactag gaaggttttt gttgttgttg ttcggggctt atccccttcg gacagggccc taacagogat gatgccacag aaggaacctg}
ttgaactggg ttgccagcac ctcctctaac acggcacctc cgagccattg cagtgcgatg
tcccgcctgc tgcatgcaga agagtgggct gaggtaaaag agttggggga ccaccatcgc
catccccagc cgcaccacgt cccgccgctg acgccacagc cacctgctac cctgcaggcg
agagaccttc ccgtctaccc ggcagaactg tccctcctgg atagcaccga cccacgcgcc
tggctgactc ccactttgca gggcctctgc acggcacgcg ccgcccagta tctgctgcat
tctccgcagc tgggtgtcct cgaggccgcg gcgccccggg acgaggctga cagccagggt
gagctggtaa ggagaagcgg ctgtggcggc ctcagcaaga gccccgggcc cgtcaaagta
cgggaacagc tgtgcaagct gaagggtggg gtgtgatggtg acgagcttgg ctgcagcgcg
cagcaagccc ttcgccagca acaagtgataat gggctacaga agcaaaagag gctggcagca
aaagcagggg aacgcgccag gatgcaaggg tgtgacccag ctcctgcacc ctcctgaccc

gltatcctcg cttcctcagaa acaagacgtaat gggctacaga agcaaaagag gctggcagca
caagatcctcg cttcctcagaa acaagacgtaat gggctacaga agcaaaagag gctggcagca

cagcaagccc ttcgccagca acaagtgataat gggctacaga agcaaaagag gctggcagca
aatcctcagc tgtgatcagc gtccttgcctt gccttggcctt gccttgctgc

tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
tctctcagaat cccctgactt tgtgctgctg tgtgctgctg tgtgctgctg tgtgctgctg
<table>
<thead>
<tr>
<th>Start Position</th>
<th>End Position</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
<td>120</td>
<td>ggaagggaaa aaaaatcaga cctgcaaga gagaactagga aggtttttgt tgggtgtgtt</td>
</tr>
<tr>
<td>130</td>
<td>130</td>
<td>cgggtctat cccotctggtt gcaagccagc ccttaacaac ggcaacctcg</td>
</tr>
<tr>
<td>140</td>
<td>140</td>
<td>agccccgcca ctgctgagctt cctcaacgcag gactctgtt</td>
</tr>
<tr>
<td>150</td>
<td>150</td>
<td>tggcggacc accatcgcgca ctcgcagcgcac ccagcttgac gcagcaagcg</td>
</tr>
<tr>
<td>160</td>
<td>160</td>
<td>cctgcgaacct tcgccggtttt cagcagctgag ccttcggtt</td>
</tr>
<tr>
<td>170</td>
<td>170</td>
<td>gagctgctgcc gcaggtgttg ccagactgtt acgggttctg gcagctgagt</td>
</tr>
<tr>
<td>180</td>
<td>180</td>
<td>ggcagttcct cggcttgacct gcaggtgttc gcaggtttctg gcagctgagt</td>
</tr>
<tr>
<td>190</td>
<td>190</td>
<td>ccccacgctct cccttcagtttt gccggatgctt gcagctgagt</td>
</tr>
<tr>
<td>200</td>
<td>200</td>
<td>ggctgcgtgt ccaggtgttttt gcaggttttctt ccagctgagt</td>
</tr>
<tr>
<td>210</td>
<td>210</td>
<td>aaaaatcaga gagaactagga aggtttttgt tgggtgtgtt</td>
</tr>
</tbody>
</table>

**SEQ ID NO 4**
**LENGTH: 1572**  
**TYPE: DNA**  
**ORGANISM: Homo sapiens**  
**FEATURE: modified_base**  
**LOCATION: (1497) .. (1497)**

**OTHER INFORMATION:** a, c, g, t, unknown or other
**US 9,127,291 B2**

-continued

<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1504) .. (1504)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1526) .. (1526)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1564) .. (1564)
<223> OTHER INFORMATION: a, c, g, t, unknown or other

<400> SEQUENCE: 4

```
gtcctctgca cacaagaact ttttcggggt tgtaaaaaact ctttgatggt ctgctcgcac 60

ggagggagcc ggcgctgcca cgcgggtgcag gagggggttg 120
gaggggggc accaacaggg gcgcctgccc gcgccctcca ttggctgaga agacacgcga ccggcgcgag gagggggttg 180
tttggttgcag cttgtgctgc actctctctgg cgcgtgccagcct ttgctgagg 240
cagcccgctg ccagagagtc gggggggaga ctcagctgac agacacgcga ccggcgcgag gagggggttg 300
cagcccgctg ccagagagtc gggggggaga ctcagctgac gcgtccgagcc tttgcagtg 360
```

<210> SEQ ID NO 5
<211> LENGTH: 1065
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 6

```
atgtccctgcagagctgcatgc agaagagtgg gcgtgcgcta cccagtgcctgga ttcctcgcctgact cttctcgcctgagtc 60
```
cgccagcccc agccgcatca tctcccgcaa ccgccgccgc cgccgcagcc acctgcaact ttgcaggcga gagagcatcc cgtctacccg cctgagctgt ccctcctgga cagcaccgac ccacgcgcct ggctggctcc cactttgcag ggcatctgca cggcacgcgc cgcccagtat ttgctacatt ccccggagct gggtgcctca gaggccgctg cgccccggga cgaggtggac ggccgggggg agctggtaag gaggagcagc ggcggtgcca gcagcagcaa gagccccggg ccggtgaaag tgcgggaaca gctgtgcaag ctgaaaggcg gggtggtggt agacgagctg ggctgcagcc gccaacgggc cccttccagc aaacaggtga atggggtgca gaagcagaga cggctagcag ccaacgccag ggagcggcgc aggatgcatg ggctgaacca cgccttcgac cagctgcgca atgttatccc gtcgttcaac aacgacaaga agctgtccaa atatgagace ctgcagatgg cccaaatcta catcaacgcc ttgtccgagc tgctacaaac gcccagcgga ggggaacagc caccgccgcc tccagcctcc tgcaaaagcg accaccacca ccttcgcacc gcggcctcct atgaaggggg cgcgggcaac gcgaccgcag ctggggctca gcaggcttcc ggagggagcc agcggccgac cccgcccggg agttgccgga ctcgcttctc agccccagct tctgcgggag ggtactcggt gcagctggac gctctgcact tctcgacttt cgaggacagc gccctgacag cgatgatggc gcaaaagaat ttgtctcctt ctctccccgg gagcatcttg cagccagtgc aggaggaaaa cagcaaaact tcgcctcggt cccacagaag cgacggggaa tttttcccccc attcccatta cagtgactcg gatgaggcaa gttag

<210> SEQ ID NO 6
<211> LENGTH: 2393
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 6

agaggcaga gagtaggggc tgggccactt gtgcggagt ggttaaagaac tatactgaat 60
ccatattacg aaaggtcga tttctgaact cacccttaag agggaacaac ccgcaagaat 120
gttaaatgc aggacattgc gtggctggct gttgcttggaa tttgcgtcctg 180
aaacgcctcc cattttttcc ccctagagga ccggagagcc ttacgctgcc tctgccgaac 240
ggcttgaaag aaacacagag tcactattgg tcactcagct gcgcttttcg 300
cattgctcc tcacaaacgc cttgagccag cttggccgact ccgacgccag 360
gaggacggc ccgcgggcgg cgctctgcc attcggagcc cactcctccgc tcccaggggtg 420
tcctcgctgg aggcgcttgg ccagcgttgg aaccccaatca aacccgagc tggcggaca 480
gtggggaga gcacaggtgg tgcggcgcgt ggccgctggc acggctttgg cgacctccac 540
tggccgtaa aaagcaacccttctctaccc ctcgagcttc ctggagctca agttataaa 600
cggcgttgg aaagacttgg ttcacgctgct ggcctgctac tccgagcagc 660
gttgatcag aattcggaga tggctcgtgg gctctctgca gcggagtccg ggaggactac 720
tagctccgc cggcggaggg tgcgcaacc cgcggcggc ggcgggaa 780
gggacgaggg cggagagcg tggctctcct ctgctggcccc ctgtccccgt cgtcctcgcc 840
cggagcgggg cgggacgctg cggctcggcc ctgctgctgc ctgctgctgc ctgctgctgc 900
tctgagctgc ctggcggaac atctctttcca ctggggctgc atctctgctgc ctgctgctgc 960
tctctctgc atgtcttgaa gcacccgcag cggagggcgc ccgctctcctg atatatataa 1020
caaagacctttctgcagacgt ggtctggagct cggagtcgcg cagagcagc cggaggggaa 1080
tgggagacac atccatcctaa ctgagctgct cggagagcc aaggggaact ttcatactta 1140
ggcatgtttt gcctgttttc cctgcaaaag gcaactttcc acaggagagg agtttgcttt
1200
ggtggaagag aagactctct gcagacgtca tttagctgc atgtctggta atttaaaaag
1260
agaaagagaa aatggaagtg taggtaggtg ggaggtgcc ctcctcacaag agcagagtg
1320
taaccatcca aaaccgcaaa aaacagctcg gaccagcttt acacagcagc aagttcagg
1380
tatgcaaca taatgctgcc agggaacaaa cccagatgca cgcacacccc agaagatggc
1440
agaagggaga gcagttgcaag aagtgttgat acaggttggg tttctgaaatt tggagacagc
1500
cccaaaaga cagtcagctc ctaaattccc atcctcacc ccaagttcag cagtcaccct
1560
cctcaggtcg tttcaccaca tgttagaaga aatggcctat ttcgctcag tgcoccaaga
1620
tggaagcttg ttaaagctgt gcctagagtt ttttgaggttc cttggccaaa caaactcttg
1680
acctccagcc tggctaaccc atcggaagtc acaacgtgca ataatcata cccataattt
1740
ttttccgga tagacttggat taaggtatct tttattagta taataacaca
1800
tgtgaagat atatacgtta attttattta tcacaccaaa cacaactctttttccttcattg
1860
gctgcccagg tagttctaca ctcggtggtct gccagcaact ttttaattta ctctcattttt
1920
tgttaaactt atgcttacca gaagttcaac aataaccaaa ctttttttct tgcagtggaaa
1980
tagttaccc ttcggtgtat cttgtaatttt atttgctaat aataaggctt ttctgcgtttaaa
2040
agaagggact gaaagttgtg cattatgattt ccaatgatttttt gtaaatgttttttttca
2100
tagtttctat ctcgcaagtt ttttttattta acctttattttttttttttaatcattc
2160
atatacctt aaggttagtt gttggtaaaa gactedaatttt tttttttttttttccatcattc
2220
tgaagacac gccgccaggag cagccatatg tctgttgattt ttttgctcattt atatacaact
2280
ggtgtgtgcgt gttaccagttt ctttggttta aagtttttatt ttttttttttttttttttttt
2340
atataaccttc ttccttttttt ctgtaaaaa tttttttttttttttttaaa
2393
atcacctggc ctgctttgct tgttttcttg gcaagaggcc gctgtccacg ggagaggagt
900
tgcgccttggt ggaggagaag gtcctctgta gagtgcactt tgactgcatg ctggacaatc
960
tgaagagaga agtggagaac ggtaatggga ttagtgtgga aggagccctt ctcacagagc
1020
aagacgtcaa tcatccaaag ccagccaaaa gagctggcag cagctttaca ggcgtacgc
1080
tccacgaggg aagacgtcaa ttaacagcag ggtgttggtt gcaggttactg taacagcggc
1140
<210> SEQ ID NO 8
<211> LENGTH: 751
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 8
tgttttgtag ctcaacaaac aactctggac tccagccctt gttacctatg caatgaccc
60
agctgccaat aagcttacca ttacctcctt ttccagggat aagaatgtag ggtcttataa
120
acttgcctat aatatgatat aaataaactc tggaagata ttaatgtgaa tttttttttt
180
aacaactcga gcatccctga catctctcog cgccacaggt atgtatctcg aaggggcttg
240
cagacactc tcacatcatt tcctatcag tcacagtaat tttttttttt taaaaatattt
300
aagacactc gaactacaaac tttttttttt actgtcttta gatagtgctt ctcagtttctg
360
gtctattac acaatctctc tggagagac aagcttcttg aagagccagt cgtgcttctg
420
cgacacccag aactatctta cttcttcatc tatactcaac ctaatctcct caagttcttg
480
atgtttataa tattatagc tattttttct tataaatgc tatactctt cgtgttcttg
540
tctgtttaa actaatctac tggagaggct cagctttctc tggagagct cgtgcttctg
600
ggagcaagtg aagcgtaatt tgaggaagat ggaggtacgc cggaggaggc aacccaaace
660
cgccccgccc ccccctccct ccttatgagc gagagagcgc ggcgccggag ccacactgcg
120
cagagcccgc gcccgcggcc caccctcgcc ccgacgcccgg cagagcaggca tgcgtgcttg
180
cgcggcgggc acgcgggggc cgggccagcg cccgagagc ggggccgccc cgggccgggg
240
cggggcgggc ggcgggggac gccgagcccc cgggagcggg cgggagcgac tggccggttc
300
tgggccgcc gcgagagcgg tcgacaggg cggccaggc cggccaggc cggagcgcctc atgtcgcct
360
tccacgaggg aagctttctc gcaggttactg taacagcggc
420
cgagcagccc gcaggttactg taacagcggc
ggagcaagtg aagcgtaatt tgaggaagat ggaggtacgc cggaggaggc aacccaaace
480
<210> SEQ ID NO 9
<211> LENGTH: 1452
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 9
ggagcaagtgg aagcttattt tggagagact ggtgtctgcc gggaggagac aacccaaace
60
cgcggcggcc cccccctccc cttatgagcc ggcgcggagc cgcggcgcgg cccagctgcg
120
cagagcgcgc gcgcgagggg caccctcggcc ccagccccgg cagcaggcga tggctgctgc
180
cgcggcgggc cggcgaggcc ccgcgaggcc ccgcgaggcc ccgcgaggcc ccgcgaggcc
240
cggggcgggc ggcgggggac gccgagcccc cgggagcggg cgggaggcgg cgggagcgac
300
tgggccgcc gcgagagcgg tcgacaggg cggccaggc cggccaggc cggagcgcctc atgtcgcct
360
tccacgaggg aagctttctc gcaggttactg taacagcggc
cgcggccgg gcgagagcgg tcgacaggg cggccaggc cggccaggc cggagcgcctc atgtcgcct
<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>gacgggtgcg cggcgccggc gcggcgccggc gcggcaccat gcagggaagc tgccaggggc cgtgggcagc</td>
</tr>
<tr>
<td>2</td>
<td>gccgctttc gccacccgct tcttcgacta caaagatctt tctgtagcgg ggtttccaga</td>
</tr>
<tr>
<td>3</td>
<td>gcacccgagc accttggcag cggcgcggc gcggcgcagc gcggccagcag ccagccaggc</td>
</tr>
<tr>
<td>4</td>
<td>ggcgcgttga cgccgcggc tggagcggc ggcggtggct ggtggggtcc</td>
</tr>
</tbody>
</table>

**SEQ ID NO 10**
**LENGTH: 1632**
**TYPE: DNA**
**ORGANISM: Homo sapiens**

**SEQUENCE: 10**

```
gcgggtgcg cggcgccggc gcggcgccggc gcggcaccat gcagggaagc tgccaggggc cgtgggcagc
```

```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
agggaaggga cccgcgtggc gcgaccctga ccgatcccac ctcaacagct ccctgactct 1140
cggggggaga aggggctccc aacatgaccc tgagtcccct ggattttgca ttcactcctg 1200
cggagaccta ggaacttttt ctgtcccacg cgcgtttgtt cttgcgcacg ggagagtttg 1260
tggcggcgat tatgcagcgt gcaatgagtg atcctgcagc ctggtgtctt agctgtcccc 1320
caggagtgcg cctccgagag tccatgggca cccccggttg gaactgggac tgagctcggg 1380
cacgcagggc ctgagatctg gccgcccatt ccgcgagcca gggccgggcg cccgggcctt 1440
tgctatctcg cgtgcgcgcc cccacgcacc cacccgtatt tatgttttta cctattgcct 1500
taaaataa gaatcctccc cccattaagag aagtgcgctt gaaaaaaa aaaaaa 1560
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1620
aaaaaaaaaa aa 1632

<210> SEQ ID NO 11
<211> LENGTH: 1406
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 11

gagcgcccga gcggagaggc ggcccgggag caggggggcg gcccccactc cggccgggtg 60
cccggcccct ggcccggccc cgccctctag atcgccgccg cagccgccgc tactgggagt 120
tcgcctgttg caggacgcac tagcctctcc tcacagggagt tggcgcgtgc gacgagggca 180
gagccgctgac gcgcgcctgt gtgcgttgcgc gacgcgatgg ccagctgcct ccagtcaagcc 240
ggcgcgtcg gtaagagcgc gacgacgccg cagccggtcg tggcctgcgc gcggcgttgcc 300
ggcgcgttg ccggccagcc cccccctcgc cagcccgccg gacgccgcag ccagtacaatc 360
acccagccag acctagagcc gccttccgct cggacgctta cccgacatcg 420
agaggccgcc aggggatggc ctgcgtggtgc gcgcgccgc gcgcgccgc gcgcgccgc 480
ttgccggcct ggcggccgct gccccctgct gcggccgggg gcggggtggtc gcgtggttgt 540
aaaggggctg gcggccggcgc gcggccggcc gcggccgggg gcggccgggg gcggccgggg 600
ggtcactcgt ggcgcgtcg tcgggagttg gcgccggcgc gcgcgcgcgc gcgcgcgcgc 660
tttccccgt gcgtcggcct gcggcgcgc gcgcggcgc gcgcgcgcgc gcgcgcgcgc 720
ccacccggtg cccccctcgc ccgccccgtg ccgccccctg gcggccccct gcgcgcgcgc 780
cgcggccccgt ccgccccctg ccgccccctg ccgccccctg gcggccccct gcgcgcgcgc 840
gttccccggt gcggccccct gcgcggcgcgc gcgcggcgcgc gcgcggggtg gcgcggcgcgc 900
gcggccccgt gcgcggggtg gcgcggcgcgc gcgcggggtg gcgcggcgcgc gcgcggggtg 960
aagccgcaac gggcgccgct ccgccccgtg ccgccccctg gcggccccct gcgcgcgcgc 1020
aaccttgggc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 1080
caaccccccg gacgggcccc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 1140
tggtgcggcgg cgggggggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg 1200
agaggccgct gcggccggcgc gcgcggggtg gcgcggggtg gcgcggggtg gcgcggggtg 1260
tcggccgct gcgggagttt cggcggcggc gcgcgcggtc gcgcgcggtc gcgcgcggtc 1320
ccgcgcgcgc gacgagcggg gcggccgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 1380
gacgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc 1406

<210> SEQ ID NO 12
<211> LENGTH: 20

<table>
<thead>
<tr>
<th>Start</th>
<th>End</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>1140</td>
<td>1406</td>
<td>267</td>
</tr>
</tbody>
</table>
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 12

tgcggtcag tccgctgtct

<210> SEQ ID NO 13
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 13

c tgctttgg atgattgacg

<210> SEQ ID NO 14
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 14

t cgcatactgc agcgctccga tctt

<210> SEQ ID NO 15
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 15

gc tctgttca ccttacatcc ttcg

<210> SEQ ID NO 16
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic 6xHis tag

<400> SEQUENCE: 16

His His His His His 1

<210> SEQ ID NO 17
<211> LENGTH: 94
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 17

tacaagtccg gactcagatc tcgagctcaa gcttcgaatt ctgcagtcga cggtaccgcg

ggccgagat cacaaggtca tagataacta atca
Met Ser Arg Leu Leu His Ala Glu Glu Trp Ala Glu Val Lys Glu Leu
1  5   10  15
Gly Asp His Asp Pro Gln Pro His His Val Pro Pro Leu Thr
20  25  30
Pro Gln Pro Pro Ala Thr Leu Gin Ala Arg Asp Leu Pro Val Tyr Pro
35  40  45
Ala Gin Leu Ser Leu Leu Asp Ser Thr Asp Pro Arg Ala Trp Leu Thr
50  55  60
Pro Thr Leu Gin Gly Leu Cys Thr Ala Gin Ala Gin Tyr Leu Leu
65  70  75  80
His Ser Pro Gin Leu Leu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
85  90  95
Ala Gin Ser Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
100 105 110
Ser Lys Ser Pro Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
115 120 125
Lys Gly Gly Val Val Asp Gin Leu Leu Cys Ser Gin Gin Gin Gin Gin
130 135 140
Pro Ser Ser Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
145 150 155 160
Ala Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
165 170 175
Asp Gin Leu Arg Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
180 185 190
Ser Lys Tyr Gin Thr Leu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
195 200 205
Ser Gin Leu Leu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
210 215 220
Thr Ala Ser Cys Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
225 230 235 240
Tyr Gin Gly Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
245 250 255
Pro Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
260 265 270
Phe Ser Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
275 280 285
Leu His Phe Pro Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
290 295 300
Gln Lys Gin Leu Ser Pro Leu Gin Gin Gin Gin Gin Gin Gin Gin
305 310 315 320
Gln Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
325 330 335
Glu Phe Ser Pro His Ser Gin Gin Gin Gin Gin Gin Gin Gin Gin
340 345 350
<400> SEQUENCE: 19

Met Ser Arg Leu Leu His Ala Glu Glu Trp Ala Glu Val Lys Glu Leu
1  5  10  15
Gly Asp His Arg His Pro Gln Pro His His Val Pro Pro Leu Thr
20 25  30
Pro Gln Pro Pro Ala Thr Leu Gln Ala Arg Asp Leu Pro Val Tyr Pro
35 40  45
Ala Glu Leu Ser Leu Leu Asp Ser Thr Asp Pro Arg Ala Trp Leu Thr
50 55  60
Pro Thr Leu Gln Gly Leu Cys Thr Ala Arg Ala Ala Gln Tyr Leu Leu
65 70  75  80
His Ser Pro Glu Leu Gly Ala Ser Glu Ala Ala Ala Pro Arg Asp Glu
85 90  95
Ala Asp Ser Gly Glu Leu Val Arg Arg Ser Gly Cys Gly Gly Leu
100 105 110
Ser Lys Ser Pro Gly Pro Val Lys Val Arg Glu Gin Leu Cys Lys Leu
115 120 125
Lys Gly Gly Val Val Asp Leu Gln Gly Cys Ser Arg Gin Arg Ala
130 135 140
Pro Ser Ser Lys Gin Val Asn Gly Val Gin Lys Gin Arg Gin Arg Ala
145 150 155 160
 Ala Asn Ala Arg Glu Arg Arg Arg Ser His Gly Leu Asn His Ala Phe
165 170 175
Aep Gin Leu Arg Asn Val Ile Pro Ser Phe Asn Asp Asp Lys Lys Leu
180 185 190
Ser Lys Tyr Glu Thr Leu Gln Met Ala Gin Ile Tyr Ile Asn Ala Leu
195 200 205
Ser Glu Leu Leu Gin Thr Pro Asn Val Gly Glu Gin Pro Pro Pro Pro
210 215 220
Thr Ala Ser Cys Lys Asn Asp His His His Leu Arg Thr Ala Ser Ser
225 230 235 240
Tyr Glu Gly Gly Ala Gly Ala Ser Ala Val Ala Gly Ala Gin Pro Ala
245 250 255
Pro Gly Gly Pro Arg Pro Thr Pro Pro Gly Cys Arg Thr Arg
260 265 270
Phe Ser Gly Pro Ala Ser Ser Gly Gly Tyr Ser Val Gin Leu Asp Ala
275 280 285
Leu His Phe Pro Ala Phe Glu Asp Arg Ala Leu Thr Ala Met Met Ala
290 295 300
Gln Lys Asp Leu Ser Pro Ser Leu Pro Gly Gly Ile Leu Gin Pro Val
305 310 315 320
Gln Glu Asp Asn Ser Lys Thr Ser Pro Arg Ser His Arg Ser Asp Gly
325 330 335
Glu Phe Ser Pro His Ser His Tyr Ser Asp Ser Asp Glu Ala Ser
340 345 350

<210> SEQ ID NO 20
<211> LENGTH: 354
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 20

Met Ser Arg Leu Leu His Ala Glu Glu Trp Ala Glu Val Lys Glu Leu
1  5  10  15
Gly Asp His Arg Gln Pro Gln Pro His His Leu Pro Gln Pro Pro
Pro Pro Pro Gln Pro Pro Ala Thr Leu Gln Ala Arg Glu His Pro Val
Tyr Pro Pro Glu Leu Ser Leu Leu Asp Ser Thr Asp Pro Arg Ala Trp
Leu Ala Pro Thr Leu Gln Gly Ile Cys Thr Ala Arg Ala Ala Gln Tyr
Leu Leu His Ser Pro Glu Leu Gly Ala Ser Glu Ala Ala Ala Pro Arg
Asp Glu Val Asp Gly Arg Gly Leu Val Arg Arg Ser Ser Gly Gly
Ala Ser Ser Ser Lys Ser Pro Gly Pro Val Lys Val Arg Glu Gln Leu
Cys Lys Leu Lys Gly Gly Val Val Asp Glu Gly Cys Ser Arg
Gln Arg Ala Pro Ser Ser Lys Gln Val Asn Gly Val Val Lys Gln Arg
Arg Leu Ala Ala Asn Ala Arg Arg Arg Met His Gln Leu Asn
His Ala Phe Asp Gln Leu Arg Asn Val Ile Pro Ser Phe Asn Asp
Lys Lys Leu Ser Lys Tyr Glu Thr Leu Gln Met Ala Gln Ile Tyr Ile
Asn Ala Leu Ser Glu Leu Gln Thr Pro Ser Gly Gly Glu Gln Pro
Pro Pro Pro Pro Ala Ser Cys Lys Ser Asp His His His Leu Arg Thr
Ala Ala Ser Tyr Glu Gly Gly Ala Gly Asn Ala Thr Ala Ala Gly Ala
Gln Gln Ala Ser Gly Gly Ser Gln Arg Pro Thr Pro Pro Gly Ser Cys
Arg Thr Arg Phe Ser Ala Pro Ala Ser Ala Gly Tyr Ser Val Gln
Leu Asp Ala Leu His Phe Ser Thr Leu Glu Asp Ser Ala Leu Thr Ala
Met Met Ala Gln Lys Asn Leu Ser Pro Ser Leu Pro Gly Ser Ile Leu
Gln Pro Val Gln Gln Gln Asn Ser Lys Thr Ser Pro Arg Ser His Arg
Ser Asp Gly Glu Phe Ser Pro His Ser His Tyr Ser Asp Ser Asp Gln
Ala Ser

<210> SEQ ID NO 21
<211> LENGTH: 354
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 21
Met Ser Arg Leu Leu His Ala Glu Glu Trp Ala Glu Val Lys Glu Leu
Gly Asp His Arg Gln Pro Gln Pro His His Leu Pro Gln Pro Pro
<table>
<thead>
<tr>
<th>Pro</th>
<th>Pro</th>
<th>Pro</th>
<th>Gln</th>
<th>Pro</th>
<th>Pro</th>
<th>Ala</th>
<th>Thr</th>
<th>Leu</th>
<th>Gln</th>
<th>Ala</th>
<th>Arg</th>
<th>Glu</th>
<th>His</th>
<th>Pro</th>
<th>Val</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35</td>
<td>40</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>Pro</td>
<td>Pro</td>
<td>Glu</td>
<td>Leu</td>
<td>Ser</td>
<td>Leu</td>
<td>Leu</td>
<td>Asp</td>
<td>Ser</td>
<td>Thr</td>
<td>Asp</td>
<td>Pro</td>
<td>Arg</td>
<td>Ala</td>
<td>Trp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>55</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Ala</td>
<td>Pro</td>
<td>Thr</td>
<td>Leu</td>
<td>Gln</td>
<td>Gly</td>
<td>Ile</td>
<td>Cys</td>
<td>Thr</td>
<td>Ala</td>
<td>Arg</td>
<td>Ala</td>
<td>Ala</td>
<td>Gln</td>
<td>Tyr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
<td>His</td>
<td>Ser</td>
<td>Pro</td>
<td>Glu</td>
<td>Leu</td>
<td>Gly</td>
<td>Ala</td>
<td>Ser</td>
<td>Glu</td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>85</td>
<td>90</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Glu</td>
<td>Val</td>
<td>Asp</td>
<td>Gly</td>
<td>Arg</td>
<td>Gly</td>
<td>Leu</td>
<td>Val</td>
<td>Arg</td>
<td>Ser</td>
<td>Ser</td>
<td>Gly</td>
<td>Gly</td>
<td>Gly</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td>105</td>
<td>110</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ser</td>
<td>Ser</td>
<td>Ser</td>
<td>Lys</td>
<td>Ser</td>
<td>Pro</td>
<td>Gly</td>
<td>Pro</td>
<td>Val</td>
<td>Lys</td>
<td>Val</td>
<td>Arg</td>
<td>Glu</td>
<td>Gln</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>110</td>
<td>120</td>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cys</td>
<td>Lys</td>
<td>Leu</td>
<td>Lys</td>
<td>Gly</td>
<td>Val</td>
<td>Val</td>
<td>Asp</td>
<td>Glu</td>
<td>Leu</td>
<td>Gly</td>
<td>Cys</td>
<td>Ser</td>
<td>Arg</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>130</td>
<td>135</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Arg</td>
<td>Ala</td>
<td>Pro</td>
<td>Ser</td>
<td>Ser</td>
<td>Lys</td>
<td>Gln</td>
<td>Val</td>
<td>Asn</td>
<td>Gly</td>
<td>Val</td>
<td>Gln</td>
<td>Lys</td>
<td>Gln</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>145</td>
<td>150</td>
<td>155</td>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Leu</td>
<td>Ala</td>
<td>Ala</td>
<td>Arg</td>
<td>Ala</td>
<td>Arg</td>
<td>Arg</td>
<td>Arg</td>
<td>Met</td>
<td>His</td>
<td>Gly</td>
<td>Leu</td>
<td>Asn</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>165</td>
<td>170</td>
<td>175</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>Ala</td>
<td>Phe</td>
<td>Asp</td>
<td>Glu</td>
<td>Leu</td>
<td>Arg</td>
<td>Asn</td>
<td>Val</td>
<td>Ile</td>
<td>Pro</td>
<td>Ser</td>
<td>Phe</td>
<td>Asn</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>180</td>
<td>185</td>
<td>190</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>Lys</td>
<td>Leu</td>
<td>Ser</td>
<td>Lys</td>
<td>Tyr</td>
<td>Glu</td>
<td>Thr</td>
<td>Leu</td>
<td>Gln</td>
<td>Met</td>
<td>Ala</td>
<td>Gln</td>
<td>Ile</td>
<td>Tyr</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>195</td>
<td>200</td>
<td>205</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Thr</td>
<td>Arg</td>
<td>Ser</td>
<td>Ala</td>
<td>Pro</td>
<td>Ala</td>
<td>Ser</td>
<td>Ala</td>
<td>Gly</td>
<td>Ser</td>
<td>Gln</td>
<td>Thr</td>
<td>Pro</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>210</td>
<td>215</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Pro</td>
<td>Pro</td>
<td>Ala</td>
<td>Ser</td>
<td>Cys</td>
<td>Lys</td>
<td>Ser</td>
<td>Asp</td>
<td>His</td>
<td>His</td>
<td>His</td>
<td>Leu</td>
<td>Arg</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>225</td>
<td>230</td>
<td>235</td>
<td>240</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td>Ser</td>
<td>Tyr</td>
<td>Glu</td>
<td>Gly</td>
<td>Ala</td>
<td>Gly</td>
<td>Arg</td>
<td>Ala</td>
<td>Ala</td>
<td>Thr</td>
<td>Ala</td>
<td>Ala</td>
<td>Gly</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>245</td>
<td></td>
<td>250</td>
<td>255</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Gln</td>
<td>Ala</td>
<td>Ser</td>
<td>Gly</td>
<td>Gly</td>
<td>Ser</td>
<td>Gln</td>
<td>Arg</td>
<td>Pro</td>
<td>Thr</td>
<td>Pro</td>
<td>Pro</td>
<td>Gly</td>
<td>Ser</td>
<td>Cys</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>260</td>
<td></td>
<td>265</td>
<td>270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Thr</td>
<td>Arg</td>
<td>Phe</td>
<td>Ser</td>
<td>Ala</td>
<td>Pro</td>
<td>Ala</td>
<td>Ser</td>
<td>Ala</td>
<td>Gly</td>
<td>Tyr</td>
<td>Ser</td>
<td>Val</td>
<td>Gln</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>275</td>
<td></td>
<td>280</td>
<td>285</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Asp</td>
<td>Ala</td>
<td>Leu</td>
<td>His</td>
<td>Phe</td>
<td>Thr</td>
<td>Phe</td>
<td>Glu</td>
<td>Asp</td>
<td>Ser</td>
<td>Ala</td>
<td>Leu</td>
<td>Thr</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>290</td>
<td></td>
<td>295</td>
<td>300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>Met</td>
<td>Ala</td>
<td>Gln</td>
<td>Lys</td>
<td>Arg</td>
<td>Ser</td>
<td>Leu</td>
<td>Ser</td>
<td>Leu</td>
<td>Pro</td>
<td>Ser</td>
<td>Leu</td>
<td>Gly</td>
<td>Ser</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>305</td>
<td></td>
<td>310</td>
<td>315</td>
<td>320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Pro</td>
<td>Val</td>
<td>Gln</td>
<td>Glu</td>
<td>Arg</td>
<td>Ser</td>
<td>Lys</td>
<td>Thr</td>
<td>Ser</td>
<td>Pro</td>
<td>Arg</td>
<td>Ser</td>
<td>His</td>
<td>Arg</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>325</td>
<td></td>
<td>330</td>
<td>335</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Asp</td>
<td>Gly</td>
<td>Glu</td>
<td>Phe</td>
<td>Ser</td>
<td>Pro</td>
<td>His</td>
<td>Ser</td>
<td>His</td>
<td>Tyr</td>
<td>Ser</td>
<td>Asp</td>
<td>Ser</td>
<td>Asp</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>340</td>
<td></td>
<td>345</td>
<td>350</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ser</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**<210>** SEQ ID NO 22  
**<211>** LENGTH: 356  
**<212>** TYPE: PRT  
**<213>** ORGANISM: Homo sapiens  
**<400>** SEQUENCE: 22  

| Met | Gln | Ile | Leu | Ser | Arg | Cys | Gln | Gly | Leu | Met | Ser | Ser | Gln | Glu | Cys | Gly |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     | 1   | 5   | 10  | 15  |     |     |     |     |     |     |     |     |     |     |     |
| Arg | Thr | Thr | Ala | Ala | Ala | Gly | Arg | Thr | Arg | Lys | Gly | Ala | Gly | Glu |     |
|     | 20  | 25  |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Glu | Gly | Leu | Val | Ser | Pro | Glu | Gly | Ala | Gly | Asp | Glu | Asp | Ser | Cys | Ser |
|     | 35  | 40  |     |     |     |     |     |     |     |     |     |     |     |     |     |

**<210>** SEQ ID NO 22  
**<211>** LENGTH: 356  
**<212>** TYPE: PRT  
**<213>** ORGANISM: Homo sapiens  
**<400>** SEQUENCE: 22  

| Met | Gln | Ile | Leu | Ser | Arg | Cys | Gln | Gly | Leu | Met | Ser | Ser | Gln | Glu | Cys | Gly |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     | 1   | 5   | 10  | 15  |     |     |     |     |     |     |     |     |     |     |     |
| Arg | Thr | Thr | Ala | Ala | Ala | Gly | Arg | Thr | Arg | Lys | Gly | Ala | Gly | Glu |     |
|     | 20  | 25  |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Glu | Gly | Leu | Val | Ser | Pro | Glu | Gly | Ala | Gly | Asp | Glu | Asp | Ser | Cys | Ser |
|     | 35  | 40  |     |     |     |     |     |     |     |     |     |     |     |     |     |
Ser Ser Ala Pro Leu Ser Pro Ser Ser Ser Pro Arg Ser Met Ala Ser
Gly Ser Gly Cys Pro Pro Gly Lys Cys Val Cys Asn Ser Cys Gly Leu
Glu Ile Val Asp Lys Tyr Leu Leu Lys Val Asn Asp Leu Cys Trp His
50
Val Arg Cys Leu Ser Cys Ser Val Cys Arg Thr Ser Leu Gly Arg His
100
Thr Ser Cys Tyr Ile Lys Asp Lys Asp Ile Phe Cys Lys Leu Asp Tyr
150
Phe Arg Arg Tyr Gly Thr Arg Cys Ser Arg Cys Gly Arg His Ile His
200
Ser Thr Asp Trp Val Arg Ala Lys Gly Asn Val Tyr His Leu Ala
250
Cys Phe Ala Cys Phe Ser Cys Lys Arg Glu Leu Ser Thr Gly Glu Glu
300
Phe Ala Leu Val Glu Glu Val Leu Cys Arg Val His Tyr Asp Cys
350
Met Leu Asp Asn Leu Lys Arg Glu Val Glu Asn Gly Asn Gly Ile Ser
400
Val Glu Gly Ala Leu Thr Glu Gln Asp Val Asn His Pro Lys Pro
450
Ala Lys Arg Ala Arg Thr Ser Phe Thr Ala Asp Gin Leu Gin Val Met
500
Gln Ala Gin Phe Ala Gin Asn Asn Pro Asp Ala Gin Thr Leu Gin
550
Lys Leu Ala Glu Arg Thr Gly Leu Ser Arg Arg Val Ile Gin Val Trp
600
Phe Gin Asn Cys Arg Ala Arg His Lys His His Val Ser Pro Asn His
650
Ser Ser Ser Thr Pro Val Thr Ala Val Pro Pro Arg Leu Ser Pro
700
Pro Met Leu Glu Ala Val Tyr Ser Ala Tyr Val Pro Gin Asp Gly
750
Thr Met Leu Thr Ala Leu His Ser Tyr Met Asp Ala His Ser Pro Thr
800
Thr Leu Gly Leu Gin Pro Leu Leu Pro His Ser Met Thr Gin Leu Pro
850
Ile Ser His Thr
900

<210> SEQ ID NO 23
<211> LENGTH: 365
<212> TYPE: PRT
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 23
Met Tyr Trp Lys Ser Asp Gin Met Phe Val Cys Lys Leu Glu Gly Lys
1
Glu Met Pro Glu Leu Ala Val Pro Arg Glu Met Cys Pro Gly Leu Met
20
Ser Glu Cys Gly Arg Pro Ala Ala Gly Ala Gly Arg Thr Arg Lys
35
Gly Ser Gly Glu Glu Leu Val Asn Pro Glu Gly Ala Gly Asp Glu
50

<210> SEQ ID NO 23
<211> LENGTH: 365
<212> TYPE: PRT
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 23
Met Tyr Trp Lys Ser Asp Gin Met Phe Val Cys Lys Leu Glu Gly Lys
1
Glu Met Pro Glu Leu Ala Val Pro Arg Glu Met Cys Pro Gly Leu Met
20
Ser Glu Cys Gly Arg Pro Ala Ala Gly Ala Gly Arg Thr Arg Lys
35
Gly Ser Gly Glu Glu Leu Val Asn Pro Glu Gly Ala Gly Asp Glu
50
Asp Ser Cys Ser Ser Ser Gly Pro Leu Ser Pro Ser Ser Ser Pro Gln
  65   70   75   80
Ser Met Ala Ser Gly Pro Met Cys Pro Pro Gly Lys Cys Val Cys Ser
  85   90
Ser Cys Gly Leu Glu Ile Val Asp Lys Tyr Leu Leu Lys Val Asn Asp
 100 105 110
Leu Cys Trp His Val Arg Cys Leu Ser Cys Ser Val Cys Arg Thr Ser
 115 120 125
Leu Gly Arg His Thr Ser Cys Tyr Ile Lys Asp Lys Asp Ile Phe Cys
 130 135 140
Lys Leu Asp Tyr Phe Arg Arg Tyr Gly Thr Arg Cys Ser Arg Cys Gly
 145 150 155 160
Arg His Ile His Ser Thr Asp Trp Val Arg Arg Ala Lys Gly Asn Val
 165 170 175
Tyr His Leu Ala Cys Phe Ala Cys Phe Ser Cys Lys Arg Gln Leu Ser
 180 185
Thr Gly Leu Glu Phe Ala Val Glu Glu Lys Val Leu Cys Arg Val
 190 195 200 205
His Phe Asp Cys Met Leu Asp Asn Leu Lys Arg Gln Val Glu Asn Gly
 210 215
Asn Gly Ile Ser Val Glu Gly Ala Leu Leu Thr Glu Gln Asp Val Asn
 220 225 230 235 240
His Pro Lys Pro Ala Lys Arg Ala Arg Thr Ser Phe Thr Ala Asp Gln
 245 250 255
Leu Gln Val Met Gln Ala Gln Asp Asn Asn Pro Asp Ala
 260 265 270
Gln Thr Leu Gln Lys Leu Ala Glu Arg Thr Gly Leu Ser Arg Arg Val
 270 275 280 285
Ile Gln Val Trp Phe Gln Asn Cys Arg Ala Arg His Lys Lys His Val
 290 295
Ser Pro Asn His Ser Ser Ser Ala Pro Val Thr Ala Val Pro Ser Ser
 305 310 315 320
Arg Leu Ser Pro Pro Ile Leu Glu Glu Met Ala Tyr Ser Ala Tyr Asp
 325 330 335
Pro Gln Asp Asp Gly Met Leu Thr Ala His Ser Tyr Leu Asp Ala His
 340 345 350
Gln Gln Leu Leu Asp Ser Ser Pro Cys Tyr Pro Ile Gln
 355 360 365

What is claimed is:

1. A method of biasing differentiation of a neural stem cell comprising introducing an expression construct comprising a Hath1 gene sequence into the neural stem cell in vitro, wherein expression of the Hath1 gene sequence is effective to bias the neural stem cell to a desired end-stage cell type, in vitro, wherein the desired end-stage cell type comprises characteristics of an inner-ear hair cell, the characteristics comprising inner ear hair cell projections.

2. A method of biasing differentiation of a neural stem cell in vitro comprising:
   a. providing the neural stem cell;
   b. preparing a nucleic acid sequence comprising a promoter operatively linked to an expressible sequence that comprises a Hath1 gene sequence, the nucleic acid sequence comprising a transcription termination site; and
   c. transfecting said neural stem cell with said nucleic acid sequence, in vitro;
   wherein expression of the expressible sequence results in biasing the neural stem cell to a desired end-stage cell type, in vitro, wherein the desired end-stage cell type comprises characteristics of an inner-ear hair cell, the characteristics comprising inner ear hair cell projections.

3. A method of biasing differentiation of one or more cells in a population of cells comprising neural stem cells, comprising:
   a. providing the population of cells in a vessel;
   b. adding to the vessel a plurality of copies of a nucleic acid sequence comprising Hath1 gene sequence under conditions to express said gene sequence; and
   wherein expression of the gene sequence in one or more cells transfected, in vitro, with a copy of the nucleic acid
sequence is effective to bias the one or more cells to a desired end-stage cell type, wherein the desired end-stage cell type comprises characteristics of an inner-ear hair cell, the characteristics comprising inner ear hair cell projections.

4. The method of claim 3, the population of cells additionally comprising one or more biasable progeny cells.

5. The method of claim 3, the neural stem cells comprising human neural stem cells.