

METHOD OF GENERATING ADIPOCYTES

FIELD OF THE INVENTION

The invention relates to methods of generating adipocytes by overexpressing one or more polypeptides that have the activity of combinations of transcription factors and/or combinations of transcription factors themselves, i.e. through forward reprogramming.

BACKGROUND OF THE INVENTION

Adipocytes, also known as lipocytes and fat cells, are the main component of adipose tissue, which plays an essential role in vertebrate energy metabolism. They have a variety of uses in the fields of medical research and therapy, such as tissue generation and wound healing. There is also a use of adipocytes in cultured meat production where a mixture of myocytes and adipocytes is desirable to yield better tasting meat products.

Currently, methods for differentiating stem cells into adipocytes generally include treating stem cells with differentiation-inducing materials such as insulin, dexamethasone and isobutylmethylxanthine, and culturing them for a long time. However, these differentiation-inducing materials are expensive and the efficiency of cell differentiation is low. By contrast, forward programming strategies provide mature human cell types with unprecedented speed and efficiency. Forward programming involves directly converting pluripotent stem cells, including human pluripotent stem cells (hPSCs), to mature cell types through the forced expression of polypeptides having the activity of key lineage transcription factors and/or the key lineage transcription factors themselves, in order to convert the stem cell into a particular mature cell type.

There is a need in the art to provide methods for generating adipocytes suitable for use as potential therapeutic agents, in research and in tissue engineering.

SUMMARY OF THE INVENTION

According to a first aspect of the invention, there is provided a method of generating adipocytes comprising increasing the expression of one or more transcription factors selected from the group consisting of: a peroxisome proliferator-activated receptor (PPAR) protein, HOXC8, EBF1, EBF2, ZNF467, ZNF423, and variants thereof, in a cell population and culturing the cell population to obtain adipocytes.

35

- According to a further aspect of the invention, there is provided a method of generating adipocytes comprising expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of one or more transcription factors, the transcription factors selected from the group consisting of: one or more PPAR proteins (such as PPARG), HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CCAAT enhancer binding proteins (CEB proteins, such as CEBPA and/or CEBPB) and variants thereof, in a cell population, preferably a pluripotent stem cell population, more preferably a human induced pluripotent stem cell (hiPSC) population, and culturing the cell population to obtain adipocytes.
- 10 In a preferred embodiment, the transcription and translation (expression) of the polypeptides having transcription factor activity and/or the transcription factors is controlled within the cell, preferably through the use of external stimuli.
- 15 According to a further aspect of the invention, there is provided a method for the production of adipocytes cells from a source cell, preferably a pluripotent stem cell, more preferably a hiPSC, comprising the steps of:
- a) insertion (preferably targeted insertion) of a gene encoding a transcriptional regulator protein into a first genomic safe harbour site of the source cell; and
 - 20 b) insertion (preferably targeted insertion) of at least one nucleotide sequence encoding one or more polypeptides having the activity of one or more transcription factors and/or encoding one or more transcription factors, the transcription factors selected from the group consisting of: one or more PPAR proteins (such as PPARG), HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB) and
 - 25 variants thereof, operably linked to an inducible promoter into a second genomic safe harbour site of the source cell, wherein said inducible promoter is regulated by the transcriptional regulator protein; and
 - c) culturing the source cell(s) comprising the insertions to obtain adipocytes.
- 30 According to a further aspect of the invention, there is provided a method for the production of adipocytes cells from a source cell, comprising the steps of:
- a) targeted insertion of a gene encoding a transcriptional regulator protein into a first genomic safe harbour site of the source cell; and
 - b) targeted insertion of at least one nucleotide sequence encoding one or more
 - 35 transcription factors selected from the group consisting of: a PPAR protein, HOXC8, EBF1, EBF2, ZNF467, ZNF423 and variants thereof, operably linked to an inducible promoter into a

second genomic safe harbour site of the source cell, wherein said inducible promoter is regulated by the transcriptional regulator protein; and

c) culturing the source cell(s) comprising the insertions to obtain adipocytes.

According to a further aspect of the invention, there is provided a use of at least two or more
5 transcription factors, wherein the two or more transcription factors are selected from the group consisting of: PPARA, PPARG, HOXC8, EBF1, EBF2, ZNF467, ZNF423 and variants thereof, to generate adipocytes.

According to a further aspect of the invention, there is provided a use of one or more
10 polypeptides having the activity of one or more transcription factors and/or one or more transcription factors, wherein the one or more transcription factors are selected from the group consisting of: PPARA, PPARG, HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB) and variants thereof, to generate adipocytes, preferably human adipocytes.

15 According to a further aspect of the invention, there is provided a cell obtainable by any one of the methods defined herein.

According to a further aspect of the invention, there is provided a cell comprising one or more
20 exogenous expression cassettes comprising nucleotide sequences encoding at least one or more transcription factors, wherein the one or more transcription factors are selected from the group consisting of: a PPAR protein, HOXC8, EBF1, EBF2, ZNF467, ZNF423 and variants thereof.

25 According to a further aspect of the invention, there is provided a cell, preferably a pluripotent stem cell, more preferably a hiPSC, comprising one or more exogenous expression cassettes comprising nucleotide sequences encoding one or more polypeptides having the activity of one or more transcription factors and/or encoding at least one or more transcription factors, wherein the one or more transcription factors are selected from the group consisting of: one or
30 more PPAR proteins (such as PPARA and/or PPARG), HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB) and variants thereof.

According to a further aspect of the invention, there is provided a cell, preferably a human cell, as defined herein, for use in therapy, *in vitro* diagnostics, drug screening or preparing cultured
35 meat.

According to a further aspect of the invention, there is provided a kit for differentiating a cell into an adipocyte comprising:

(i) a source cell and an agent that activates or increases the expression or amount of at least one or more transcription factors; and/or

5 (ii) one or more expression cassette(s) comprising nucleotide sequences encoding at least one or more transcription factors,

wherein the one or more transcription factors are selected from the group consisting of: a PPAR protein, HOXC8, EBF1, EBF2, ZNF467, ZNF423 and variants thereof.

10 According to a further aspect of the invention, there is provided a kit for differentiating a cell, preferably a pluripotent stem cell, more preferably a hiPSC, into an adipocyte comprising:

(i) a source cell and an agent that activates or increases the expression or amount of at least one or more transcription factors; and/or

(ii) one or more expression cassette(s) comprising nucleotide sequences encoding one
15 or more polypeptides having the activity of one or more transcription factors and/or encoding at least one or more transcription factors,

wherein the one or more transcription factors are selected from the group consisting of: one or more PPAR proteins (such as PPARG and/or PPARG), HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB) and variants
20 thereof.

According to a further aspect of the invention, there is provided a use of a kit as defined herein, for differentiating a cell into an adipocyte.

25 According to a further aspect of the invention, there is provided a method of drug screening comprising contacting an adipocyte generated using the method or an adipocyte as defined herein, with the drug and observing a change in the adipocyte induced by the drug.

According to a further aspect of the invention, there is provided a method of treating a subject
30 having or at risk of a disease or disorder comprising administering to the subject a therapeutically effective amount of adipocytes generated using the method or adipocytes as defined herein.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. Single-cell gene expression data visualized on a UMAP plot including undifferentiated iPSCs (G10), endpoint sorted live cells (NG), FABP4-GFP positive cells, PLIN1-mCherry positive cells, marker-negative cells (FN).

5

Figure 2. Distribution of detected transcription factors (eTFs) within the UMAP plot.

Figure 3. Cell identities within the UMAP plot assigned by using CellTypist against a reference dataset of gene expression profiles of adipose tissue cells. Besides adipocytes, 4 cell types are shown (“Cell type 1-4”).

10

Figure 4. Dot plot showing expression levels of adipokines, adipocyte marker genes and cardiac marker genes (as a control) across the FABP4-GFP and PLIN1-mCherry positive cell populations (shown in Figure 1). Dot sizes indicate the percentage of cells in each population expressing the gene, and the gray levels indicate the average expression level.

15

Figure 5. Dot plot showing expression levels of marker genes of brown adipose tissue across the different cell types annotated by CellTypist (shown in Figure 3). Dot sizes indicate the percentage of cells in each population expressing the gene, and the gray levels indicate the average expression level.

20

Figure 6. Dot plot showing expression levels of marker genes of white adipose tissue across the different cell types annotated by CellTypist (shown in Figure 3). Dot sizes indicate the percentage of cells in each population expressing the gene, and the gray levels indicate the average expression level.

25

Figure 7. Log-likelihood enrichment of each eTF combination in the adipocytes population compared to the control (NG) population. Upper row: adipocytes in FABP4-GFP-positive cells in each experimental replicate; Lower row: adipocytes in PLIN1-mCherry-positive cells in each experimental replicate.

30

Figure 8. RT-qPCR of adipocytes markers in iPSC line encoding PPARG, PPARA, HOXC8, ZNF423, EBF1 and EBF2 on day 10 of reprogramming. Expression levels shown in relation to the reference hydroxymethylbilane synthase (HMBS).

35

Figure 9. Immunocytochemistry for the adipocyte markers FABP4 and PLIN1 in iPSC line encoding PPARG, PPARA, HOXC8, ZNF423, EBF1 and EBF2 on day 10 of reprogramming. Lipid accumulation was assessed by LipidTOX staining. Scale bar is 100µm.

5 DETAILED DESCRIPTION

The present invention provides methods for producing adipocytes from source cells, preferably pluripotent stem cells, by expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of a select group of transcription factors which the present inventors have identified as inducing cell differentiation into adipocytes.

Definitions

Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by a person skilled in the art to which this invention belongs. As used herein, the following terms have the meanings ascribed to them below.

References to “transcription factor” as used herein, refer to proteins that are involved in gene regulation in both prokaryotic and eukaryotic organisms. In one embodiment, transcription factors can have a positive effect on gene expression and, thus, may be referred to as an “activator” or a “transcriptional activation factor”. In another embodiment, a transcription factor can negatively affect gene expression and, thus, may be referred to as “repressors” or a “transcription repression factor”. Activators and repressors are generally used terms and their functions may be discerned by those skilled in the art.

The term “increasing the expression of” or “increasing the amount of” with respect to increasing an amount, level or expression of a transcription factor, refers to increasing the quantity of the transcription factor in a cell of interest (e.g., a source cell). In some embodiments, the amount of transcription factor is increased in a cell (e.g., via an expression cassette directing expression of a polynucleotide encoding one or more transcription factors) when the quantity of transcription factor is at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more relative to a control (e.g., a source cell without said expression cassette(s) or a control cell where the baseline expression is zero or negligible). In some of the embodiments, increasing the expression comprises “overexpressing” the transcription factor, *i.e.*, increasing the expression of the transcription factor above the endogenous expression level of the transcription factor in the cell.

Methods of the invention may be used in a "cell population", *i.e.*, a collection of cells which may be differentiated into the desired cell type. Said cell population may comprise "source cells", also referred to as "starting cells", *i.e.*, a cell type prior to differentiation into the desired cell type.

5

References herein to "pluripotent" refer to cells which have the potential to differentiate into all types of cell found in an organism. One form of pluripotent stem cell, known as induced pluripotent stem cells, are of particular interest to the present invention. "Induced pluripotent stem cells" (iPSCs) are cells that have been reprogrammed to an embryonic stem cell-like state by being forced to express genes and factors important for maintaining the defining properties of embryonic stem cells. In 2006, it was shown that overexpression of four specific transcription factors could convert adult cells into pluripotent stem cells. Oct-3/4 and certain members of the Sox gene family have been identified as potentially crucial transcriptional regulators involved in the induction process. Additional genes including certain members of the Klf family, the Myc family, Nanog, and Lin28, may increase the induction efficiency. Examples of the genes which may be used as reprogramming factors to generate iPSCs include Oct3/4, Sox2, Sox1, Sox3, Sox15, Sox17, Klf4, Klf2, c-Myc, N-Myc, L-Myc, Nanog, Lin28, Fbx15, ERas, ECAT15-2, Tcl1, beta-catenin, Lin28b, Sall4, Esrrb, Tbx3 and Glis1, GATA3, GATA6 and these reprogramming factors may be used singly, or in combination of two or more kinds thereof. In particular, the reprogramming factors may comprise at least the Yamanaka factors, *i.e.*, Oct3/4, Sox2, Klf4 and c-Myc. These reprogramming factors may also be used in combination with the transcription factors of interest in the present invention.

References herein to "somatic" refer to any type of cell that makes up the body of an organism, excluding germ cells. Somatic cells therefore include, for example, skin, heart, muscle, bone or blood cells and their stem cells. Somatic cells may also be referred to as differentiated cells. In one embodiment, the somatic cell may be an adult cell or a cell derived from an adult which displays one or more detectable characteristics of an adult or non-embryonic cell.

Methods of the invention (*e.g.*, cellular reprogramming of iPSCs) are for use in generating "adipocytes", which may also be referred to as "lipocytes" or "fat cells". The term "adipocyte" as used herein is meant to refer to cells that are related to the adipose tissue of vertebrates. This term includes both white adipocytes, which are the primary site of triglyceride/energy storage, and brown adipocytes, which play an important role in energy expenditure in the form of thermogenesis. In one embodiment, the adipocyte is a white adipocyte. In alternative embodiments, the adipocyte is a brown adipocyte. The term "adipocyte" includes adipocyte-

35

like cells that exhibit some but not all characteristics of adult adipocytes, as well as mature, fully functional and/or metabolically active adult adipocyte cells. This term also includes adult and fetal adipocyte progenitor cells (also known as preadipocytes) and fetal adipocytes. This term includes further cells with the capacity to engraft fat tissue when transplanted *in vivo*. The
5 adipocytes produced by this method may be at least as functional as the adipocytes produced by directed differentiation to date.

References herein to “culturing” include the addition of cells (*e.g.*, the cell population, *i.e.*, the source cells), to media comprising growth factors and/or essential nutrients. It will be
10 appreciated that such culture conditions may be adapted according to the cells or cell population to be generated according to methods of the invention.

References to a “variant” when referring to a polypeptide could be, for example, an amino acid sequence at least 80%, 85%, 90%, 95%, 98%, or 99% identical to the original, full-length
15 polypeptide. When referring to a nucleic acid sequence, the term “variant” could be, for example, a nucleic acid sequence at least 80%, 85%, 90%, 95%, 98%, or 99% identical to the original, full-length nucleic acid sequence. The variant could be a fragment of full-length polypeptide, in particular a functional fragment of the polypeptide. The fragment may be at least 50%, 60%, 70%, 80%, 85%, 90%, 95%, 98%, or 99% as long as the full-length wild type
20 polypeptide or a domain thereof having an activity of interest such as the ability to differentiate a source cell into an adipocyte. Variations known in the art to eliminate or substantially reduce the activity of the protein are preferably avoided. In some embodiments, the variant lacks an N- and/or C-terminal portion of the full-length polypeptide, *e.g.*, up to 10, 20, or 50 amino acids from either terminus is lacking. In some embodiments, a functional variant or fragment has at
25 least 50%, 60%, 70%, 80%, 90%, 95% or more of the activity of the full-length wild type polypeptide. One of skill in the art will be aware of, or will readily be able to ascertain, whether a particular polypeptide variant or fragment is functional using assays known in the art. For example, the ability of a variant of a transcription factor as listed in **Table 1** to generate adipocytes can be assessed using the assays as described herein. In particular, the variant
30 may be a biologically active variant. A “biologically active variant” includes any variant of a molecule having substantially, at least in part, the same functional and/or biological properties of said molecule, such as binding properties, and/or the same structural features, such as binding domain. It also refers to a molecule that exhibits the functional features as the transcription factors disclosed herein.

35

In one embodiment, the variant is an isoform of the listed transcription factor. Many transcription factors have one or more isoforms which result, for example, from alternative splicing or from a shifted transcription initiation. Based on the different transcript variants (i.e. mRNA), different polypeptides are generated. It is possible that different transcript variants
5 have different translation initiation sites.

A “promoter” is a nucleotide sequence which is recognised by proteins involved in initiating and regulating transcription of a polynucleotide sequence. An “inducible promoter” is a nucleotide sequence where expression of a genetic sequence operably linked to the promoter
10 is controlled by an analyte, co-factor, regulatory protein, etc. It is intended that the term “promoter” or “control element” includes full-length promoter regions and functional (e.g., controls and/or affects transcription or translation) segments of these regions.

The term “operably linked” refers to an arrangement of elements wherein the components so
15 described are configured so as to perform their usual function. Thus, a given promoter operably linked to a genetic sequence is capable of effecting the expression of that sequence when the regulatory factors are present. The promoter need not be contiguous with the sequence, so long as it functions to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between the promoter sequence and the genetic
20 sequence and the promoter sequence can still be considered “operably linked” to the genetic sequence. Thus, the term “operably linked” is intended to encompass any spacing or orientation of the promoter element and the genetic sequence in the inducible cassette which allows for initiation of transcription of the inducible cassette upon recognition of the promoter element by a transcription complex.

25 The term “vector”, as used herein, is intended to refer to a nucleic acid molecule which is used as a vehicle to carry genetic material into a cell. One type of vector is a “plasmid”, which refers to a circular double stranded DNA loop or circle into which additional DNA segments may be ligated. Another type of vector is an infectious but non-pathogenic viral vector, wherein
30 additional DNA segments may be ligated to certain viral genetic elements. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian and yeast vectors). Other vectors (e.g., non-episomal mammalian vectors) can be integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host
35 genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as “recombinant expression

vectors" (or simply, "expression vectors"). In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, lentiviral vectors, adenoviruses, Sendai viruses and adeno-
5 associated viruses), which serve equivalent functions, and also bacteriophage and phagemid systems. Another type of vector includes synthetic and *in vitro* transcribed RNA molecules, e.g., mRNA and stabilised RNA, to carry coding genetic information to the cells. This also includes synthetic-self- replicating RNA vectors.

10 References to "subject", "patient" or "individual" refer to a subject, in particular a mammalian subject, to be treated. Mammalian subjects include humans, non-human primates, farm animals (such as cows), sports animals, or pet animals, such as dogs, cats, guinea pigs, rabbits, rats or mice. In some embodiments, the subject is a human. In alternative embodiments, the subject is a non-human mammal, such as a mouse.

15 The term "sufficient amount" means an amount sufficient to produce a desired effect. The term "therapeutically effective amount" is an amount that is effective to ameliorate a symptom of a disease or disorder. A therapeutically effective amount can be a "prophylactically effective amount" as prophylaxis can be considered therapy.

20 As used herein, the term "about" when used herein includes up to and including 10% greater and up to and including 10% lower than the value specified, suitably up to and including 5% greater and up to and including 5% lower than the value specified, especially the value specified. The term "between" includes the values of the specified boundaries.

25 It will be understood that any method as described herein may have one or more, or all, steps performed *in vitro*, *ex vivo* or *in vivo*.

Transcription factor activity

30 The method described herein may comprise increasing the expression (in particular, the protein expression) of a sufficient number of polypeptides having the activity of the transcription factors (or the transcription factors themselves) (e.g., as listed in **Table 1** and variants and isoforms thereof) capable of causing differentiation of a cell population to adipocytes, therefore differentiating the cell population into adipocytes. In the context of the present invention, these
35 factors may also be referred to as "reprogramming factors". As described herein, the

expression of an exogenous or endogenous (in particular an exogenous) transcription factor may be increased.

According to an aspect of the invention, there is provided a method of generating adipocytes comprising increasing the expression of one or more transcription factors selected from the group consisting of: a peroxisome proliferator-activated receptor (PPAR) protein, HOXC8, EBF1, EBF2, ZNF467, ZNF423, and variants thereof, in a cell population and culturing the cell population to obtain adipocytes.

According to another aspect of the invention, there is provided a method of generating adipocytes comprising expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of one or more transcription factors selected from the group consisting of: one or more PPAR proteins (such as PPARG), HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB) and variants thereof, in a cell population, preferably a pluripotent stem cell population, more preferably a hiPSC population, and culturing the cell population to obtain adipocytes.

References herein to a “peroxisome proliferator-activated receptor” (PPAR) or a “PPAR protein” refer to transcription factor in the group of nuclear receptor proteins known as peroxisome proliferator-activated receptors. There are three types of PPAR in the group: PPAR alpha (PPARA), PPAR beta/delta (PPARD) and PPAR gamma (PPARG). All PPARs heterodimerize with the retinoid X receptor (RXR) and bind to peroxisome proliferator hormone response elements (specific regions on the DNA of target genes). These specific regions have a DNA consensus sequence of 5'-AGGTCANAGGTCA-3' (SEQ ID NO: 1), with N being any nucleotide.

In one embodiment, the PPAR is selected from the group consisting of: PPARG. It will be understood that the one or more transcription factors, as referenced herein, includes more than one type of PPAR protein, such as PPARG in combination with PPARG. In one embodiment, the PPAR is PPARG. In an alternative embodiment, the PPAR is PPARG. It will be understood that if the expression of one or more polypeptide having the activity of one or more transcription factor and/or of more than one transcription factor is increased, this may include one or more PPAR protein. Therefore, in one embodiment, the method comprises increasing the expression of one or more polypeptide having the activity of PPARG and/or of PPARG and PPARG themselves.

References herein to “PPARA” or “PPAR α ” or “Peroxisome Proliferator Activated Receptor Alpha” relates to a member of the PPAR subfamily of nuclear hormone receptors. In one embodiment, the PPARA is human PPARA. Wild type human PPARA is identified by UniProt ID: Q07869, and is encoded by the *PPARA* gene, identified by Ensembl Gene ID: ENSG00000186951.

References herein to “PPARG” or “PPAR γ ” or “Peroxisome Proliferator Activated Receptor Gamma” relates to a member of the PPAR subfamily of nuclear hormone receptors, and includes all isoforms of PPARG, such as PPARG1 and PPARG2. In one embodiment, the PPARG is human PPARG. Wild type human PPARG is identified by UniProt ID: P37231, and is encoded by the *PPARG* gene, identified by Ensembl Gene ID: ENSG00000132170.

References herein to “HOXC8” relates to Homeobox Protein Hox-C8. In one embodiment, the HOXC8 is human HOXC8. Wild type human HOXC8 is identified by UniProt ID: P31273, and is encoded by the *HOXC8* gene, identified by Ensembl Gene ID: ENSG00000037965. HOXC8 belongs to the homeobox family of genes that encode a highly conserved family of transcription factors. HOXC8 is involved in the regulation of cartilage differentiation, HOXC8 is involved with cell junction organization and the regulation of CDH11 expression and function.

References herein to “EBF1” relates to Early B Cell Factor 1 (also known as Transcription factor COE1). In one embodiment, the EBF1 is human EBF1. Wild type human EBF1 is identified by UniProt ID: Q9UH73, and is encoded by the *EBF1* gene, identified by Ensembl Gene ID: ENSG00000164330. EBF1 is involved in the olfactory signalling pathway and in nervous system development. EBF1 activates B-cell-specific genes such as BCR or CD40 and represses genes associated with T cell fates, such as GATA3 and TCF7.

References herein to “EBF2” relates to Early B Cell Factor 2 (also known as Transcription factor COE2). In one embodiment, the EBF2 is human EBF2. Wild type human EBF2 is identified by UniProt ID: Q9HAK2, and is encoded by the *EBF2* gene, identified by Ensembl Gene ID: ENSG00000221818. EBF2 regulates osteoclast differentiation by activating the decoy receptor for RANKL, TNFRSF11B.

References herein to “ZNF467” or “Zinc Finger Protein 467” relates to a zinc finger protein. In one embodiment, the ZNF467 is human ZNF467. Wild type human ZNF467 is identified by

UniProt ID: Q7Z7K2, and is encoded by the *ZNF467* gene, identified by Ensembl Gene ID: ENSG00000181444. *ZNF467* binds to STAT3 at the consensus sequence 5'-CTTCTGGGAAGA-3' (SEQ ID NO: 2).

- 5 References herein to “ZNF423” relates to Zinc Finger Protein 423. In one embodiment, the ZNF423 is human ZNF423. Wild type human ZNF423 is identified by UniProt ID: Q2M1K9, and is encoded by the *ZNF423* gene, identified by Ensembl Gene ID: ENSG00000102935. ZNF423 plays a central role in BMP signaling and olfactory neurogenesis. ZNF423 acts as a transcriptional repressor via its interaction with EBF1, a transcription factor involved in terminal
- 10 olfactory receptor neurons differentiation. ZNF423 is involved in olfactory neurogenesis by participating in a developmental switch that regulates the transition from differentiation to maturation in olfactory receptor neurons.

- References herein to “CEB proteins” relates to CCAAT enhancer binding proteins. In one
- 15 embodiment, the CEB protein is a human CEB protein. The CEB protein may be, for example CEBPA (also known as CEBP α or C/EBP α) and/or CEBPB (also known as CEBP β or C/EBP β). Wild type human CEBPA is identified by UniProt ID: P49715, and is encoded by the *CEBPA* gene, identified by Ensembl Gene ID: ENSG00000245848. Wild type human CEBPB is identified by UniProt ID: P17676, and is encoded by the *CEBPB* gene, identified by
- 20 Ensembl Gene ID: ENSG00000172216. CEB proteins, by their name interact with the CCAAT box motif present in several gene promoters. CEB proteins have a highly conserved basic-leucine zipper domain at the C-terminus. CEB proteins also contain activation domains at the N-terminus and regulatory domains. CEB proteins recruit co-activators, such as CREB-binding proteins, that in turn open up the chromatin structure or recruit basal transcription factors.

25

In one embodiment, the CEB protein is CEBPB. In another embodiment, the CEB protein is not CEBPA.

- In one embodiment, the method comprises expressing one or more polypeptides having the
- 30 activity of two or more transcription factors, in particular three or more, four or more or five transcription factors and/or increasing the expression of two or more transcription factors, in particular three or more, four or more, five or more and six or more, the transcription factors selected from the group consisting of: PPARA, PPARG, HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB); and variants thereof.

35

In one embodiment, the method comprises expressing one or more polypeptides having the activity of one or more transcription factors, in particular two or more, three or more, four or more or five transcription factors and/or increasing the expression of one or more transcription factors, in particular two or more, three or more, four or more, five or more and six or more, the
5 transcription factors selected from the group consisting of: PPARA, PPARG, HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB); and variants thereof.

Once the activity of a transcription factor is appreciated, the endogenous transcription
10 machinery can be modulated using not only the transcription factors themselves, but also polypeptides engineered to replicate the action of the transcription factor, such as synthetic transcription factors or artificial transcription factors. For example, CRISPR (clustered regularly interspaced palindromic repeats), TALE (transcriptional activator-like effector) or Zinc Finger technologies can be used to modulate the expression of endogenous cellular genes, to allow
15 for faster and more efficient nuclear reprogramming under conditions amenable for clinical and commercial applications. This is set out in, for example, US2016/362705, incorporated herein.

Alternatively, with the development of highly accurate protein structure prediction with artificial intelligence tools such as AlphaFold, it is now straightforward for polypeptides to be developed
20 that have very similar structure and/or activity to a transcription factor of interest whilst at the same time having an amino acid sequence that has very little resemblance to that of the transcription factor of interest. For example, large language models trained on biological diversity have been used to develop proteins only around 70% identical to CRISPR-Cas proteins that occur in nature and yet with comparable or improved biological activity and
25 specificity (Ruffolo et al. (2024) bioRxiv, doi: <https://doi.org/10.1101/2024.04.22.590591>). Such polypeptides are covered within the scope of the invention.

In some embodiments of the present invention, a polypeptide (in particular a single polypeptide) is engineered to mimic the activity of more than one transcription factor of interest.
30 In a further embodiment, polypeptides having the activity of one or more transcription factors is expressed, in combination with increasing the expression of another transcription factor. For example, a polypeptide having the activity of PPARA can be expressed in combination with increasing the expression of HOXC8 (or *vice versa*).

35 According to another aspect of the invention, there is provided a method of generating adipocytes comprising increasing the expression of three or more transcription factors, wherein

the one or more transcription factors are selected from the group consisting of: PPARA, PPARG, HOXC8, EBF1, EBF2, ZNF467, ZNF423; and variants thereof, in a non-adipocyte cell population and culturing the cell population to obtain adipocytes.

- 5 According to another aspect of the invention, there is provided a method of generating adipocytes comprising expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of the one or more transcription factors themselves, wherein the one or more transcription factors are selected from the group consisting of: PPARA, PPARG, HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB
10 proteins (such as CEBPA and/or CEBPB); and variants thereof, in a non-adipocyte cell population and culturing the cell population to obtain adipocytes.

In one embodiment, the (e.g. at least one or more) transcription factors are selected from the group consisting of: PPARA, HOXC8, EBF1, EBF2 and variants thereof.

15

In one embodiment, the (e.g. at least two or more) transcription factors are selected from the group consisting of: PPARA, HOXC8, EBF1, EBF2 and variants thereof.

- In a further embodiment, the method comprises expressing one or more polypeptides having
20 the activity of PPARA and HOXC8 and/or increasing the expression of PPARA and HOXC8.

In a further embodiment, the method comprises expressing one or more polypeptides having the activity of one or more transcription factors, and/or increasing the expression of the transcription factors themselves, where the transcription factors are PPARA, EBF1 and EBF2.

- 25 According to a further aspect of the invention, there is provided a method of generating adipocytes comprising expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of the transcription factors themselves, where the transcription factors are PPARA, EBF1 and EBF2, in a cell population and culturing the cell population to obtain adipocytes.

30

In a further embodiment, the method comprises expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of the transcription factors themselves, where the transcription factors are HOXC8, EBF1 and EBF2.

- 35 According to a further aspect of the invention, there is provided a method of generating adipocytes comprising expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of the transcription factors themselves,

where the transcription factors are HOXC8, EBF1 and EBF2, in a cell population and culturing the cell population to obtain adipocytes.

In one embodiment, the (e.g. at least two or more) transcription factors are selected from the
5 group consisting of: PPARA, PPARG, HOXC8, ZNF467, ZNF423 and variants thereof.

In a further embodiment, the method comprises expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of the transcription factors themselves, where the transcription factors are PPARA, PPARG and
10 ZNF467. According to a further aspect of the invention, there is provided a method of generating adipocytes comprising expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of the transcription factors themselves, where the transcription factors are PPARA, PPARG and ZNF467, in a cell population and culturing the cell population to obtain adipocytes.

15 In a further embodiment, the method comprises expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of the transcription factors themselves, where the transcription factors are PPARA, PPARG and HOXC8. According to a further aspect of the invention, there is provided a method of
20 generating adipocytes comprising expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of the transcription factors themselves, where the transcription factors are PPARA, PPARG and HOXC8, in a cell population and culturing the cell population to obtain adipocytes.

25 In a further embodiment, the method comprises expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of the transcription factors themselves, where the transcription factors are PPARA, PPARG and ZNF423. According to a further aspect of the invention, there is provided a method of
30 generating adipocytes comprising expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of the transcription factors themselves, where the transcription factors are PPARA, PPARG and ZNF423 in a cell population and culturing the cell population to obtain adipocytes.

In one embodiment, the method comprises expressing one or more polypeptides having the
35 activity of one or more additional transcription factors and/or increasing the expression of the

one or more additional transcription factors themselves. The additional transcription factors may be one or more of the transcription factors listed in **Table 1**.

In one embodiment, the transcription factor comprises PPARA. PPARA may be used in
5 combination with one or more, such as one, two, three, four, or five transcription factors
selected from the list in **Table 1**. In a further embodiment, the method comprises expressing
one or more polypeptides having the activity of between two and seven transcription factors
and/or increasing the expression of the between two and seven transcription factors
themselves, the transcription factors selected from PPARA in combination with PPARG,
10 HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or
CEBPB) or variants thereof. In an even further embodiment, the method comprises expressing
one or more polypeptides having the activity of between two and five transcription factors
and/or increasing the expression of the between two and five transcription factors themselves,
the transcription factors selected from PPARA in combination with PPARG, EBF1, EBF2,
15 ZNF467, one or more CEB proteins (such as CEBPA and/or CEBPB) or variants thereof.

In one embodiment, the transcription factor comprises PPARG. PPARG may be used in
combination with one or more, such as one, two, three, four, or five transcription factors
selected from the list in **Table 1**. In a further embodiment, the method comprises expressing
20 one or more polypeptides having the activity of between two and seven transcription factors
and/or increasing the expression of the between two and seven transcription factors
themselves, the transcription factors selected from PPARG in combination with PPARA,
HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or
CEBPB) or variants thereof. In an even further embodiment, the method comprises expressing
25 one or more polypeptides having the activity of between two and five transcription factors
and/or increasing the expression of the between two and five transcription factors themselves,
the transcription factors selected from PPARG in combination with PPARA, HOXC8, EBF1,
EBF2, ZNF467, one or more CEB proteins (such as CEBPA and/or CEBPB) or variants
thereof.

30 In one embodiment, the transcription factor comprises ZNF467. ZNF467 may be used in
combination with one or more, such as one, two, three, four, or five transcription factors
selected from the list in **Table 1**. In a further embodiment, the method comprises expressing
one or more polypeptides having the activity of between two and seven transcription factors
and/or increasing the expression of the between two and seven transcription factors
35 themselves, the transcription factors selected from ZNF467 in combination with PPARA,

PPARG, HOXC8, EBF1, EBF2, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB) or variants thereof. In an even further embodiment, the method comprises expressing one or more polypeptides having the activity of between two and three transcription factors and/or increasing the expression of the between two and three transcription factors themselves, the transcription factors selected from ZNF467 in combination with PPARA, PPARG, or variants thereof.

In one embodiment, the transcription factor comprises ZNF423. ZNF423 may be used in combination with one or more, such as one, two, three, four, or five transcription factors selected from the list in **Table 1**. In a further embodiment, the method comprises expressing one or more polypeptides having the activity of between two and seven transcription factors and/or increasing the expression of the between two and seven transcription factors themselves, the transcription factors selected from ZNF423 in combination with PPARA, PPARG, HOXC8, EBF1, EBF2, ZNF467, one or more CEB proteins (such as CEBPA and/or CEBPB) or variants thereof. In an even further embodiment, the method comprises expressing one or more polypeptides having the activity of between two and three transcription factors and/or increasing the expression of the between two and three transcription factors themselves, the transcription factors selected from ZNF423 in combination with PPARA, HOXC8, or variants thereof.

In one embodiment, the transcription factor comprises EBF1. EBF1 may be used in combination with one or more, such as one, two, three, four, or five transcription factors selected from the list in Table 1. In a further embodiment, the method comprises expressing one or more polypeptides having the activity of between two and seven transcription factors and/or increasing the expression of the between two and seven transcription factors themselves, the transcription factors selected from EBF1 in combination with PPARA, PPARG, HOXC8, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB) or variants thereof. In an even further embodiment, the method comprises expressing one or more polypeptides having the activity of between two and four transcription factors and/or increasing the expression of the between two and four transcription factors themselves, the transcription factors selected from EBF1 in combination with PPARA, HOXC8, EBF2 or variants thereof.

In one embodiment, the transcription factor comprises EBF2. EBF2 may be used in combination with one or more, such as one, two, three, four, or five transcription factors selected from the list in Table 1. In a further embodiment, the method comprises expressing

one or more polypeptides having the activity of between two and seven transcription factors and/or increasing the expression of the between two and seven transcription factors themselves, the transcription factors selected from EBF2 in combination with PPARA, PPARG, HOXC8, EBF1, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB) or variants thereof. In an even further embodiment, the method comprises expressing one or more polypeptides having the activity of between two and four transcription factors and/or increasing the expression of the between two and four transcription factors themselves, the transcription factors selected from EBF2 in combination with PPARA, HOXC8, EBF1 or variants thereof.

Methods of the invention encompass the use of variants of the transcription factors of interest (i.e., as described in **Table 1**). References to the transcription factors also encompasses species variants, isoforms, homologues, allelic forms, mutant forms, and equivalents thereof, including conservative substitutions, additions, deletions therein not adversely affecting the structure and/or function. Changes in the nucleic acid sequence of the transcription factor gene can result in conservative changes or substitutions in the amino acid sequence. Therefore, the invention includes polypeptides having conservative changes or substitutions. The invention includes sequences where conservative substitutions are made that do not alter the activity of the transcription factor protein of interest.

Table 1. Transcription factors for generation of adipocytes, including accession numbers (as accessed on 27 July 2023)

Transcription Factor Protein name	Transcription Factor Gene name	Ensembl Gene ID
AT-rich interaction domain 5B (ARID5B)	ARID5B	ENSG00000150347
Cyclin dependent kinase inhibitor 1A	CDKN1A	ENSG00000124762
CCAAT enhancer binding protein beta	CEBPB	ENSG00000172216
cAMP responsive element binding protein 1	CREB1	ENSG00000118260
Early B cell Factor 1	EBF1	ENSG00000164330
Early B cell Factor 2	EBF2	ENSG00000221818
estrogen related receptor alpha	ESRRA	ENSG00000173153
forkhead box A3	FOXA3	ENSG00000170608
Forkhead box O1	FOXO1	ENSG00000150907

Homeobox protein C10	HOXC10	ENSG00000180818
Homeobox protein C8	HOXC8	ENSG00000037965
Krüppel-like factor 15	KLF15	ENSG00000163884
Nuclear factor 1 A-type	NFIA	ENSG00000162599
Peroxisome proliferator-activated receptor alpha	PPARA	ENSG00000186951
Peroxisome proliferator-activated receptor gamma	PPARG	ENSG00000132170
Retinoid X receptor alpha	RXRA	ENSG00000186350
Zinc finger and BTB domain-containing protein 7B	ZBTB7B	ENSG00000160685
Zinc finger protein 423	ZNF423	ENSG00000102935
Zinc finger protein 467	ZNF467	ENSG00000181444
Zinc finger protein 516	ZNF516	ENSG00000101493

Cell Types

The method may be used on any cell type, including stem cells. In the case of stem cells, the generation of adipocytes using the method may be referred to as “cellular reprogramming”, “forward reprogramming”, “direct programming” or “direct differentiation”, i.e., the pluripotent stem cell is differentiated into an adipocyte. Furthermore, adipocyte cellular reprogramming may be used as generic terminology referring to the use of transcription factors to differentiate a source cell into adipocytes.

10

Sources of cells suitable for methods of the invention may include, for example, any stem cells or non-adipocyte cells. For example, the stem cells may be pluripotent stem cells, for example induced pluripotent stem cells, embryonic stem cells or pluripotent stem cells derived by nuclear transfer or cell fusion. It may be preferred that the embryonic stem cell is derived without destruction of the embryo, particularly where the cells are human. In some embodiments, the stem cells are not derived from human or animal embryos, i.e., the invention does not extend to any methods which involve the destruction of human or animal embryos. The stem cells may also include multipotent stem cells, oligopotent stem cells, or unipotent stem cells. The stem cells may also include fetal stem cells or adult stem cells, such as hematopoietic stem cells, mesenchymal stem cells, neural stem cells, epithelial stem cells, skin stem cells. In certain aspects, the stem cells may be isolated from umbilical, placenta,

20

amniotic fluid, chorion villi, blastocysts, bone marrow, adipose tissue, brain, peripheral blood, cord blood, menstrual blood, blood vessels, skeletal muscle, skin and liver.

5 In one embodiment, the cell population is of human origin. The source cell e.g., a non-adipocyte cell, may be of human origin. It is well known that, compared with non-human pluripotent cells, genome engineering in human pluripotent stem cells is challenging due to, for example, partially due to low transfection/transduction efficiency and high apoptosis under stresses such as low-density plating, drug-selection and sorting (Cerbini *et al.*, *PLOS ONE*, 10(1), e0116032).

10

In one embodiment, the cell population is of animal origin. The source cell e.g., a non-adipocyte cell, may be of animal origin. In certain aspects, the cell is preferably one from a livestock animal. Livestock animals include, for example, pigs, cows, horses, buffalo, bison, goats, sheep, deer, reindeer, donkeys, bantengs, yaks, chickens, ducks and turkeys.

15

In one embodiment, the cell population comprises stem cells, e.g., induced pluripotent stem cells (iPSCs), embryonic stem cells (ESCs), haematopoietic stem cells, mesenchymal stem cells or neuronal stem cells. In a further embodiment, the cell population comprises pluripotent stem cells, e.g., iPSCs or ESCs.

20

In one embodiment, the source cell is a stem cell, e.g., an iPSC, an ESC, a haematopoietic stem cell, a mesenchymal stem cell or a neuronal stem cell. In a further embodiment, the source cell is a pluripotent stem cell, e.g., an iPSC or an ESC. In some embodiments, the source cell is an iPSC.

25

Methods of preparing induced pluripotent stem cells are also known in the art. Induction of iPSCs typically require the expression of or exposure to at least one member from Sox family and at least one member from Oct family. Sox and Oct are thought to be central to the transcriptional regulatory hierarchy that specifies ES cell identity. For example, Sox may be Sox-1, Sox-2, Sox-3, Sox-15, or Sox-18; Oct may be Oct-4. Additional factors may increase the reprogramming efficiency, like Nanog, Lin28, Klf4, or c-Myc; specific sets of reprogramming factors may be a set comprising Sox-2, Oct-4, Nanog and, optionally, Lin-28; or comprising Sox-2, Oct4, Klf and, optionally, c-Myc. In one method, iPSC may be generated by transfecting cells with transcription factors Oct4, Sox2, c-Myc and Klf4 using viral transduction. In an alternative method, iPSC may be generated by transfecting cells with RNA encoding

30
35

transcription factors inducing the development of stem cell characteristics, such as transcription factors selected from Oct4, Sox2, c-Myc and Klf4.

In one embodiment, the adipocytes are human adipocytes.

5

In one embodiment, the induced pluripotent stem cells are derived from somatic or germ cells of the patient. Such use of autologous cells would remove the need for matching cells to a recipient. Alternatively, commercially available iPSC may be used, such as those available from WICELL (WiCell Research Institute, Inc, Wisconsin, US). Alternatively, the cells may be a tissue-specific stem cell which may also be autologous or donated.

10

Delivery of transcription factors

It will be understood that methods for expressing polypeptides having transcription factor activity and/or increasing the expression of the transcription factors in the cells to be programmed into adipocytes may include any method known in the art, for example, by induction of expression of one or more expression cassettes previously introduced into the cells, or by introduction of nucleic acids (such as DNA or RNA), polypeptides, or small molecules to the cells to stimulate expression of the endogenous or exogenous transcription factors. Increasing the expression of certain endogenous but transcriptionally repressed genes may also reverse the silencing or inhibitory effect on the expression of these genes by regulating the upstream transcription factor expression or epigenetic modulation. Therefore, methods of the invention may involve culturing the cell population under conditions to artificially increase the expression level of one or more of the transcription factors described herein.

20

In one embodiment, the expression of the polypeptides having transcription factor activity and/or the transcription factors themselves is increased by contacting the cell population with the polypeptides and/or the transcription factors (i.e., the proteins encoding the transcription factors). Delivery of the transcription factors may occur using direct electroporation of transcription factor proteins to the cells.

25

In an alternative embodiment, the expression of the transcription factors is increased by introducing a promoter (e.g. a strong promoter) ahead of an endogenous gene encoding the transcription factor(s).

30

35

In a further alternative embodiment, the expression of the polypeptides having transcription factor activity and/or the transcription factors themselves is increased by contacting the cell population with one or more agents that activate or increase the expression amount of the (exogenous or endogenous) transcription factors. In the case of polypeptides having transcription factor activity or exogenous transcription factors, the agents may still be used after the genes for the polypeptides and/or transcription factors have been inserted into the cell.

In one embodiment, the agent is selected from the group consisting of: a nucleic acid (i.e., polynucleotide, e.g., messenger RNA (mRNA), coding DNA sequence), a protein, an aptamer and small molecule, ribosome, RNAi agent, guide RNA (gRNA) and peptide nucleic acid (PNA) and analogues or variants thereof. In one embodiment, the agent is a transcriptional activation system (e.g., a gRNA for use in a gene activation system such as CRISPR/Cas or TALEN) for increasing the expression of the one or more endogenous transcription factors.

The method of inducing differentiation of the cell population (i.e., source cells), may comprise delivering to the cells a nucleic acid comprising an open reading frame encoding one or more of the polypeptides having transcription factor activity, one or more of the transcription factors themselves (e.g., in an expression cassette), the transcription factor protein, and/or an activator of transcription of the open reading frame encoding the polypeptide and/or transcription factor. This results in the amount of the transcription factor in the cells being increased, and the cells differentiate to form adipocytes. Said open reading frame may be part of a recombinant expression cassette.

In one embodiment, the nucleic acid comprises a recombinant or exogenous expression cassette comprising the one or more transcription factor sequences (or genes) in a sufficient number to cause cellular reprogramming of source cells to adipocytes. The exogenous expression cassette may comprise an externally inducible transcriptional regulatory element for inducible expression of the one or more transcription factors, such as an inducible promoter, e.g., comprising a tetracycline response element or variant thereof.

If expression of the transcription factors is increased by introducing an exogenous sequence encoding the transcription factor (e.g., the transcription factor gene), then it would be understood that any suitable system for delivering the sequence may be used. The gene delivery system may be a transposon system; a viral gene delivery system; an episomal gene delivery system; or a homologous recombination system such as utilizing a zinc finger

nuclease, a transcription activator-like effector nuclease (TALENs), a meganuclease, or CRISPR/Cas, or the like.

Alternatively, introduction of a nucleic acid, such as DNA or RNA, into cells may use any
5 suitable methods for nucleic acid delivery for transformation of a cell, as described herein or
as would be known to one of ordinary skill in the art. Such methods include, but are not limited
to, direct delivery of DNA such as by *ex vivo* transfection, by injection (including microinjection),
by electroporation, by calcium phosphate precipitation, by using DEAE-dextran followed by
polyethylene glycol, by direct sonic loading, by liposome mediated transfection, by receptor-
10 mediated transfection, by microprojectile bombardment, by agitation with silicon carbide fibers,
by Agrobacterium-mediated transformation, and any combination of such methods. Through
the application of these techniques, cells may be stably or transiently transformed.

Further, the expression cassette (e.g., an inducible recombinant expression cassette) may
15 include cleavable sequences. Such sequences are sequences that are recognised by an entity
capable of specifically cutting DNA, and include restriction sites, which are the target
sequences for restriction enzymes or sequences for recognition by other DNA cleaving entities,
such as nucleases, recombinases, ribozymes or artificial constructs. At least one cleavable
sequence may be included, but preferably two or more are present. These cleavable
20 sequences may be at any suitable point in the cassette, such that a selected portion of the
cassette, or the entire cassette, can be selectively removed if desired. The cleavable sites may
thus flank the part/all of the genetic sequence that it may be desired to remove. The method
may therefore also comprise removal of the expression cassette and/or the genetic material.

25 In an alternative embodiment, the cell population is contacted with one or more agents that
has the same effect as activating or increasing the expression or amount of the transcription
factors (i.e. an indirect method of increasing the expression transcription factor). In this aspect
of the invention, the method comprises introducing an exogenous agent which mimics the
effect of increasing the expression of the transcription factors described herein. For example,
30 such a method may comprise introducing a protein (e.g. an engineered zinc finger nuclease)
that has a DNA-binding activity analogous to the transcription factor. For instance, PPAR
proteins bind to peroxisome proliferator responsive elements, so the activity of these
transcription factors could be reproduced by a zinc finger nuclease engineered to bind the
same domain.

35

It will be understood that a combination of one or more of the methods for expressing the polypeptides having transcription factor activity or increasing the expression of the transcription factors may be used where the combination overall results in activity necessary for the forward programming to adipocytes.

5

Vectors

In one embodiment, the polypeptides having transcription factor activity or the transcription factors themselves (e.g., combinations of polypeptides and/or transcription factors) are introduced into the cell population using a vector. One of skill in the art would be well equipped to construct a vector through standard recombinant techniques. Vectors include but are not limited to plasmids, cosmids, viruses (bacteriophage, animal viruses, and plant viruses), and artificial chromosomes (e.g., YACs).

In one embodiment, a nucleic acid sequence encoding one or more transcription factors is introduced into a cell by a transposon system (i.e. involving a transposon plasmid). A transposon delivery system is comprised of two plasmids, one encoding the transposase and one encoding the transcription factor(s). The transposase protein mediates random integration of the transcripts encoded in the transposon plasmid into the genome. In one embodiment, the transposon system is selected from a PiggyBac or Sleeping Beauty transposon system. The transposon plasmid encodes a payload flanked by two ITRs (internal terminal repeats). The payload may comprise a Tet inducible promoter, the transcription factor(s), and optionally a selection marker, e.g. an antibiotic selection cassette under a constitutive promoter.

In one embodiment, the transposase and transposon plasmids are delivered by nucleofection or lipofection into the cells. The number of integration events, and therefore the number of copies of payload per cell can be in part controlled by adjusting the total and relative amounts of transposase and transposon plasmid DNA. This allows the combinatorial delivery of transcription factors at a single cell level.

In one embodiment, the vector is a viral vector. The viral gene delivery system may be an RNA-based or DNA-based viral vector. Viral vectors include retroviral vectors, lentiviral vectors (e.g., derived from HIV-1, HIV-2, SIV, BIV, FIV etc.), gammaretroviral vectors, adenoviral (Ad) vectors (including replication competent, replication deficient and gutless forms thereof), adeno-associated virus-derived (AAV) vectors, simian virus 40 (SV-40) vectors, bovine papilloma virus vectors, Epstein-Barr virus vectors, herpes virus vectors, vaccinia virus vectors, Harvey murine sarcoma virus vectors, murine mammary tumour virus vectors, Rous

sarcoma virus vectors and Sendai virus vectors. In a further embodiment, the viral vector is selected from: a lentiviral vector, an adeno-associated virus vector or a Sendai virus vector. In a yet further embodiment, the viral vector is a lentiviral vector.

- 5 Lentiviral vectors are well known in the art. Lentiviral vectors are complex retroviruses capable of integrating randomly into the host cell genome, which, in addition to the common retroviral genes *gag*, *pol*, and *env*, contain other genes with regulatory or structural function (e.g., accessory genes *Vif*, *Nef*, *Vpu*, *Vpr*). Lentiviral vectors have the advantage of being able to infect non-dividing cells and can be used for both *in vivo* and *ex vivo* gene transfer and
10 expression of nucleic acid sequences. For example, recombinant lentiviral vector capable of infecting a non-dividing cell wherein a suitable host cell is transfected with two or more vectors carrying the packaging functions, namely *gag*, *pol* and *env*, as well as *rev* and *tat*.

- 15 In one embodiment, the vector is a self-replicating RNA vector expression system. For example, the system may comprise self-replicating RNA vectors that remain ectopic to the host cell genome and encode the transcription factors that induce reprogramming. Self-replicating RNA vectors are known in the art and many are based on positive strand RNA viruses, such as alphaviruses.

- 20 In one embodiment, the viral vector is used at a high multiplicity of infection (MOI). A high MOI helps to ensure that more than one transcription factor is introduced into the source cell. In one embodiment, the MOI is greater than 0.5, such as 1.0 or above.

- 25 In one embodiment, a nucleic acid sequence encoding the one or more polypeptides having transcription factor activity and/or transcription factors is introduced into a cell by a plasmid. In one embodiment, at least one nucleic acid sequence encoding the polypeptides having transcription factor activity and/or the transcription factors is introduced into a cell on a single plasmid.

- 30 In one embodiment, the plasmid is episomal. Episomal vectors are able to introduce large fragments of DNA into a cell but are maintained extra-chromosomally, replicated once per cell cycle, partitioned to daughter cells efficiently, and elicit substantially no immune response. In alternative embodiments, an Epstein-Barr virus (EBV)-based episomal vector, a yeast-based vector, an adenovirus-based vector, a simian virus 40 (SV40)-based episomal vector, or a
35 bovine papilloma virus (BPV)-based vector may be used.

Site-specific delivery

Any suitable technique for insertion of a nucleic acid sequence into a specific sequence may be used, and several are described in the art. Suitable techniques include any method which introduces a break at the desired location and permits recombination of the vector into the gap.

5 Thus, a crucial first step for targeted site-specific genomic modification is the creation of a double-strand DNA break (DSB) at the genomic locus to be modified. Distinct cellular repair mechanisms can be exploited to repair the DSB and to introduce the desired sequence, and these are non-homologous end joining repair (NHEJ), which is more prone to error; and homologous recombination repair (HR).

10

Several techniques exist to allow customized site-specific generation of DSB in the genome. Many of these involve the use of customized endonucleases, such as zinc finger nucleases, TALENs or the clustered regularly interspaced short palindromic repeats/CRISPR associated protein (CRISPR/Cas, e.g. CRISPR/Cas9) system.

15

Zinc finger nucleases are artificial enzymes which are generated by fusion of a zinc-finger DNA-binding domain to the nuclease domain of the restriction enzyme FokI. The latter has a non-specific cleavage domain which must dimerise in order to cleave DNA. This means that two zinc finger nuclease monomers are required to allow dimerisation of the FokI domains and
20 to cleave the DNA. The DNA binding domain may be designed to target any genomic sequence of interest, is a tandem array of Cys2His2 zinc fingers, each of which recognises three contiguous nucleotides in the target sequence. The two binding sites are separated by 5-7bp to allow optimal dimerization of the FokI domains. The enzyme thus is able to cleave DNA at a specific site, and target specificity is increased by ensuring that two proximal DNA-binding
25 events must occur to achieve a double-strand break.

Transcription activator-like effector nucleases, or TALENs, are dimeric transcription factor/nucleases. They are made by fusing a TAL effector DNA-binding domain to a DNA cleavage domain (a nuclease). Transcription activator-like effectors (TALEs) can be
30 engineered to bind practically any desired DNA sequence, so when combined with a nuclease, DNA can be cut at specific locations. TAL effectors are proteins that are secreted by *Xanthomonas* bacteria, the DNA binding domain of which contains a repeated highly conserved 33–34 amino acid sequence with divergent 12th and 13th amino acids. These two positions are highly variable and show a strong correlation with specific nucleotide recognition.
35 This straightforward relationship between amino acid sequence and DNA recognition has allowed for the engineering of specific DNA-binding domains by selecting a combination of

repeat segments containing appropriate residues at the two variable positions. TALENs are thus built from arrays of 33 to 35 amino acid modules, each of which targets a single nucleotide. By selecting the array of modules, almost any sequence may be targeted. Again, the nuclease used may be FokI or a derivative thereof.

5

Three types of CRISPR mechanisms have been identified, of which type II is the most studied. The CRISPR/Cas9 system (type II) utilises the Cas9 nuclease to make a double-stranded break in DNA at a site determined by a short guide RNA. The CRISPR/Cas system is a prokaryotic immune system that confers resistance to foreign genetic elements. CRISPR are segments of prokaryotic DNA containing short repetitions of base sequences. Each repetition is followed by short segments of "protospacer DNA" from previous exposures to foreign genetic elements. CRISPR spacers recognize and cut the exogenous genetic elements using RNA interference. The CRISPR immune response occurs through two steps: CRISPR-RNA (crRNA) biogenesis and crRNA-guided interference. CrRNA molecules are composed of a variable sequence transcribed from the protospacer DNA and a CRISPR repeat. Each crRNA molecule then hybridizes with a second RNA, known as the trans-activating CRISPR RNA (tracrRNA) and together these two eventually form a complex with the nuclease Cas9. The protospacer DNA encoded section of the crRNA directs Cas9 to cleave complementary target DNA sequences, if they are adjacent to short sequences known as protospacer adjacent motifs (PAMs). This natural system has been engineered and exploited to introduce DSB breaks in specific sites in genomic DNA, amongst many other applications. In particular, the CRISPR type II system from *Streptococcus pyogenes* may be used. At its simplest, the CRISPR/Cas9 system comprises two components that are delivered to the cell to provide genome editing: the Cas9 nuclease itself and a gRNA. The gRNA is a fusion of a customised, site-specific crRNA (directed to the target sequence) and a standardised tracrRNA.

25

Once a DSB has been made, a donor template with homology to the targeted locus is supplied; the DSB may be repaired by the homology-directed repair (HDR) pathway allowing for precise insertions to be made.

30

Derivatives of this system are also possible. Mutant forms of Cas9 are available, such as Cas9D10A, with only nickase activity. This means it cleaves only one DNA strand, and does not activate NHEJ. Instead, when provided with a homologous repair template, DNA repairs are conducted via the high-fidelity HDR pathway only. Cas9D10A may be used in paired Cas9 complexes designed to generate adjacent DNA nicks in conjunction with two sgRNAs

35

complementary to the adjacent area on opposite strands of the target site, which may be particularly advantageous.

5 The elements for making the double-strand DNA break may be introduced in one or more vectors, such as plasmids, for expression in the cell.

10 Thus, any method of making specific, targeted double strand breaks in the genome in order to effect the insertion of a gene/inducible cassette may be used in the method of the invention. It may be preferred that the method for inserting the gene/inducible cassette utilises any one or more of zinc finger nucleases, TALENs and/or CRISPR/Cas9 systems or any derivative thereof.

15 Once the DSB has been made by any appropriate means, the gene/inducible cassette for insertion may be supplied in any suitable fashion as described below. The gene/inducible cassette and associated genetic material form the donor DNA for repair of the DNA at the DSB and are inserted using standard cellular repair machinery/pathways. How the break is initiated will alter which pathway is used to repair the damage, as noted above.

20 Other methods in the art for site specific delivery include the use of homologous recombination (HR) and recombinase mediated cassette exchange (RMCE). DNA damage mediated site specific insertion methods (such as CRISPR/Cas) can also be used to perform site specific integration of DNA recognition sequences ('att' sites) which in turn mediate site specific insertion via the activity of tyrosine and serine recombinases or integrases. These sites (e.g. attP) once inserted into the genome, can mediate site specific HR and RMCE. Insertion of
25 exogenous nucleic acid sequences occurs through homologous recombination between cognate attP and attB sites mediated by the expression of the appropriate and cognate recombinase (e.g. Flp, Cre) or integrase (PhiC31, Bxb1). Using targeting vectors, as described above, flanked by attB sites, site specific exogenous DNA insertion of transgenes can be achieved.

30

Controlled expression

In one embodiment, expression of the transcription factors is under inducible control. In this aspect of the invention, the transcription and translation (expression) of the transcription factors may be controlled within the cell. This permits overexpression of the transcription
35 factor(s), if required.

In an alternative embodiment, expression of the polypeptides having transcription factor activity and/or the transcription factors themselves is under inducible control. In this aspect of the invention, the transcription and translation (expression) of the polypeptides having transcription factor activity and/or the transcription factors may be controlled within the cell.

5 This permits overexpression of the transcription factor(s), if required, preferably in response to external stimuli.

An exogenous expression cassette carrying the polypeptides having transcription factor activity and/or the transcription factors themselves may comprise an externally inducible transcriptional regulatory element (i.e., an inducible promoter) for rapid induction of protein
10 expression in response to external stimuli, i.e. inducible gene (or transgene) expression. The presence or addition of the appropriate external stimuli (e.g. protein, compound or chemical) to cell culture media modulates the controlled expression of the genetic sequence within the inducible expression cassette; and may be administered continuously or transiently to
15 modulate transcription as required.

Expression of the transcription factors described herein may be increased using a dual cassette expression system, such as the system described in WO2018096343, which is incorporated herein by reference. In this instance, induced transgene over-expression is
20 achieved by using the Tet-ON system components with transgene expression controlled by doxycycline. The components are split between two genetic safe harbour sites (GSH) to reduce the risk of epigenetic gene silencing. The components are (i) transcriptional activator protein (reverse tetracycline trans-activator (rtTA)), which in the presence of doxycycline binds (ii) tetracycline response element (TRE; multiple TetO repeat sequences & minimal
25 Cytomegalovirus (CMV) promoter). TRE binding by rtTA trans-activates transgene expression. Trans-activatable coding sequences for transgenes may be of human origin.

Therefore, in one embodiment, a sequence encoding one or more (e.g., two or more or three or more) of the polypeptides having transcription factor activity and/or the transcription factors
30 is introduced into the cell population, preferably a pluripotent stem cell population, more preferably a hiPSC population, using a method comprising:

- insertion (preferably targeted insertion) of a coding sequence for a transcriptional regulator protein into a first genomic safe harbour site of a source cell present in the cell population; and
- 35 - insertion (preferably targeted insertion) of one or more inducible cassettes into one or more second genomic safe harbour sites of the source cell, wherein said one or more inducible

cassettes comprises said sequence encoding the one or more polypeptides having the activity of one or more transcription factors and/or the transcription factors operably linked to an inducible promoter, and said promoter is regulated by the transcriptional regulator protein.

- 5 In one embodiment, a sequence encoding one or more (e.g., two or more or three or more) of the transcription factors is introduced into the cell population using a method comprising:
- targeted insertion of a coding sequence for a transcriptional regulator protein into a first genomic safe harbour site of a source cell present in the cell population; and
 - targeted insertion of an inducible cassette into a second genomic safe harbour site of
- 10 the source cell, wherein said inducible cassette comprises said sequence encoding one or more transcription factors operably linked to an inducible promoter, and said promoter is regulated by the transcriptional regulator protein.

According to another aspect of the invention, there is provided a dual expression system

15 comprising:

- (a) a first expression cassette comprising a gene encoding a transcriptional regulator protein flanked by one or more homology arms targeting the first expression cassette to a first genomic safe harbour site; and
 - (b) a second expression cassette comprising a sequence encoding one or more
- 20 transcription factors operably linked to an inducible promoter, flanked by one or more homology arms targeting the second expression cassette to a second genomic safe harbour site, wherein the inducible promoter is regulated by the transcriptional regulator protein of the first expression cassette, and wherein the one or more transcription factors are selected from the group consisting of: a PPAR protein, HOXC8, EBF1, EBF2, ZNF467, ZNF423 and variants
- 25 thereof.

According to another aspect of the invention, there is provided a dual expression system comprising:

- (a) a first expression cassette comprising a gene encoding a transcriptional regulator
- 30 protein flanked by one or more homology arms targeting the first expression cassette to a first genomic safe harbour site; and
- (b) a second expression cassette comprising a sequence encoding one or more polypeptides having the activity of one or more transcription factors and/or transcription factors operably linked to an inducible promoter, flanked by one or more homology arms targeting the
- 35 second expression cassette to a second genomic safe harbour site, wherein the inducible promoter is regulated by the transcriptional regulator protein of the first expression cassette,

and wherein the one or more transcription factors are selected from the group consisting of: one or more PPAR proteins (such as PPARA and/or PPARG), HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB) and variants thereof.

5 This embodiment of the invention provides a dual expression cassette system. The insertion of the gene encoding a transcriptional regulator protein into the first GSH site provides the control mechanism for the expression of the inducible cassette which is operably linked to the inducible promoter and inserted into a second GSH site. In one embodiment, the first and second GSH site are different (i.e. are located at different positions in the genome). It will be
10 understood that if more than one transcription factor is to be introduced into the cell using the dual expression system, then the transcription factors may be introduced into the second GSH site (such as within a multicistronic cassette at the same GSH site), or into multiple GSH sites (i.e. as separate cassettes across different GSH sites).

15 Alternatively, the dual expression cassette system utilises different alleles of the same GSH site. In this embodiment, the inducible cassette may be inserted into one allele of the GSH site and the system controlling the expression of the inducible cassette into the other allele of the GSH site (e.g. as described in DeKolver *et al.*, 2010, *Genome Res.*, 20, 1133-43 and Qian *et al.*, 2014, *Stem Cells*, 32, 1230-8).

20

A GSH site is a locus within the genome wherein a gene or other genetic material may be inserted without any deleterious effects on the cell or on the inserted genetic material. Most beneficial is a GSH site in which expression of the inserted gene sequence is not perturbed by any read-through expression from neighbouring genes and expression of the inducible
25 cassette minimizes interference with the endogenous transcription programme. More formal criteria have been proposed that assist in the determination of whether a particular locus is a GSH site in future (Papapetrou *et al.* (2011) *Nature Biotechnology*, 29(1): 73–8) These criteria include a site that is (i) 50 kb or more from the 5' end of any gene, (ii) 300 kb or more from any gene related to cancer, (iii) 300 kb or more from any microRNA (miRNA), (iv) located outside
30 a transcription unit and (v) located outside ultraconserved regions (UCR). It may not be necessary to satisfy all of these proposed criteria, since GSH sites already identified do not fulfil all of the criteria. It is thought that a suitable GSH site will satisfy at least 2, 3, 4 or all of these criteria. Any suitable GSH site may be used in the method of the invention, on the basis that the site allows insertion of genetic material without deleterious effects to the cell and
35 permits transcription of the inserted genetic material. Those skilled in the art may use these simplified criteria to identify a suitable GSH site, and/or the more formal criteria set out above.

Insertion of the coding sequence for a transcriptional regulator protein and/or the inducible cassette may be carried out through direct delivery methods as described above. It is understood that although such direct delivery methods may lead to the random insertion of the genetic material, screening may be carried out in order to identify clones that show no deleterious effects, are able to express the genetic material and are able to be forward programmed or reprogrammed to adipocytes, and by doing so one is able to confirm that the transcriptional regulator protein / inducible cassette has been inserted into a GSH site.

10 In one embodiment the insertion of the transcriptional regulator protein or the inducible cassette is targeted. In a further embodiment, the insertion of the transcriptional regulator protein and the inducible cassette is targeted. "Targeted insertion", as with site-specific delivery, is understood as the insertion of the genetic material into a pre-chosen GSH site. As discussed above, this can be carried out using techniques known in the art such as zinc finger
15 nucleases, TALENs or the clustered regularly interspaced short palindromic repeats/CRISPR associated protein (CRISPR/Cas, e.g. CRISPR/Cas9) system.

In one embodiment, the first and second genetic safe harbour sites (GSH sites) are selected from (in particular any two) of the hROSA26 locus, the AAVS1 locus, the *CLYBL* gene, the
20 *CCR5* gene or the *HPRT* gene. Insertions specifically within genetic safe harbour sites is preferred over random genome integration, since this is expected to be a safer modification of the genome, and is less likely to lead to unwanted side effects such as silencing natural gene expression or random insertional mutagenesis.

25 The adeno-associated virus integration site 1 locus (AAVS1) is located within the protein phosphatase 1, regulatory subunit 12C (PPP1R12C) gene on human chromosome 19, which is expressed uniformly and ubiquitously in human tissues. AAVS1 has been shown to be a favourable environment for transcription, since it comprises an open chromatin structure and native chromosomal insulators that enable resistance of the inducible cassettes against
30 silencing. There are no known adverse effects on the cell resulting from disruption of the PPP1R12C gene. Moreover, an inducible cassette inserted into this site remains transcriptionally active in many diverse cell types.

The human ROSA26 (hROSA26) site has been identified on the basis of sequence analogy
35 with a GSH site from mice (ROSA26 – reverse oriented splice acceptor site #26). The hROSA26 locus is on chromosome 3 (3p25.3), and can be found within the Ensembl database

(GenBank:CR624523). The integration site lies within the open reading frame (ORF) of the THUMPD3 long non-coding RNA (reverse strand). Since the hROSA26 site has an endogenous promoter, the inserted genetic material may take advantage of that endogenous promoter, or alternatively may be inserted operably linked to a promoter.

5

Intron 2 of the Citrate Lyase Beta-like (*CLYBL*) gene, on the long arm of Chromosome 13, was identified as a suitable GSH site since it is one of the identified integration hot-spots of the phage derived phiC31 integrase. Studies have demonstrated that randomly inserted inducible cassettes into this locus are stable and expressed. It has been shown that insertion of inducible cassettes at this GSH site does not perturb local gene expression (Cerbini *et al.* (2015) PLOS One, 10(1): e0116032). *CLYBL* thus provides a GSH site which may be suitable for use in the present invention.

CCR5, which is located on chromosome 3 (position 3p21.31) is a gene which codes for HIV-1 major co-receptor. Interest in the use of this site as a GSH site arises from the null mutation in this gene that appears to have no adverse effects, but predisposes to HIV-1 infection resistance. Zinc-finger nucleases that target the third exon have been developed, thus allowing for insertion of genetic material at this locus.

The hypoxanthine-guanine phosphoribosyltransferase (*HPRT*) gene encodes a transferase enzyme that plays a central role in the generation of purine nucleotides through the purine salvage pathway.

Other GSH sites have been described in the art, such as in Sadelain *et al.* (2012) Nature Reviews 12:51-58 and in WO2021/152086, which are herein incorporated by reference.

GSH sites in other organisms have been identified and include ROSA26, *HRPT* and *Hipp11* (*H11*) loci in mice. Mammalian genomes may include GSH sites based upon pseudo attP sites. For such sites, phiC31 integrase, the *Streptomyces* phage-derived recombinase, has been developed as a non-viral insertion tool, because it has the ability to integrate an inducible cassette-containing plasmid carrying an attB site into pseudo attP sites.

Technically, the insertions into the first and/or second GSH site may occur on one chromosome, or on both chromosomes. The GSH sites exist at the same genetic loci on both chromosomes of diploid organisms. Insertion within both chromosomes is advantageous since

it may enable an increase in the level of transcription from the inserted genetic material within the inducible cassette, thus achieving particularly high levels of transcription.

Specific insertion of genetic material into the particular GSH site based upon customised site-specific generation of DNA double-strand breaks at the GSH site may be achieved. The genetic material may then be introduced using any suitable mechanism, such as homologous recombination. Any method of making a specific DSB in the genome may be used, but preferred systems include CRISPR/Cas9 and modified versions thereof, zinc finger nucleases and the TALEN system, or via HR or RCME mediated integration or recombination.

One or more genetic sequences may be controllably transcribed from within the second and/or further GSH sites. Indeed, the inducible cassette may contain 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 genetic sequences (e.g., transcription factor sequences) which it is desired to insert into the GSH site and the transcription of which be controllably induced. Therefore, the transcription factors required by the present invention may be included within the same cassette introduced into the second genetic safe harbour site. For example, the three or more transcription factors may be included in, for example, three mono-cistronic constructs, one mono-cistronic and one bi-cistronic construct or one tri-cistronic construct. It will be understood that similar combinations of constructs may be used to achieve higher orders of transcription factor expression.

Alternatively, if a combination of transcription factors is used, the individual transcription factors may be introduced into separate GSH sites and/or under the control of the same, different or orthogonal inducible promoters. Therefore, in one embodiment, the transcription factors are introduced into separate GSH sites. For example, this may be achieved by utilising three or more different GSH sites for three or more transcription factors (i.e., wherein the transcription factors are introduced as mono-cistronic cassettes). Alternatively, this may be achieved by utilising the fact that a GSH site exists at the same genetic loci on both chromosomes of diploid organisms, e.g., introducing one transcription factor into the GSH site on one chromosome and a different transcription factor into the same GSH site on the other chromosome. This embodiment is advantageous if different expression levels or timing of expression of the transcription factors is desired. In one embodiment, the method comprises targeted insertion of the transcription factors, each operably linked to an inducible promoter into a second, third and fourth genetic safe harbour site of the source cell. The inducible promoter may be the same of each transcription factor and therefore are all regulated by the transcriptional regulator protein.

A transcriptional regulator protein is a protein that binds to DNA, preferably sequence-specifically to a DNA site located in or near a promoter, and either facilitating the binding of the transcription machinery to the promoter, and thus transcription of the DNA sequence (a
5 transcriptional activator) or blocks this process (a transcriptional repressor).

The DNA sequence that a transcriptional regulator protein binds to is called a transcription factor-binding site or response element, and these are found in or near the promoter of the regulated DNA sequence. Transcriptional activator proteins bind to the response element and
10 promote gene expression. Such proteins are preferred in the methods of the present invention for controlling inducible cassette expression. Transcriptional repressor proteins bind to the response element and prevent gene expression.

Transcriptional regulator proteins may be activated or deactivated by a number of mechanisms
15 including binding of a substance, interaction with other transcription factors (e.g., homo- or hetero-dimerization) or coregulatory proteins, phosphorylation, and/or methylation. The transcriptional regulator protein may be controlled by activation or deactivation.

If the transcriptional regulator protein is a transcriptional activator protein, it is preferred that
20 the transcriptional activator protein requires activation. This activation may be through any suitable means, but it is preferred that the transcriptional regulator protein is activated through the addition to the cell of an exogenous substance. The supply of an exogenous substance to the cell can be controlled, and thus the activation of the transcriptional regulator protein can be controlled. Alternatively, an exogenous substance can be supplied in order to deactivate a
25 transcriptional regulator protein, and then supply withdrawn in order to activate the transcriptional regulator protein.

If the transcriptional regulator protein is a transcriptional repressor protein, it is preferred that the transcriptional repressor protein requires deactivation. Thus, a substance is supplied to
30 prevent the transcriptional repressor protein repressing transcription, and thus transcription is permitted.

Any suitable transcriptional regulator protein may be used, preferably one that may be activated or deactivated. It is preferred that an exogenous substance may be supplied to
35 control the transcriptional regulator protein. Such transcriptional regulator proteins are also called inducible transcriptional regulator proteins.

Tetracycline-Controlled Transcriptional Activation is a method of inducible gene expression where transcription is reversibly turned on or off in the presence of the antibiotic tetracycline or one of its derivatives (e.g., doxycycline which is more stable). In this system, the transcriptional activator protein is reverse tetracycline-controlled transactivator (rtTa, which may also be referred to as tetracycline – responsive transcriptional activator protein) or a derivative thereof. The rtTA protein is able to bind to DNA at specific TetO operator sequences. Several repeats of such TetO sequences are placed upstream of a minimal promoter (such as the CMV promoter), which together form a tetracycline response element (TRE). There are two forms of this system, depending on whether the addition of tetracycline or a derivative activates (Tet-On) or deactivates (Tet-Off) the rtTA protein.

In a Tet-Off system, tetracycline or a derivative thereof binds rtTA and deactivates the rtTA, rendering it incapable of binding to TRE sequences, thereby preventing transcription of TRE-controlled genes. This system was first described in Gossen *et al.* (1992) PNAS 89 (12): 5547–5551.

The Tet-On system is composed of two components; (1) the constitutively expressed reverse tetracycline-controlled transactivator (rtTa) and the rtTa-sensitive inducible promoter (Tet Responsive Element, TRE). This may be bound by tetracycline or its more stable derivatives, including doxycycline (dox), resulting in activation of rtTa, allowing it to bind to TRE sequences and inducing expression of TRE-controlled genes. The use of this may be preferred in the method of the invention.

Thus, the transcriptional regulator protein may thus be a reverse tetracycline-controlled transactivator (rtTa) protein, which can be activated or deactivated by the antibiotic tetracycline or one of its derivatives, which are supplied exogenously. If the transcriptional regulator protein is rtTA, then the inducible promoter inserted into the second GSH site includes the tetracycline response element (TRE). The exogenously supplied substance is the antibiotic tetracycline or one of its derivatives.

Variants and modified rtTa proteins may also be used in the methods of the invention, these include Tet-On Advanced transactivator (also known as rtTA2S-M2) and Tet-On 3G (also known as rtTA-V16, derived from rtTA2S-S2).

The tetracycline response element (TRE) generally consists of 7 repeats of the 19bp bacterial TetO sequence separated by spacer sequences, together with a minimal promoter. Variants and modifications of the TRE sequence are possible, since the minimal promoter can be any suitable promoter. Preferably the minimal promoter shows no or minimal expression levels in the absence of rtTa binding. The inducible promoter inserted into the second GSH site may thus comprise a TRE.

A modified system based upon tetracycline control is the T-REX System (Thermo-Fisher Scientific), in which the transcriptional regulator protein is a transcriptional repressor protein, TetR. The components of this system include (i) an inducible promoter comprising a strong human cytomegalovirus immediate-early (CMV) promoter and two tetracycline operator 2 (TetO2) sites, and a Tet repressor (TetR). In the absence of tetracycline, the Tet repressor forms a homodimer that binds with extremely high affinity to each TetO2 sequence in the inducible promoter, and prevent transcription from the promoter. Once added, tetracycline binds with high affinity to each Tet repressor homodimer rendering it unable to bind to the Tet operator. The Tet repressor: tetracycline complex then dissociates from the Tet operator and allows induction of expression. In this instance, the transcriptional regulator protein is TetR and the inducible promoter comprises two TetO2 sites. The exogenously supplied substance is tetracycline or a derivative thereof.

Other inducible expression systems are known and can be used in the method of the invention. These include the Complete Control Inducible system from Agilent Technologies. This is based upon the insect hormone ecdysone or its analogue ponasterone A (ponA) which can activate transcription in mammalian cells which are transfected with both the gene for the Drosophila melanogaster ecdysone receptor (EcR) and an inducible promoter comprising a binding site for the ecdysone receptor. The EcR is a member of the retinoid-X-receptor (RXR) family of nuclear receptors. In humans, EcR forms a heterodimer with RXR that binds to the ecdysone-responsive element (EcRE). In the absence of PonA, transcription is repressed by the heterodimer.

Thus, the transcriptional regulator protein can be a repressor protein, such as an ecdysone receptor or a derivative thereof. Examples of the latter include the VgEcR synthetic receptor from Agilent technologies which is a fusion of EcR, the DNA binding domain of the glucocorticoid receptor and the transcriptional activation domain of Herpes Simplex Virus VP16. The inducible promoter comprises the EcRE sequence or modified versions thereof together with a minimal promoter. Modified versions include the E/GRE recognition sequence

of Agilent Technologies, in which mutations to the sequence have been made. The E/GRE recognition sequence comprises inverted half-site recognition elements for the retinoid-X-receptor (RXR) and GR binding domains. In all permutations, the exogenously supplied substance is ponasterone A, which removes the repressive effect of EcR or derivatives thereof
5 on the inducible promoter, and allows transcription to take place.

Alternatively, inducible systems may be based on the synthetic steroid mifepristone as the exogenously supplied substance. In this scenario, a hybrid transcriptional regulator protein is inserted, which is based upon a DNA binding domain from the yeast GAL4 protein, a truncated
10 ligand binding domain (LBD) from the human progesterone receptor and an activation domain (AD) from the human NF- κ B. This hybrid transcriptional regulator protein is available from Thermo-Fisher Scientific (Gene Switch™). Mifepristone activates the hybrid protein, and permits transcription from the inducible promoter which comprises GAL4 upstream activating sequences (UAS) and the adenovirus E1b TATA box. This system is described in Wang *et al.*
15 (1994) PNAS 91: 8180-8184.

The transcriptional regulator protein can thus be any suitable regulator protein, either an activator or repressor protein. Suitable transcriptional activator proteins are tetracycline-responsive transcriptional activator protein or the Gene Switch hybrid transcriptional regulator
20 protein. Suitable repressor proteins include the Tet-Off version of rtTA, TetR or EcR. The transcriptional regulator proteins may be modified or derivatised as required.

The inducible promoter can comprise elements which are suitable for binding or interacting with the transcriptional regulator protein. The interaction of the transcriptional regulator protein
25 with the inducible promoter is preferably controlled by the exogenously supplied substance.

The exogenously supplied substance can be any suitable substance that binds to or interacts with the transcriptional regulator protein. Suitable substances include tetracycline (or derivatives thereof, such as doxycycline), ponasterone A and mifepristone.
30

It is preferred that the gene encoding the transcriptional regulator protein is operably linked to a constitutive promoter. Alternatively, the first GSH site can be selected such that it already has a constitutive promoter than can also drive expression of the transcriptional regulator protein gene and any associated genetic material. Constitutive promoters ensure sustained
35 and high-level gene expression. Commonly used constitutive promoters, including the human β -actin promoter (ACTB), cytomegalovirus (CMV), elongation factor-1 α , (EF1 α),

phosphoglycerate kinase (PGK) and ubiquitin C (UbC). The CAG promoter is a strong synthetic promoter frequently used to drive high levels of gene expression and was constructed from the following sequences: (C) the cytomegalovirus (CMV) early enhancer element, (A) the promoter, the first exon and the first intron of chicken beta-actin gene, and (G) the splice
5 acceptor of the rabbit beta-globin gene.

According to a further aspect of the invention, there is provided a method for the production of adipocytes from a source cell, preferably a pluripotent stem cell, more preferably a hiPSC, comprising the steps of:

- 10 a) insertion (preferably targeted insertion) of a gene encoding a transcriptional regulator protein into a first genomic safe harbour site of the source cell; and
- b) insertion (preferably targeted insertion) of at least one nucleotide sequence encoding one or more polypeptides having the activity of one or more transcription factors and/or one or more transcription factors, the transcription factors selected from the group consisting of: one
15 or more PPAR proteins (such as PPAR and/or PPARG), HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB) and variants thereof, operably linked to an inducible promoter into a second genomic safe harbour site of the source cell, wherein said inducible promoter is regulated by the transcriptional regulator protein; and
- c) culturing the source cell(s) comprising the insertions to obtain adipocytes.

20

According to a further aspect of the invention, there is provided a method for the production of adipocytes from a source cell, comprising the steps of:

- a) targeted insertion of a gene encoding a transcriptional regulator protein into a first genomic safe harbour site of the source cell; and
- 25 b) targeted insertion of at least one nucleotide sequence encoding one or more transcription factors selected from the group consisting of: a PPAR protein, HOXC8, EBF1, EBF2, ZNF467, ZNF423 and variants thereof, operably linked to an inducible promoter into a second genomic safe harbour site of the source cell, wherein said inducible promoter is regulated by the transcriptional regulator protein; and
- 30 c) culturing the source cell(s) comprising the insertions to obtain adipocytes.

It will be understood that these aspects of the invention may be used with any of the combinations of transcription factors described herein.

Obtaining adipocytes

In one embodiment, the method additionally comprises monitoring the cell population for at least one characteristic of an adipocyte. Cells may be monitored throughout culturing to identify
5 expression of key lineage markers.

For example, monitoring may be through the use of engineered 'reporter' cell lines (i.e. endogenously tagged proteins or positive selection markers under the control of adipocyte specific promoters) or immunostaining and detected, using fluorescence microscopy or flow
10 cytometry. Such material includes genes for markers or reporter molecules, such as genes that induce visually identifiable characteristics including fluorescent and luminescent proteins. Examples include the gene that encodes jellyfish green fluorescent protein (GFP), which causes cells that express it to glow green under blue/UV light, luciferase, which catalyses a reaction with luciferin to produce light, and the red fluorescent protein from the gene dsRed.

15 The cell may further comprise a positive selection marker and/or selectable reporter expression cassette, e.g., comprising an adipocyte-specific promoter operably linked to a reporter gene.

20 Selectable markers may include resistance genes to antibiotics or other drugs. Examples of drug resistance genes may include: a puromycin resistance gene, an ampicillin resistance gene, a neomycin resistance gene, a tetracycline resistance gene, a kanamycin resistance gene or a chloramphenicol resistance gene. Cells can be cultured on a medium containing the appropriate drug (i.e., a selection medium) and only those cells which incorporate and express
25 the drug resistance gene will survive. Therefore, by culturing cells using a selection medium, it is possible to select for cells comprising and expressing a drug resistance gene, positively enriching for a target cell population.

Examples of fluorescent protein genes which may be used as markers include: a green
30 fluorescent protein (GFP) gene, yellow fluorescent protein (YFP) gene, red fluorescent protein (RFP) gene or aequorin gene. Cells expressing the fluorescent protein can be detected using a fluorescence microscope and fluorescence activated cell sorting (FACS) used to identify and select cell populations based on the expression of fluorescent proteins.

35 Fluorescent protein genes may be tagged with a nuclear localization signal peptide to confine expression of the fluorescent proteins to the nucleus. This may be helpful in cell types with a

high lipid content which may not be suitable for FACS. This allows end-point fluorescence-activated cell sorting to be carried out on either whole cell populations, or purified nuclei which maintain an intact fluorescent signal.

- 5 Examples of chromogenic enzyme genes which may be used as markers, and known in the art, include but are not limited to: β -galactosidase gene, β -glucuronidase gene, alkaline phosphatase gene, or secreted alkaline phosphatase SEAP gene. Cells expressing these chromogenic enzyme genes can be detected by applying the appropriate chromogenic substrate (e.g., X-gal for β galactosidase) so that cells expressing the marker gene will produce
10 a detectable colour (e.g., blue in a blue-white screen test).

- The method may therefore comprise a selection or enrichment step for adipocytes provided from the methods described herein. In one embodiment, the method comprises the step of
15 sorting the adipocytes using fluorescence activated cell sorting (FACS) or immunomagnetic sorting methods based on the expression of adipocyte markers and/or absence of non-adipocyte cell markers. A labelled binding agent directed to target cell surface proteins may be used. Any binding agent capable of specific binding to a particular epitope may be used for this purpose, for example an antibody or a fragment thereof, a peptide or a synthetic binder such as a plastic antibody, or an aptamer or oligonucleotide, capable of specific binding to an
20 epitope. The binding agent may be labelled with a detectable marker, such as a luminescent, fluorescent (e.g. fluorochrome), enzyme or radioactive marker; alternatively or additionally an affinity tag, e.g. a biotin, avidin, streptavidin or His (e.g. hexa-His) tag. In one embodiment, fluorochrome conjugated antibodies targeting cell surface proteins (e.g. adipocyte markers) may be used to sort target cells.

- 25 In another embodiment, adipocytes are enriched by drug-resistance selection from genetically engineered source cells expressing an antibiotic-resistance gene under the control of an adipocyte-specific promoter.

- 30 The method may generate cells (i.e., differentiated cells) exhibiting at least one characteristic of an adipocyte. One or more characteristics may be used to select for the adipocytes generated by methods of the invention.

- Characteristics include but are not limited to the detection or quantitation of expressed cell
35 markers, enzymatic activity, and the characterization of morphological features and intercellular signalling. The biological function of an adipocyte may also be evaluated, for

example using functional assays, e.g. secretion of adipokines (e.g., adiponectin and leptin) and response to insulin (i.e. sensitivity to insulin and/or production of insulin sensitizing and anti-inflammatory agents).

- 5 In one embodiment, the characteristic (i.e., of an adipocyte, in particular a human adipocyte) is selected from one or more of:
- (i) expression of one or more cell markers, such as FABP4, PLIN1, or a combination thereof;
 - (ii) adipokine expression or response to insulin; or
 - 10 (iii) adipocyte morphological features.

In one embodiment, the cells are sorted on the basis of acquisition of expression of a mature adipocyte marker, such as FABP4 and PLIN1. In another embodiment, the cells are sorted on the basis of acquisition of expression of an adipocyte marker, such as CEBPA, CEBPB and

15 CD36. In another embodiment, the cells are sorted on the basis of acquisition of expression of a brown adipocyte marker, such as UCP1.

The adipocyte markers may be markers obtained by transcriptome analysis. For example, single cell RNA sequencing has been used to provide detailed transcriptional profiles of human

20 adipocytes obtained from primary human tissues. This information can be used to identify adipocytes generated by the methods described herein. Additional resources, such as Human Cell Atlas and CellTypist may also be used to identify markers of adipocytes.

The method may comprise assaying the differentiated cells obtained by the method described herein and determining a set of transcribed genes; comparing the set of transcribed genes of the differentiated cells to one or more reference sets of transcribed genes from one or more reference adipocytes; and identifying a match between the differentiated cells and a reference

25 adipocyte.

30 In one embodiment, the method comprises the step of identifying differentiated cells as a type of adipocyte by assaying morphological features of the differentiated cells and matching the morphological features to a reference tissue or cell's morphological features.

In one embodiment, the method comprises the step of identifying differentiated cells as a type

35 of adipocyte by assaying protein marker expression of the differentiated cells and matching the protein marker expression to a reference adipocyte protein marker expression.

In one embodiment, the method comprises the step of identifying differentiated cells as a type of adipocyte by assaying a function and matching the function to a function of a reference adipocyte.

5

In one embodiment, the cells obtained by the methods of the invention express an adipocyte cell phenotype. This phenotype may be defined through expression (+) or non-expression (-) of one of more of the following markers: FABP4+, PLIN1+, CEBPA+, CEBPB+, CD36+ and UCP1+. The target cells may also be negative for markers of pluripotency.

10

Alternatively, certain differentiated cells may be sorted from other differentiated cells and from cells on the basis of their expression of a lineage-specific cell surface antigen. Yet another means is by assessing expression at the RNA level, e.g., by RT-qPCR methods or by single cell RNA sequencing without any sorting or pre-selection step. Such techniques are known in the art.

15

Cell culturing

In one embodiment, the method includes culturing the cell population for a sufficient time and under conditions to allow differentiation to an adipocyte. Generally, cells of the present invention are cultured in a culture medium, which is a nutrient-rich buffered solution capable of sustaining cell growth.

20

The cell culture medium may contain any of the following in an appropriate combination: salt(s), buffer(s), amino acids, glucose or other sugar(s), antibiotics, serum or serum replacement, and other components such as peptide growth factors, etc. Cell culture media ordinarily used for particular cell types are known to those skilled in the art. For example, the media may comprise Basal Medium (e.g. DMEM/F12 or STEMPRO-34) supplemented with GLUTAMAX, antibiotics (such as penicillin or streptomycin), B27 supplement and/or N2 supplement (all available from Thermo Fisher Scientific). The media may then be further supplemented at different time points during the culturing process. For example, one or more peptide hormones and/or cytokines can be at 2, 4 and/or 10 days during the culturing process.

25

30

In one embodiment, the culture media comprises one or more components selected from the group consisting of: Bone Morphogenetic Protein 4 (BMP4), Activin A, Fibroblast Growth Factor 2 (FGF2), Insulin, Ascorbic acid and Dexamethasone. In one embodiment, the culture

35

media comprises one or more peptide hormones and/or cytokines selected from the group consisting of: BMP4, Activin A, FGF2 and Insulin.

Adipocytes may be obtained using methods of the invention at least about 3, 4, 5, 6, 7, 8, 9,
5 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 days after culturing. In one embodiment, the method
comprises culturing under suitable conditions for at least 4 days, such as at least 7 days or
about 10 days. In further embodiments, method comprises culturing cells for a duration (e.g.,
at least 4 days, at least 5 days, at least 6 days, at least 7 days, at least 8 days, at least 9 days,
at least 10 days, at least 11 days, at least 12 days, at least 13 days, at least 14 days, at least
10 21 days, at least 28 days, or longer, e.g., from 5 days to 40 days, from 7 days to 35 days, from
14 days to 28 days, or about 21 days) which is sufficient to generate adipocytes. In some
embodiments, the cells are cultured for a period of several hours (e.g., about 2, 3, 4, 5, 6, 7,
8, 9, 10, 15, 18, or 21 hours) to about 35 days (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14,
15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, or 35 days). In one
15 embodiment, the method comprises culturing the cells for at least about 5, 10, 15 or 20 days
to produce adipocytes. In one embodiment, the cells are cultured for a period of between 4
and 25 days, such as between 7 and 14 days.

After culturing, the cell population may comprise two cell types. For example, such a cell
20 population may have two cell types including the stem cells and adipocytes. In one
embodiment, the cell population comprises up to 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60,
70, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 99.5% (or any intermediate ranges) of
adipocytes in the resulting cell population.

25 Culturing the cells may either help to induce cells to commit to a more mature phenotype,
preferentially promote survival of the mature cells, or have a combination of both these effects.

According to a further aspect of the invention, there is provided a cell obtainable by any one of
the methods defined herein.

30

According to a further aspect of the invention, there is provided a cell comprising one or more
exogenous expression cassettes comprising nucleotide sequences encoding at least one or
more transcription factors, wherein the one or more transcription factors are selected from the
group consisting of: a PPAR protein, HOXC8, EBF1, EBF2, ZNF467, ZNF423, and variants
35 thereof.

According to a further aspect of the invention, there is provided a cell comprising one or more exogenous expression cassettes comprising nucleotide sequences encoding at least one or more polypeptides having the activity of one or more transcription factors and/or transcription factors, wherein the one or more transcription factors are selected from the group consisting of: one or more PPAR proteins (such as PPARA and/or PPARG), HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB) and variants thereof.

As described herein, the exogenous expression cassettes encoding the one, two or three or more polypeptides having transcription factor activity and/or transcription factors may be integrated into the genome of the cell. In a further embodiment, exogenous expression cassettes encoding the two or more polypeptides having transcription factor activity and/or transcription factors are integrated into a (specific) target site in the genome of the cell. Alternatively, exogenous expression cassettes encoding the two or more polypeptides having transcription factor activity and/or transcription factors are integrated into a non-specific target site in the genome of the cell.

Cell compositions

According to a further aspect, there is provided a pharmaceutical composition comprising the adipocytes produced by the method as described herein and a pharmaceutically acceptable carrier.

Pharmaceutical compositions may include adipocytes as described herein in combination with one or more pharmaceutically or physiologically acceptable carrier, diluents, or excipients. Such compositions may include buffers such as neutral buffered saline, phosphate buffered saline and the like; carbohydrates such as glucose, mannose, sucrose or dextrans, mannitol; proteins; polypeptides or amino acids such as glycine; antioxidants; chelating agents such as EDTA or glutathione; adjuvants (e.g., aluminium hydroxide); and preservatives. Cryopreservation solutions which may be used in the pharmaceutical compositions of the invention include, for example, DMSO.

For purposes of manufacture, distribution, and use, the adipocytes described herein may be supplied in the form of a cell culture or suspension in an isotonic excipient or culture medium, optionally frozen to facilitate transportation or storage.

Uses of adipocytes

The cells produced according to any of the methods of the invention have applications in basic and medical research, diagnostic and therapeutic methods. The cells may be used *in vitro* to study cellular development, provide test systems for new drugs, enable screening methods to be developed, scrutinise therapeutic regimens, provide diagnostic tests and the like. These uses form part of the present invention. Alternatively, the cells may be transplanted into a human or animal patient for diagnostic or therapeutic purposes. The use of the cells in therapy is also included in the present invention.

According to one aspect of the invention, there is provided an adipocyte as defined herein, for use in *in vitro* diagnostics or drug screening.

Adipocytes generated by methods of the invention may find particular use in drug screening.

Therefore, in one embodiment, the method additionally comprises contacting the adipocytes with a test substance and observing a change (e.g., an effect) in the adipocytes induced by the test substance. The change or effect may be observed using methods known in the art, for example using pharmacological or toxicological assays. In one aspect, the cells may be used in a method of assessing a test substance (e.g., a drug, such as a compound), comprising assaying a pharmacological or toxicological property of the test substance on the adipocytes provided by the methods described herein. The method may comprise: a) contacting the adipocytes described herein with the test substance; and b) assaying an effect of the test substance on the adipocytes.

Assessment of the activity of a candidate molecule may involve combining the adipocytes described herein with the candidate molecule, determining any change in the morphology, phenotype, or metabolic activity of the adipocytes that is attributable to the molecule (i.e., compared with a control, such as untreated cells or cells treated with an inert compound), and then correlating the effect of the molecule with the observed change. The screening may be done either because the candidate molecule is designed to have a pharmacological effect on adipocytes, or because the molecule is designed to have effects elsewhere but there is a need to determine if it has and unintended side effects.

Cytotoxicity can be determined in the first instance by the effect on cell viability, survival, morphology, and leakage of enzymes into the culture medium. More detailed analysis may be conducted to determine whether a test substance affects cell function without causing toxicity.

Alternatively, the cells can be used to assess changes in gene expression patterns caused by a potential drug candidate. In this embodiment, the changes in gene expression pattern from addition of the candidate drug can be compared with the gene expression pattern caused by
5 a control drug with a known effect on adipocytes.

Therefore, according to a further aspect, there is provided a method for drug screening (e.g., evaluating drug reactivity), comprising a step of using the adipocytes produced by the method as described herein. According to a further aspect of the invention, there is provided a method
10 of drug screening comprising contacting an adipocyte generated using the method as defined herein, or an adipocyte as defined herein, with the drug and observing a change in the adipocyte induced by the drug.

According to a further aspect of the invention, there is provided the adipocyte as defined herein
15 for use in therapy.

In one embodiment, the method additionally comprises transplanting the adipocytes into a patient. In this aspect of the invention, the cells used to generate the adipocytes may be autologous (i.e., mature cells removed, modified and returned to the same individual) or from
20 a donor (i.e., allogeneic, including a stem cell line). Direct reprogramming of cells into adipocytes is amenable to the production of autologous and allogeneic adipocytes.

The adipocytes can be used with a variety of materials to form a composition for purposes such as reconstructive surgery. The cells may be combined with a biomatrix to form a two
25 dimensional or three dimensional material as needed. In particular, adipocytes may be used to prepare fat pads and fatty tissues to build up an area where tissue has been removed.

In one embodiment, the therapy is a cosmetic treatment. Similar to their use in reconstructive surgery, adipocytes generated using methods of the invention may be of use in elective
30 cosmetic surgery in much the same way, i.e. to build up underlying tissue below the skin with a composite of autologous cells and biocompatible material.

Therefore, according to a further aspect of the invention there is provided a method of treating a subject having or at risk of a disease or disorder comprising administering to the subject a
35 therapeutically effective amount of adipocytes generated using the method as defined herein, or adipocytes as defined herein.

Adipocytes of the invention may find use in the development for therapies and the treatment of disorders of adipose development and function (e.g., lipodystrophy or obesity), as well as the secondary disorders of adipose dysfunction (e.g., diabetes, hyperlipidemia, hypertension or cardiovascular disease).

In a different aspect, the cells may be used in tissue engineering. Tissue engineering requires the generation of tissue which could be used to replace tissues or even whole organs of a human or animal. Methods of tissue engineering are known to those skilled in the art, but include the use of a scaffold (an extracellular matrix) upon which the cells are applied in order to generate tissues/organs. These methods can be used to generate an "artificial" tissue or organ. Methods of generating tissues may include additive manufacturing, otherwise known as three-dimensional (3D) printing, which can involve directly printing cells to make tissues. The present invention thus provides a method for generating tissues using the cells produced as described in any aspect of the invention.

Tissues generated using cells made according to the methods of the present invention may be used for *in vitro*/cultured meat. Therefore, according to one aspect of the invention, there is provided the adipocyte as defined herein for use in preparing cultured meat. According to a further aspect of the invention, there is provided a food product (e.g. cultured meat) comprising the adipocytes as defined herein. The primary cell type for cultured meat is myocytes, however, to recreate the texture and/or taste of natural meat, it is desirable to prepare cultured meat using a combination of cell types, including adipocytes (i.e. such as those of the present invention). If the aim of the engineered tissue is for cultured meat, then the source cell may be taken from a livestock animal.

References herein to "cultured meat" (which may also be referred to as *in vitro* meat, lab-grown meat, cell-based meat, cultivated meat or synthetic meat) is used to refer to meat grown from a cell or tissue culture as opposed to slaughtering animals to obtain the meat.

30

Reprogramming kits

According to a further aspect, there is provided a kit for differentiating a cell into an adipocyte comprising:

(i) a source cell and an agent that activates or increases the expression or amount of at least one or more transcription factors; and/or

(ii) one or more expression cassette(s) comprising nucleotide sequences encoding at least one or more transcription factors,

wherein the one or more transcription factors are selected from the group consisting of: a PPAR protein, HOXC8, EBF1, EBF2, ZNF467, ZNF423 and variants thereof.

5

According to a further aspect, there is provided a kit for differentiating a cell into an adipocyte comprising:

(i) a source cell, preferably a pluripotent stem cell, more preferably a hiPSC, and an agent that activates or increases the expression or amount of at least one or more transcription factors; and/or

10

(ii) one or more expression cassette(s) comprising nucleotide sequences encoding one or more polypeptides having transcription factor activity and/or at least one or more transcription factors,

wherein the one or more transcription factors are selected from the group consisting of: one or more PPAR proteins (such as PPARA and/or PPARG), HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins (such as CEBPA and/or CEBPB) and variants thereof.

15

In one embodiment, the expression cassette comprises an inducible expression construct comprising a sequence encoding one or more polypeptides having transcription factor activity and/or the transcription factors themselves.

20

As described herein, combinations of polypeptides having transcription factor activity and/or the transcription factors described herein are of particular use in the present invention. If a combination of polypeptides and/or transcription factors is required, these may be encoded on the same or on different expression cassettes. Therefore, in one embodiment, the kit comprises an expression cassette (preferably an inducible expression cassette) encoding two or more polypeptides and/or transcription factors, such as three, four, five, six, seven or eight transcription factors. Preferably, the kit comprises an expression cassette encoding three or more, more preferably four or more, polypeptides and/or transcription factors.

25

30

According to a further aspect, there is provided a use of a kit as defined herein, for differentiating a cell into an adipocyte.

35

The kit may include one or more articles and/or reagents for performance of the method. For example, one or more transcription factor genes, derivatives, variants or fragments thereof, for

use in the methods described herein may be provided in isolated form and may be part of a kit, e.g., in a suitable container such as a vial in which the contents are protected from the external environment.

- 5 In one embodiment, the kit additionally comprises at least one source cell, such as a pluripotent stem cell (such as an induced pluripotent stem cell) or a non-pluripotent, non-adipocyte cell.

In one embodiment, the kit additionally comprises a medium for culturing the cell and instructions for preparing the enhanced potency cells or reprogrammed pluripotent cells in
10 accordance with the method defined herein.

It will be understood that all embodiments described herein may be applied to all aspects of the invention.

- 15 Other features and advantages of the present invention will be apparent from the description provided herein. It should be understood, however, that the description and the specific examples while indicating preferred embodiments of the invention are given by way of illustration only, since various changes and modifications will become apparent to those skilled in the art. The invention will now be described using the following, non-limiting examples:

20

EXAMPLES

EXAMPLE 1 – Prioritisation screen

25 Methods:

- Plasmid library encoding the TFs, reporter line and transposase-mediated delivery

A large-scale library of transcription factor (TF) candidates was shortlisted from selected multi-omics datasets by scientific curation based on a tiered transcriptomics analysis of adipocyte-
30 labelled cells and applying a suite of computational tools

Each TF and a small set of negative controls were cloned separately into vectors for genomic integration mediated by PiggyBac transposase. TF expression was controlled by a tetracycline-inducible promoter. Vectors encoded a puromycin-resistance cassette for the
35 selection of cells with successful genomic integration. Each TF expression cassette encodes

a unique DNA barcode, thus allowing the identification of TFs and the quantification of their frequency in cell pools.

To identify cells successfully reprogrammed to adipocytes, an hiPSC reporter line was generated. The genomic loci of two markers for mature adipocytes, FABP4 and PLIN1, were engineered to generate polycistronic cassettes encoding the fluorescent proteins GFP and mCherry, respectively. The fluorescent markers are therefore expressed upon the activation of the endogenous markers, without disrupting their function. Since cells with high lipid content may not be suitable for flow sorting or other microfluidics-based protocols, both GFP and mCherry were tagged with a nuclear localization signal peptide which confined the expression of the fluorescent proteins to the nucleus. Therefore, end-point fluorescence-activated cell sorting could be carried out on either whole cell populations, or purified nuclei which maintained intact fluorescent signal.

Vectors were pooled in equimolar ratios to ensure even representation of TFs followed by retransformation in bacteria and large-scale DNA preparation. A mix of TF library and PiggyBac transposase was nucleofected into hiPSCs. Cell culture scale and nucleofection parameters were optimized to ensure an adequate coverage of the high-plexity combinatorial space.

• Screening outline

Two replicates of a prioritisation screening experiment were performed. Prior to nucleofection, FABP4-GFP/PLIN1-mCherry hiPSCs were expanded in TesR E8 (Stem Cell Technologies) on standard tissue culture plates coated with Vitronectin (Life Technologies). Following nucleofection, cells were cultured as above and selected for successful genomic integration by adding puromycin to the culture media. Following selection, hiPSCs were plated for reprogramming and cultured for up to 10 days in media promoting maintenance and survival of adipocytes, in the presence of doxycycline. At all stages, enough cells were maintained in culture to ensure adequate coverage for each combination of TFs. After reprogramming, cells were harvested and selected by fluorescence-activated cell sorting of GFP-positive, mCherry-positive and negative control populations. In a separate set of samples, cell nuclei were purified prior to sorting.

• Bulk TF barcode quantification and analysis

Genomic DNA was extracted from all marker-positive and control samples and the relative distribution of TF barcodes in each bulk gDNA sample was quantified by amplicon sequencing.

Data were analysed using the Mageck method (Li et al. *Genome Biol* 15, 554, 2014). TF barcode counts were analysed using the Mageck *count* function. Subsequently, the raw counts files were used in both the *test* and *maximum likelihood estimation* (MLE) functions to generate coefficients and statistics on the enrichment of each TF in FABP4-GFP and PLIN1-mCherry positive samples compared to a non-enriched control sample. The *MLE* function allows the comparison of multiple conditions and produces a beta score which is similar to the logFC value of a traditional differential expression test. This analysis considered reproducibility across the two replicates. The *test* function compares conditions pairwise and produces a logFC value and associated statistics for each pair.

Results:

The enrichment of TFs in the two marker positive populations (FABP4-GFP or PLIN1-mCherry) was measured by beta score and logFC as produced by the *MLE* and *test* functions respectively. A TF was considered a putative reprogramming factor, and selected for downstream validation (see Example 2 below) if either beta score > 0.3 or mean logFC > 0.5 in any of the following samples:

- Whole cells FABP4-GFP positive samples;
- Whole cells PLIN1-mCherry positive samples;
- Purified nuclei FABP4-GFP positive samples;
- Purified nuclei PLIN1-mCherry positive samples.

Table 2.

TF	Beta score				LogFC			
	Whole cells		Nuclei		Whole cells		Nuclei	
	FABP4-GFP	PLIN1-mCherry	FABP4-GFP	PLIN1-mCherry	FABP4-GFP	PLIN1-mCherry	FABP4-GFP	PLIN1-mCherry
ARID5B	-0.35	-0.26	-6.89	0.48	-0.48	-0.35	-10.11	0.64
CDKN1A	-0.49	0.35	1.16	0.80	0.23	1.45	1.97	1.49
CEBPB	-1.63	-0.42	-0.72	0.37	-2.44	-0.70	-1.04	0.52
CREB1	0.36	-0.16	-0.84	-0.41	0.54	-0.20	-1.16	-0.64
EBF1	-0.28	1.05	-5.45	1.44	-0.09	1.82	-7.85	2.16
EBF2	-0.57	0.93	-1.18	1.29	-0.45	1.71	-1.67	1.93
ESRRA	0.31	-0.03	-1.07	0.42	0.37	-0.11	-1.50	0.52
FOXA3	-0.74	0.25	0.18	0.54	-1.30	0.13	-0.03	0.69
FOXO1	-0.20	-0.38	0.30	0.52	0.52	0.27	0.55	0.72
HOXC10	0.42	-0.43	0.05	-0.79	0.52	-0.70	0.32	-1.31

HOXC8	0.63	0.19	0.09	-0.27	0.72	0.10	0.19	-0.44
KLF15	-1.56	-0.10	-0.74	1.36	-0.60	1.51	-0.91	2.08
NFIA	1.00	1.01	0.57	1.11	1.41	1.42	0.78	1.60
PPARA	0.97	0.21	0.92	0.00	1.29	0.19	1.48	0.11
PPARG	1.23	0.81	0.97	0.51	1.77	1.17	1.46	0.81
PPARG	0.78	0.70	0.65	0.52	1.13	1.02	1.03	0.82
RXRA	-0.03	0.36	-0.03	0.09	-0.14	0.41	-0.12	0.09
ZBTB7B	0.37	0.40	-0.94	-0.23	0.58	0.61	-1.34	-0.16
ZNF423	1.32	-0.78	1.47	-1.98	1.92	-1.12	1.93	-3.04
ZNF467	-0.22	0.21	0.91	1.20	0.52	1.14	1.39	1.93
ZNF516	-0.29	0.24	-5.25	0.87	-0.06	0.71	-7.72	1.13

Table 2 lists the 20 putative reprogramming TFs and indicates the beta score and logFC values for FABP4-GFP and PLIN1-mCherry positive samples. LogFC values are shown for both replicates, beta scores include experimental replicates when calculated. In bold are values that satisfy the conditions for downstream validation. Not shown are the TFs that failed the conditions for downstream validation.

EXAMPLE 2 – Validation Screen

10 Methods:

- Plasmid library encoding the TFs and Transposase-mediated delivery

Following the prioritisation screen (see Example 1) the TFs were assessed for reprogramming potential. Each of the TFs and negative controls were cloned separately into vectors for genomic integration mediated by PiggyBac transposase. TF expression was controlled by a tetracycline-inducible promoter. Vectors encoded a puromycin-resistance cassette for the selection of cells with successful genomic integration. Each TF expression cassette encodes a unique DNA barcode. We designed primers for the concurrent amplification of TF and cell barcodes in the 10x Genomics workflow, allowing us to assign TF combinations to single-cell transcriptome profiles.

Vectors were pooled in equimolar ratios to ensure even representation of TFs followed by retransformation in bacteria and large-scale DNA preparation. A mix of TF library and PiggyBac transposase was nucleofected into iPSCs. Cell culture scale and nucleofection parameters were optimized to ensure an adequate coverage of the high-plexity combinatorial space.

- Screening outline

Two replicates of a validation screening experiment were performed using the same pooled plasmid library. Prior to nucleofection, FABP4-GFP/PLIN1-mCherry hiPSCs were expanded in
5 TesR E8 (Stem Cell Technologies) on standard tissue culture plates coated with Vitronectin (Life Technologies). Following nucleofection, cells were cultured as above and selected for successful genomic integration by adding puromycin to the culture media. Following selection, iPSCs were plated for reprogramming and cultured for up to 9 days in media promoting maintenance and survival of adipocytes, in presence of doxycycline. At all stages, enough cells
10 were maintained in culture to ensure adequate coverage for each combination of TFs.

After reprogramming, cells were harvested and selected by fluorescence-activated cell sorting of GFP-positive, mCherry-positive and negative control populations. Non-sorted (NS), sorted-non-gated (NG) and marker-negative (or fully negative, FN) control samples were collected.
15

- Single-cell transcriptome analysis and TF barcode capture

Nuclei of sorted cells were purified and analysed by scRNA-seq using 10x Genomics Chromium Single Cell 3' Reagent Kits v3 following manufacturer's instructions. Up to 10,000 nuclei per sample were targeted. After the cDNA amplification step, gene expression libraries
20 were created and sequenced on NovaSeq aiming for at least 25,000 reads per cell, as per the 10x Genomics 3' v3.1 protocol. In addition to the gene expression libraries, the matched cDNA was used as a template for further targeted amplification of the TF barcodes, which were sequenced on MiSeq.

- Data analysis

Single-cell gene expression data generated by the Cell Ranger pipeline were further analysed in Seurat v3 (Stuart et al., Cell, 2019) and visualized on uniform manifold approximation and projection (UMAP) plots (Becht, Nature Biotech, 2019). Cell identities were assigned using CellTypist (Dominguez Conde et al., Science, 2022), an automated cell classification tool, and
30 a curated reference single cell atlas of human adipose tissue (Emont et al., Nature, 2022). 10x cell barcodes were used to assign exogenous transcription factor (eTF) barcodes to single cells and their corresponding gene expression profiles. eTF enrichment was quantified by comparing cells in negative control (NG) and sorted (FABP4-GFP, PLIN1-mCherry) samples, as well as cells within the marker-positive populations identified as Adipocytes by CellTypist.

35

Results:

Single-cell gene expression data were visualized on UMAP plots including undifferentiated hiPSCs (G10), sorted live cells (NG), FABP4-GFP positive cells, PLIN1-mCherry positive cells as well as marker-negative cells (full negative, or FN; Figure 1). Predictably a cluster composed exclusively by undifferentiated hiPSCs is separate from all the other clusters that include the cells collected at the end of the reprogramming (NG, FN, FABP4-GFP, PLIN1-mCherry). Within the latter, a large cluster includes almost exclusively non-sorted cells, while more scattered clusters include cells from both FABP4-GFP positive and PLIN1-mCherry positive samples. The partial overlap between FABP4-GFP and PLIN1-mCherry positive cells suggests that both markers enrich for cells that share similar gene expression profiles. A small fraction low-quality nuclei is clustered at the right end of the UMAP.

eTF barcodes were successfully detected by 10x and assigned to single cells in all samples (Figure 2). While negative controls (e.g. NEG1 to NEG5) are distributed randomly across the samples, several eTFs are found in clusters within the sorted cell population (e.g. PPARA, PPARG, EBF1, EBF2).

Next, cell identities were assigned by using CellTypist against a reference dataset of gene expression profiles of adipocytes (Emont et al., Nature, 2022). Within our dataset, we identified with high confidence cells belonging to at least 5 cell types, including adipocytes (Figure 3). In total 4,670 cells were identified as adipocytes, most of them in FABP4-GFP and PLIN1-mCherry positive samples (Table 3 below).

Table 3. Counts of cells identified as Adipocytes, per sample.

Sample	Adipocytes
ADlv_R10_FABP4	285
ADlv_R10_PLIN1	1641
ADlv_R10_FN	45
ADlv_R10_NG	47
ADlv_R10_NS	65
ADlv_R11_FABP4	650
ADlv_R11_PLIN1	1854
ADlv_R11_NG	49
ADlv_R11_NS	34

The expression of FABP4-GFP and PLIN1-mCherry defines two distinct, partially overlapping adipocyte clusters, suggesting that each marker identifies a specific subpopulation of adipocytes. We assessed whether these hiPSC-derived adipocytes expressed any functional genes associated with this cell type. For both markers and across all experimental replicates, cells show high expression levels of several genes commonly associated with adipose tissue (Figure 4). Conversely, no markers for cardiac cells were found. Of the 5 cell types identified by Celltypist, adipocytes showed the highest expression levels of markers commonly associated to brown adipose (Figure 5) and white adipose tissue (Figure 6).

To identify eTFs reprogramming hiPSCs to adipocytes cells, we first calculated the log-likelihood of finding sets of up to 4 eTFs in FABP4-GFP and PLIN1-mCherry positive cells identified as adipocytes by Celltypist, compared to control, and ranked the sets based on the enrichment. As shown in Figure 7, several enriched sets were found with a small number showing particularly high enrichment. Table 4 shows the most enriched combinations for FABP4-GFP and PLIN1-mCherry positive adipocytes. Not surprisingly, the two data sets identify partially overlapping eTF sets, which drive the reprogramming of distinct adipocyte subtypes. The eTF combination including EBF2, EBF1, PPARA and/or HOXC8 is identified as reprogramming in the PLIN1-mCherry positive dataset. The eTF combination including PPARG, PPARA, HOXC8 and/or ZNF467, ZNF423 is identified as reprogramming in the FABP4-GFP positive dataset.

Table 4. The most enriched reprogramming combinations for each adipocyte subpopulation

Comparison	Rank	Combination
PLIN1-mCherry adipocytes	1	PPARA, EBF2, EBF1
	2	HOXC8, EBF2, EBF1
FABP4-GFP adipocytes	1	PPARG, ZNF467, PPARA
	2	PPARG, HOXC8, PPARA
	3	HOXC8, ZNF423, PPARA

EXAMPLE 3 – Further review of TF combinations

Following the Validation screen (see Example 2) some of the top scoring TFs (PPARG, PPARA, HOXC8, ZNF423, EBF1 and EBF2) were assessed in combination for reprogramming potential.

Methods:

Each of the TFs were cloned separately into vectors for genomic integration mediated by PiggyBac transposase. TF expression was controlled by a tetracycline-inducible promoter. Vectors encoded a puromycin-resistance cassette for the selection of cells with successful
5 genomic integration. Vectors were pooled in equimolar ratios to ensure even representation of TFs. A mix of TF pool and PiggyBac transposase was nucleofected into hiPSCs. Cell culture scale and nucleofection parameters were optimized to ensure an adequate fraction of cells encoded the full set of 6 TFs.

- 10 Two nucleofection replicates were performed using the same plasmid pool. Prior to nucleofection hiPSCs were expanded in GIBCO ESSENTIAL 8 Medium (Gibco) on standard tissue culture plates coated with Vitronectin (Life Technologies). Following nucleofection, cells were cultured as above and selected for successful genomic integration by adding puromycin to the culture media. Following selection, hiPSCs were plated for reprogramming and cultured
15 for 10 days in media promoting maintenance and survival of adipocytes, in the presence of doxycycline. TF delivery by transposase may result in a variable number of insertion events, therefore biasing the analysis of reprogramming efficiency.

- After reprogramming, cells were harvested and processed for RNA extraction, and reverse
20 transcription and quantitative PCR was performed for adipocyte markers and the housekeeping control gene hydroxymethylbilane synthase (HMBS). Alternatively, cells were fixed and immunostained for adipocyte markers, and intracellular lipid accumulation was assessed by LipidTox™ (*Thermo Fisher Scientific*) staining.

Results:

Analysis of transcript levels by qPCR showed the expression of key adipocyte markers PLIN1, FABP4, CEBPB, CEBPA, CD36, and UCP1 enriched in cells reprogrammed with the pool of 6TFs (Figure 8). Of these adipocyte markers, UCP1 is the key regulator of brown adipocyte fate.

- 30 To assess reprogramming efficiency, immunocytochemistry (ICC) for markers FABP4 and PLIN1 was performed. Protein expression of both markers was readily detectable (Figure 9). Furthermore, intracellular lipid accumulation was detected by LipidTox staining.

- 35 Overall these data show that one or more combinations of TFs including PPARG, PPARA, HOXC8, ZNF423, EBF1 and EBF2 can drive reprogramming of hiPSCs into adipocytes.

CLAIMS

1. A method of generating adipocytes comprising expressing one or more polypeptides having the activity of one or more transcription factors and/or increasing the expression of one or more transcription factors, the transcription factors selected from the group consisting of: one or more PPAR proteins such as PPARA and/ PPARG, HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins such as CEBPA and/or CEBPB and variants thereof, in a cell population and culturing the cell population to obtain adipocytes.
2. The method of claim 1, wherein expression of two or more transcription factors selected from the group consisting of: a PPAR protein, HOXC8, EBF1, EBF2, ZNF467, ZNF423, is increased.
3. The method as defined in claim 1 or claim 2, wherein the PPAR is selected from the group consisting of: PPARA, PPARG and PPARA in combination with PPARG.
4. The method as defined in any one of claims 1 to 3, wherein the transcription factors are selected from the group consisting of: PPARA, EBF1, EBF2, HOXC8 and variants thereof.
5. The method as defined in any one of claims 1 to 3, wherein the transcription factors are selected from the group consisting of: PPARA, PPARG, HOXC8, ZNF467, ZNF423 and variants thereof.
6. The method as defined in claim 1 or claim 2, wherein the transcription factors comprise one of the following combinations:
 - (i) PPARA, EBF2 and EBF1;
 - (ii) HOXC8, EBF2 and EBF1;
 - (iii) PPARG, ZNF467 and PPARA;
 - (iv) PPARG, HOXC8 and PPARA;
 - (v) HOXC8, ZNF423 and PPARA;
 - (vi) PPARA and CEBPA;
 - (vii) PPARA and CEBPB;
 - (viii) PPARG and CEBPA; or
 - (ix) PPARG and CEBPB.

7. The method as defined in any one of claims 1 to 6, which comprises increasing the expression of one or more additional transcription factors as listed in **Table 1**.
8. The method as defined in any one of claims 1 to 7, wherein the method comprises increasing the expression of between three and seven transcription factors.
9. The method as defined in any one of claims 1 to 8, wherein the transcription and translation of the one or more polypeptides having the activity of one or more transcription factors and/or the transcription factors themselves are controlled within the cell.
10. The method as defined in claim 9, wherein the transcription and translation of the transcription factors is controlled within the cell.
11. The method as defined in any one of claims 1 to 10, wherein the cell population comprises pluripotent stem cells, in particular induced pluripotent stem cells.
12. The method as defined in any one of claims 1 to 11, wherein the method comprises generating adipocytes by cellular reprogramming of pluripotent stem cells or induced pluripotent stem cells, preferably human induced pluripotent stem cells.
13. The method as defined in any one of claims 1 to 12, wherein the adipocytes are human adipocytes.
14. The method as defined in any one of claims 1 to 13, which additionally comprises monitoring the cell population for at least one characteristic of an adipocyte.
15. The method as defined in claim 14, wherein the characteristic is selected from one or more of:
- (i) expression of one or more cell markers, such as FABP4, PLIN1, CEBPA, CEBPB, CD36, UCP1 or a combination thereof;
 - (ii) adipokine expression or response to insulin; and
 - (iii) adipocyte morphological features.
16. The method as defined in any one of claims 1 to 15, wherein the expression of the transcription factors is increased by contacting the cell population with one or more exogenous

expression cassettes encoding one or more of the genes, or one or more agents that activate or increase the expression or amount of the transcription factors.

17. The method as defined in any one of claims 1 to 16, wherein expression of the genes is under controlled transcription.

18. The method as defined in any one of claims 1 to 17, wherein a sequence encoding one or more of the transcription factors is introduced into the cell population using a method comprising:

- insertion (preferably targeted insertion) of a coding sequence for a transcriptional regulator protein into a first genomic safe harbour site of a source cell present in the cell population; and

- insertion (preferably targeted insertion) of an inducible cassette into a second genomic safe harbour site of the source cell, wherein said inducible cassette comprises said sequence encoding one or more transcription factors operably linked to an inducible promoter, and said promoter is regulated by the transcriptional regulator protein.

19. The method as defined in any one of claims 1 to 18, which comprises culturing under suitable conditions for at least 4 days, such as at least 7 days, in particular about 10 days.

20. The method as defined in claim 19, which comprises culturing in media comprising one or more components selected from the group consisting of: BMP4, Activin A, FGF2, Insulin, Ascorbic acid and Dexamethasone.

21. A method for the production of adipocytes from a source cell, preferably a pluripotent stem cell, more preferably a human induced pluripotent stem cell comprising the steps of:

- a) insertion (preferably targeted insertion) of a gene encoding a transcriptional regulator protein into a first genomic safe harbour site of the source cell; and

- b) insertion (preferably targeted insertion) of at least one nucleotide sequence encoding one or more polypeptides having the activity of one or more transcription factors and/or encoding one or more transcription factors, the transcription factors selected from the group consisting of: one or more PPAR proteins, such as PPARA and/or PPARG, HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins such as CEBPA and/or CEBPB and variants thereof, operably linked to an inducible promoter into a second genomic safe harbour site of the source cell, wherein said inducible promoter is regulated by the transcriptional regulator protein; and

c) culturing the source cell(s) comprising the insertions to obtain adipocytes.

22. Use of one or more transcription factors, wherein the one or more transcription factors are selected from the group consisting of: one or more PPAR proteins such as PPARA and/or PPARG, HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins such as CEBPA and/or CEBPB and variants thereof, to generate adipocytes.

23. A cell obtainable by any one of the methods defined in claims 1 to 22.

24. A cell, preferably a pluripotent stem cell, more preferably a human induced pluripotent stem cell, comprising one or more exogenous expression cassettes comprising nucleotide sequences encoding one or more polypeptides having the activity of one or more transcription factors and/or encoding at least one or more transcription factors, wherein the one or more transcription factors are selected from the group consisting of: one or more PPAR proteins such as PPARA and/or PPARG, HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins, such as CEBPA and/or CEBPB and variants thereof.

25. The cell as defined in claim 23 or claim 24, wherein the nucleotide sequences encoding the one or more transcription factors are integrated into the genome of the cell.

26. The cell as defined in claim 25, wherein the nucleotide sequences encoding the one or more transcription factors are integrated into a target site of the cell.

27. A cell as defined in any one of claims 23 to 26, for use in therapy.

28. A cell as defined in any one of claims 23 to 26, for use in *in vitro* diagnostics or drug screening.

29. A cell as defined in any one of claims 23 to 26, for use in preparing cultured meat.

30. A kit for differentiating a cell, preferably a pluripotent stem cell, more preferably a human induced pluripotent stem cell, into an adipocyte comprising:

(i) a source cell and an agent that activates or increases the expression or amount of at least one or more transcription factors; and/or

(ii) one or more expression cassette(s) comprising nucleotide sequences encoding one or more polypeptides having the activity of one or more transcription factors and/or encoding at least one or more transcription factors,

wherein the one or more transcription factors are selected from the group consisting of: one or more PPAR proteins such as PPARG and/or PPARG, HOXC8, EBF1, EBF2, ZNF467, ZNF423, one or more CEB proteins such as CEBPA and/or CEBPB and variants thereof.

31. Use of a kit as defined in claim 30, for differentiating a cell, preferably a pluripotent stem cell, more preferably a human induced pluripotent stem cell, into an adipocyte.

32. A method of drug screening comprising contacting an adipocyte generated using the method as defined in any one of claims 1 to 21, or an adipocyte as defined in any one of claims 23 to 26, with a drug and observing a change in the adipocyte induced by the drug.

33. A method of treating a subject having or at risk of a disease or disorder comprising administering to the subject a therapeutically effective amount of adipocytes generated using the method as defined in any one of claims 1 to 21, or adipocytes as defined in any one of claims 23 to 26.