An investigation of the early molecular regulation of adipogenesis in human mesenchymal stem cells.

Claire Westwood

Thesis submitted to the University of London for the Degree of

Doctor of Philosophy

2006

Wolfson Institute for Biomedical Research

University College London

UMI Number: U592458

UMI

Dissertation Publishing

ProQuest

ABSTRACT

Human bone marrow-derived mesenchymal stem cells (hMSCs) can differentiate *in vitro* into various lineages including adipocytes, osteoblasts, and chondrocytes. Determining the earliest events of hMSC differentiation would provide a greater understanding of stem cell function, leading to novel therapeutic strategies. The molecular regulation of adipogenesis has been studied extensively using preadipocyte models that are committed to the adipocyte lineage. However, the events that occur in uncommitted stem cells such as hMSCs, and thus control the early stages of differentiation, have not been identified. In this thesis, hMSC differentiation to adipocytes was induced using a classic induction cocktail, and transcriptional profiles were obtained at early timepoints using Affymetrix U133 Plus 2.0 microarrays. Adipogenesis was shown to involve the regulation of over 4200 genes, which were clustered into distinct temporal patterns of gene expression. Recognised adipogenic regulators were expressed at the later stages of adipogenesis alongside genes involved in lipid metabolism, thus demonstrating the validity of the system. Global annotation of genes showing early expression changes revealed temporal regulation of signaling pathways known to control adipogenesis, together with molecular processes and pathway regulators not previously implicated in adipogenesis. A further detailed analysis of gene function identified 83 potential candidates for novel adipogenic regulators.

Knock-down of gene expression induced by virally-mediated shRNA is an effective method to assess gene function. To investigate whether the candidate genes identified from the transcriptional profiling played a role in hMSC adipogenesis, retroviral shRNA library screens and an additional lentiviral shRNA strategy were developed and implemented.

Finally, the effect on hMSC differentiation of the introduction of oncogenic regulators was investigated. Oncogenic transformation of hMSCs may lead to aberrant differentiation and mesenchymal tumour formation. It was found that the progressive addition of transforming genes did not block adipogenic or osteogenic differentiation, but caused significant impairment to the process.

2

CONTENTS

LIST OF FIGURES

9

construct and resulting viral copy number per cell.

LIST OF TABLES

LIST OF ABBREVIATIONS

The abbreviations (other than human gene names, see following list) used in this thesis are:

Abbreviation	Full name
293gp	Human transformed embryonal kidney cell line 293 containing copies of retroviral *gag* and *pol* genes stably integrated into the genome
ADSC	Adipose-derived stem cell
ALT	Alternative lengthening of telomeres
AML	Acute myeloid leukaemia
Amp	Ampicillin
ANOVA	Analysis of variance
APC	Allophycocyanin
BAT	Brown adipose tissue
BHA	Butylated hydroxyanisole
BrdUrd	Bromodeoxyuridine
BSA	Bovine serum albumin
cAMP	Cyclic adenosine monophosphate
CFU-F	Colony forming units-fibroblastic
CLP	Common lymphoid progenitor
CMP	Common myeloid progenitor
CMV	Cytomegalovirus
cPPT	Central copy of the polypurine tract
CSGW	HIV-1-based, self-inactivating lentiviral vector containing cPPT, SFFV promoter, eGFP and WPRE.
Cy3/5/7	Cyanine 3/5/7
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethylsulfoxide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ESC	Embryonic stem cell
FACS	Fluorescence-assisted cell sorting

FCFC	Fibroblast colony forming cells
FCS	Fetal calf serum
FDR	False discovery rate
FISH	Fluorescence in situ hybridisation
FITC	Fluorescein isothiocyanate
GFP	Green fluorescent protein
GMP	Granulocytic/monocytic-restricted progenitors
GO	Gene Ontology
GOI	Gene of interest
GTP	Guanosine triphosphate
HEK-293	Human embryonic kidney cell line
HEPES	4-2-hydroxyethyl-1-piperazineethanesulfonic acid
HFSC	Hair follicle stem cell
HPV	Human papilloma virus
HSC	Haematopoietic stem cell
IBMX	Isobutylmethylxanthine
KEGG	Kyoto Encyclopedia of Genes and Genomes
KSHV	Kaposi's sarcoma-associated herpesvirus
LB	Luria-Bertoni broth
LIMMA	Linear models for microarray analysis
LiSa2	Human liposarcoma-derived cell line
LRC	Label-retaining cell
LTHSC	Long term haematopoietic stem cell
LTR	Long terminal repeat
MACS	Magnetic-assisted cell sorting
MAPC	Multipotent adult progenitor cell
MAS5	Affymetrix microarray suite version 5.0
MBEI	Model-based expression index
MEF	Mouse embryonic fibroblast
MEP	Megakaryocytic/erythroid-restricted progenitors
MM	Mismatch
MPP	Multipotent progenitor

(h)MSC	(Human) mesenchymal stem cell
MSCV	Murine stem cell virus
NOD/SCID	Nonobese diabetic/severe combined immunodeficient
OMIM	Online Mendelian Inheritance in Man
PBS	Phosphate buffered saline
PCC	Pearson's correlation coefficient
PE	Phycoerythrin
PM	Perfect match
Q-PCR	Quantitative polymerase chain reaction
RMA (GC-RMA)	Robust multiarray analysis (guanine/cytosine content of probe)
RNAi	RNA interference
RISC	RNA-induced silencing complex
RRE	Rev response element
RSV	Rous sarcoma virus
SAP	Shrimp alkaline phosphatase
SCID	Severe combined immunodeficient
SFFV	Spleen focus-forming virus
shRNA	Small hairpin RNA
siRNA	Small interfering RNA
SL-IC	SCID-leukaemia-inducing cell
STHSC	Short term haematopoietic stem cell
SV40 ST/LT	Simian virus 40 small/large T antigen
UBSC	Umbillical cord blood stem cell
VSVG	Vesicular stomatitis virus glycoprotein
WAT	White adipose tissue
WebGestalt	WEB-based GEne SeT AnaLysis Toolkit
WPRE	Woodchuck hepatitis virus response element
β-ME	β-mercaptoethanol

The abbreviated human gene names used in this thesis are:

Abbreviation	Full name
ACTA(1/2)	Actin, alpha 1/2, smooth muscle, aorta
ACVR1	Activin A receptor, type 1
ADNP	Activity-dependent neuroprotector
ALCAM1	Activated leukocyte cell adhesion molecule
ALP	Alkaline phosphatase
AML1-ETO	Acute myeloid leukemia 1-eight twenty one protein fusion protein
aP2	Adipocyte protein 2
BMP(R)	Bone morphogenetic protein (receptor)
bFGF(R)	Basic fibroblast growth factor (receptor)
C/EBP	CCAAT/enhancer binding protein
CBFA1	Core binding factor alpha 1
CDK	Cyclin-dependent kinase
CHOP10	C/EBP homologous protein 10
CLP1	Cartilage linking protein 1
COL1A1	Collagen, alpha 1 type I
CREB	cAMP responsive element binding protein
Cyc	Cyclin
CYLD	Cylindromatosis (turban tumor syndrome)
DACT1	Dapper homolog 1 (antagonist of β-catenin)
DCN	Decorin
DKK	Dickkopf homolog (Xenopus laevis)
DLK1	Delta-like 1 homolog (Drosophila)
DLX2	Distal-less homeobox 2
DNCH1	Dynein, cytoplasmic 1, heavy chain 1
DP	Differentiation regulated transcription factor protein
DTR	Diphtheria toxin receptor
DVL	Dishevelled homolog
EGFR	Epidermal growth factor receptor
ENC1	Ectodermal-neural cortex (with BTB-like domain)

15

ERK	Extracellular signal-regulated protein kinase
ETO/MTG8	eight twenty one protein/myeloid translocation gene on 8q22
EWS	Ewing sarcoma breakpoint region 1 protein
FASN	Fatty acid synthase
FOXC2	Forkhead box C2
FOXO1A	Forkhead box O1A (rhabdomyosarcoma)
FUS	Fusion (involved in t(12;16) in malignant liposarcoma)
FZD	Frizzled receptor
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GAS1	Growth arrest-specific 1
GATA	GATA sequence-binding protein
GBP	GSK3β binding protein
G-CSF (R)	Colony stimulating factor 3 (granulocyte) (receptor)
GFAP	Glial fibrillary acidic protein
GILZ	Glucocorticoid-induced leucine zipper
GM-CSF	Colony stimulating factor 2 (granulocyte-macrophage)
GSK3β	Glycogen synthase kinase 3 beta
HES1	Hairy and enhancer of split 1
HOX	Homeobox
H-RAS	V-Ha-ras Harvey rat sarcoma viral oncogene homolog
ICAM(1/2/3)	Intercellular adhesion molecule 1/2/3
ID 2/3	Inhibitor of differentiation/DNA binding 2/3
IFNγR	Interferon gamma receptor
IGF1 (R)	Insulin-like growth factor
IL (R)	Interleukin (receptor)
INHBB	Inhibin beta B
INK4A/ARF	Cyclin-dependent kinase inhibitor 2A
INSIG1	Insulin-induced gene 1
IPR	Prostacyclin receptor
KLF	Krüppel-like factor
K-RAS	Kirsten rat sarcoma viral oncogene homolog
KROX20	Early growth response 2

LANA	Latent nuclear antigen
LFA3	Lymphocyte function-associated antigen 3
LIF (R)	Leukemia inhibitory factor (receptor)
LPL	Lipoprotein lipase
LRP	Low density lipoprotein receptor-related protein
MAPK	Mitogen-activated protein kinase
M-CSF	Colony stimulating factor 1 (macrophage)
MDM2	Mouse double minute 2 homolog
MEK	MAPK/ERK kinase
MEOX2	Mesenchyme homeobox 2
MHC class I/II	Major histocompatibility complex, class I/II
MYC	V-myc myelocytomatosis viral oncogene homolog (avian)
NEDD9	Neural precursor cell expressed, developmentally downregulated, 9
NFkB	Nuclear factor kappa-B
OCT4	POU domain, class 5, transcription factor 1
PAI1	Plasminogen activator inhibitor, type I
PDE4D	Phosphodiesterase 4D, cAMP-specific
PDGFR	Platelet-derived growth factor receptor, beta polypeptide
PDK1/4	Pyruvate dehydrogenase kinase 1/4
PECAM1	Platelet/endothelial cell adhesion molecule
PGK	Phosphoglycerate kinase
PI3K	Phosphoinositide-3-kinase
PITX1	Paired-like homeodomain transcription factor 1
PKA	Protein kinase A
PKB/AKT	Protein kinase B/v-akt murine thymoma viral oncogene homolog 1
PLD1	Phospholipase D1
PLZF	Promyelocytic leukemia zinc finger protein
PP2A	Protein phosphatase 2A
PPARγ	Peroxisome-proliferator activated receptor γ
PPP2R1B	Protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), beta isoform
PTPRC	Protein tyrosine phosphatase, receptor type, C

RAF	V-raf murine leukemia viral oncogene homolog
RAL-GEF	RAL guanine-nucleotide exchange factors
RASD1	RAS, dexamethasone-induced 1
RB	Retinoblastoma protein
REST	RE1-silencing transcription factor
Rev-Erbα	Nuclear receptor subfamily 1, group D, member 1
RORγ	Retinoid-related orphan receptor γ
RXRα	Retinoid X receptor α
SCF (R)	Stem cell factor (receptor)
sFRP	Soluble frizzled-related protein
SGK	Serum/glucocorticoid regulated kinase
SIX2	Sine oculis homeobox homolog 2 (Drosophila)
SMAD	Mothers against decapentaplegic
SMARCF1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily F, member 1
SMURF2	SMAD specific E3 ubiquitin protein ligase 2
SNAI1	Snail homolog 1
SNAI2	Snail homolog 2 (slug)
SOCS3	Suppressor of cytokine signalling 3
SOD2	Superoxide dismutase 2
SREBP1c	Sterol response element binding protein 1c
SSEA1/3/4	Stage-specific embryonic antigen 1/3/4
STAT	Signal transducer and activator of transcription
TAZ	Transcriptional coactivator with PDZ-binding motif
TBL1X	Transducin (beta)-like 1X-linked
TBX2	T-box 2
TCF4/ TCF7L1/2	Transcription factor 4/7-like 1/2
TCF8	Transcription factor 8 (represses interleukin 2 expression)
TERT	Telomerase reverse transcriptase
TGFBR(1/2/3)	Transforming growth factor β receptor, type 1/2/3
TGFβ	Transforming growth factor β

TGIF	TGFβ-induced factor (TALE family homeobox)
TIP	Tension induced/inhibited protein
TNFα (R1/2)	Tumor necrosis factor alpha (receptor, type 1/2)
TOB1	Transducer of ERBB2, 1
TRAIL	Tumor necrosis factor (ligand) superfamily, member 10
UCP1	Uncoupling protein 1
VCAM1	Vascular cell adhesion molecule 1
vWF	von Willebrand factor
Wnt	Wingless-type MMTV integration site family

ACKNOWLEDGEMENTS

CHAPTER 1. INTRODUCTION.

1.1 Introduction to stem cells

Stem cells can broadly be described as possessing two major characteristics: an extensive self-renewal capacity and the ability to form multiple mature cell types through differentiation (Hall and Watt 1989, Joseph and Morrison 2005). Other stem cell characteristics include a quiescent nature *in vivo*, the ability to undergo asymmetric division, and the ability to regenerate tissue after injury, but such properties may not be possessed by all stem cell populations, and may vary over time within a single population (Hall and Watt 1989, Potten and Loeffler 1990, Morrison 1997).

Stem cells offer great promise in advancing the field of regenerative medicine. Their ability to proliferate extensively and differentiate into several, or many, cell types (depending on the stem cell in question) may be exploited in the laboratory in order to derive healthy, replacement tissue in the instance of injury or disease. Further to this, and perhaps more fundamentally, stem cells also provide a unique opportunity to study the processes of early development and cellular differentiation.

The isolation of stem cells from embryos as well as adult tissues demonstrates their existence throughout the development of an organism, and as such, stem cells have been classified in terms of their developmental potential. Totipotency defines the ability to form all extra-embryonic and embryonic tissues, and is a property possessed by the fertilised egg (Sell 2004). Pluripotent stem cells are able to form cells of the three embryonic germ layers, namely ectoderm (giving rise to skin and nervous system), endoderm (giving rise to cells of the gastrointestinal tract and internal organs such as liver), and mesoderm (giving rise to blood, blood vessels and connective tissue such as bone); this is the case for the inner cell mass of the blastocyst, embryonic stem cells, embryonic carcinoma cells and embryonic germ cells. Stem cell populations that give rise to several cell lineages, generally within one germ layer, have been defined by terms such as multipotent and oligopotent; adult stem cells are classically described in this manner, although this has been challenged in recent years (Section 1.3). Finally, multipotent stem cells can give rise to unipotent progenitors of multiple cell types, which ultimately form all mature tissues of the organism (Sell 2004).

Several methods are uniformly used to define stem cell populations, including the expression of certain cell surface markers, and the demonstration of specific functional properties. *In vivo*, the existence of a stem cell can be demonstrated via the reconstitution of a specific tissue type by a single cell after transplantation into an animal model, including production of further stem cells that can be re-transplanted and exhibit the same properties. *In vitro*, a definitive test of stem cell nature is the ability of a clonal cell population to produce differentiated cells of multiple types.

1.2 Embryonic stem cells

Embryonic stem cells (ESCs) have been isolated through the culture of pre-implantation blastocysts from mice (Evans and Kaufman 1981, Martin 1981) and humans (Thomson *et al* 1998). ESCs have been defined in terms of the expression of specific markers, and also their differentiation potential. Mouse ES cells (mESCs) express cell surface markers including SSEA-1; in striking contrast, human ES cells (hESCs) do not express this antigen, but express SSEA-3, SSEA-4, and the TRA-1-60 and TRA-1-81 antigens (Thomson *et al* 1998). Both ESC populations express the transcription factor Oct4, which acts to maintain their undifferentiated state (Nichols *et al* 1998). In support of this, siRNA-mediated knockdown (1.7.3) of Oct4 in hESCs results in their differentiation towards trophectoderm (Matin *et al* 2004).

ESCs are pluripotent: when mouse ESCs are injected into mouse embryos at the blastocyst stage, ESC contribution to all embryonic germ layers as well as the germ line is observed (Bradley *et al* 1984). Further, teratomas formed from hESCs transplanted into immunodeficient mice contain cells of all germ layers (Thomson *et al* 1998). Both mouse and hESCs can be maintained indefinitely in culture, and both exhibit *in vitro* pluripotency: they can spontaneously differentiate into cells of ectodermal, endodermal and mesodermal origin (Reubinoff *et al* 2000, Evans and Kaufman 1981), a process that can be initiated through the formation of embryoid bodies (which resemble the early embryo but lack polarity). Enrichment of specific cell types from differentiating hESCs has been achieved, including haematopoietic progenitors, neurons and hepatocyte-like

cells. However, hESC differentiation is not specific to a single cell lineage, but generally produces a mixture of cell types (Draper *et al* 2004a, Hoffman and Carpenter 2005).

Culturing primary embryo cells on replication-deficient cell feeder layers in the presence of FCS permitted the initial isolation of ESCs due to the presence of factors (either in the serum or secreted from the feeder cells) that supported ESC self-renewal and maintenance of pluripotency (Evans and Koufman 1981, Thomson *et al* 1998). Recently, some of these factors have been identified and can be used to culture ESCs in more defined conditions. In feeder-free conditions, the presence of the cytokine LIF is sufficient to maintain the self-renewal of mESCs (Williams *et al* 1988, Smith *et al* 1988). The same effect can also be achieved in serum-free conditions using LIF in combination with BMP proteins (Ying *et al* 2003). In contrast to this, LIF is not capable of maintaining hESCs in their undifferentiated state (Daheron *et al* 2004). The combination of the BMP inhibitor noggin and high levels of bFGF maintains hESC self-renewal and differentiation potential (Xu *et al* 2005). In that study, an incompletely defined serum replacement was used, but recent work has described the isolation of two hESC lines in entirely defined conditions; that is, using proteins of purely human origin as a replacement for animal serum and matrix components (required for cell adherence) (Ludwig *et al* 2006). This is a crucial step for the use of hESCs in stem cell therapies.

Although ESCs offer great promise in deriving multiple cell types for use in stem cell therapies, several complicating issues exist. As described above, hESC cells form teratomas after injection into immunodeficient mice in their undifferentiated state, necessitating their terminal differentiation in culture, and removal of any remaining undifferentiated hESCs, prior to use in therapy. However, hESC *in vitro* differentiation is currently inefficient, and does not yield homogeneous populations of a specific cell type. Furthermore, it has been found that prolonged culture of hESCs can lead to culture adaptation - the selective outgrowth of cells with enhanced self-renewal capability due to karyotypic changes that occurred as a result of culturing methods employed (Draper *et al* 2004a). This often involves the gain of chromosomes 17q and 12p, which are commonly seen in embryonal carcinoma cells (the tumourigenic counterparts of hESCs that are derived from teratocarcinomas), meaning that caution must be taken with the conditions

used to culture hESCs (Draper *et al* 2004b). Lastly, the use of human embryos for research or therapeutic purposes is highly controversial, and has prompted many ethical debates, as the embryos are ultimately destroyed by the process. Therefore, it is also of great importance to focus parallel research on adult stem cells, which are less surrounded by the ethical and scientific issues described here.

1.3 Adult stem cells

Stem cell populations have been identified within the adult organism that are capable of giving rise to multiple cell types within a specific tissue. Table 1.1 summarises a representative selection of well characterised adult stem cell populations. Major goals in adult stem cell research are the isolation of pure stem cell populations, and the identification of the molecular mechanisms underlying their stem cell properties. As well as producing tissue-specific cell types, some studies have raised the possibility that adult stem cells may be more plastic than originally thought, and may be able to differentiate to lineages from more than one germ layer (Table 1.1). However, *in vivo* (or *in vitro* co-culture) investigations of this property are hampered by the issue of whether such observations are caused by cell fusion (stem cells fusing with mature cells of a specific type, thereby adopting their characteristics) or trans-differentiation (the true ability of stem cells to differentiate to a lineage outside their anticipated range without cell fusion). Haematopoietic and skin stem cells are leading examples of adult stem cell populations, and the subject of this thesis focuses on adult bone marrow-derived mesenchymal stem cells.

Adult stem cell population	Lineages	Reference
Haematopoietic	All blood lineages. Hepatocytes? Muscle? Neural?	Morrison and Weissman 1994
Hair follicle	Hair follicle lineages and surrounding epithelia	Levy *et al* 2005
Intestinal epithelial	All intestinal epithelial lineages	Marshman *et al* 2002
Neural	Neurons, astrocytes, oligodendrocytes, blood? Muscle?	Morrison *et al* 1999, Uchida *et al* 2000
Skeletal muscle satellite	Skeletal muscle cells, blood?	Jackson *et al* 1999
Cardiac muscle	Smooth muscle cells	Beltrami *et al* 2003
Mesenchymal	Adipocytes, osteoblasts, chondrocytes, tendon. Muscle? Neural? Hepatocytes?	Pittenger *et al* 1999
Multipotent adult progenitor cells (MAPCs)	Cells of all three germ layers including liver, brain and muscle	Jiang *et al* 2002a

Table 1.1. Adult stem cell populations (in humans as well as in animal models) and their differentiation potential. Each stem cell population in this representative selection has been characterised to varying degrees, with haematopoietic stem cells being the most well characterised stem cell population. Each adult stem cell population is responsible for forming distinct cell subsets, but there may be some degree of plasticity in their developmental potential (indicated by a ? next to the cell lineage, as stem cell plasticity is a debated, and not fully defined, phenomenon).

1.3.1 Haematopoietic stem cells

The presence of a cell within mouse bone marrow that could rescue the haematopoietic system of lethally irradiated mice provided the first evidence of the haematopoietic stem cell (HSC). Clonogenic colonies of donor origin were formed in the spleen, and were found to contain cells of the myeloerythroid lineage, as well as occasional cells that could reconstitute secondary hosts after serial transplantation (Till and McCulloch 1961, Siminovitch *et al* 1963). Identification of surface markers of bone marrow cells, followed by purification of cell populations with specific marker combinations, allowed the isolation and characterisation of mouse HSC and multiple progenitor populations (Spangrude *et al* 1988, Morrison and Weissman 1994). The only population with the ability to self renew and form cells of all blood lineages in long-term reconstitution experiments was termed long term HSC (LT-HSC, Fig 1.1A), which has a specific marker expression profile (C-kit$^+$Thy-1.1^{lo}Lin$^-$Sca-1$^+$) (Spangrude *et al* 1988, Morrison and Weissman 1994). Short term HSC (ST-HSC) and multipotent progenitors (MPP) are subsets of this population (with a Lin$^{-/lo}$ phenotype); they are capable of giving rise to all blood lineages but ST-HSC are capable of short-term self-renewal whereas MPP have minimal self-renewal ability (Morrison and Weissman 1994). Several oligopotent precursor populations also exist, such as common myeloid progenitors (CMP) and common lymphoid progenitors (CLP, Fig 1.1A), which differentiate into distinct subsets of blood lineages (Kondo *et al* 1997, Akashi *et al* 2000). Thus, HSCs exist in a developmental hierarchy, where the multipotent, self-renewing population (LT-HSCs) gives rise to various progenitor cells with progressively more restricted differentiation and/or self-renewing potential, and eventually to precursors of a single mature cell type. As a result of the expression of specific sets of surface markers, populations at all levels of this hierarchy have been prospectively isolated and characterised in detail. The ability to obtain highly purified populations of mouse HSC has allowed for their *in vivo* self-renewal and differentiation capacity to be definitively demonstrated on the single-cell level (Spangrude *et al* 1988, Morrison and Weissman 1994, Osawa *et al* 1996). Whereas mouse HSC reside mostly within the CD34$^-$ population, human HSCs have also been isolated, but are mostly CD34$^+$; they have been prospectively isolated as CD34$^+$CD38$^-$ $^{/lo}$Lin$^-$ (Baum *et al* 1992, Bhatia *et al* 1997), and are estimated to constitute 0.05-0.1%
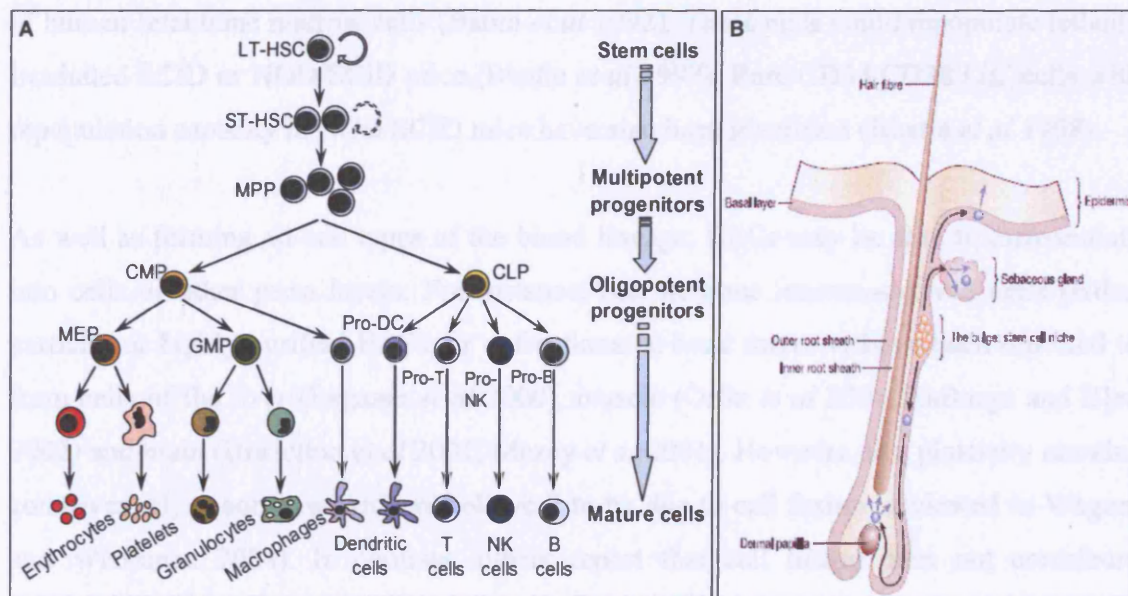
Figure 1.1. Adult stem cell populations. A, Haematopoietic stem cell lineage. LT-HSC have long term self-renewal capacity and can form all blood lineages. They give rise to ST-HSC with short-term self-renewal capacity, then MPP, with no or little self-renewal capacity. Oligopotent progenitors include CMP and CLPs. CMPs differentiate into MEPs that form erythrocytes and platelets, and also GMPs that form granulocytes and macrophages. CLPs differentiate into T lymphocytes, B lymphocytes and natural killer cells. Both CMPs and CLPs give rise to dendritic cells. Figure adapted from Passegué *et al* 2003; B, Epidermal stem cells in the hair follicle. Stem cells reside in the bulge region (orange cells) from which they can migrate to the sebaceous gland, epidermis and basal region. Here, they proliferate (blue cells) and form cells of the hair follicles, sebaceous glands and possibly interfollicular epidermis. Figure reprinted by permission from Macmillan Publishers Ltd: Nature, Reya and Clevers © 2005.

of human fetal bone marrow cells (Baum *et al* 1992). These cells could repopulate lethally irradiated SCID or NOD/SCID mice (Bhatia *et al* 1997). Rare CD34⁻CD38⁻Lin⁻ cells with repopulation capacity in NOD/SCID mice have also been identified (Bhatia *et al* 1998).

As well as forming all cell types of the blood lineage, HSCs may be able to differentiate into cells of other germ layers. For instance, murine bone marrow-derived cells (either partially or highly purified HSCs, or unfractionated bone marrow) have been reported to form cells of the liver (Lagasse *et al* 2000), muscle (Orlic *et al* 2001, LaBarge and Blau 2002) and brain (Brazelton *et al* 2001, Mezey *et al* 2001). However, this plasticity remains controversial, as some researchers believe it to be due to cell fusion (reviewed in Wagers and Weissman 2004). In contrast, others report that cell fusion does not contribute significantly to the observations, but such studies used un-fractionated bone morrow (Harris *et al* 2004), or a partially characterised HSC population (Jang *et al* 2004). Definitive demonstration of transdifferentiation will require the use of highly purified, phenotypically characterised cell populations, combined with functional studies of the "transdifferentiated" cell type, and importantly evidence (e.g. using FISH to detect a diploid chromosome complement) that cell fusion has not occurred (Wagers and Weissman 2004).

1.3.2 Stem cells of the skin: hair follicle stem cells

The mammalian epidermis consists of multiple layers of epithelium punctuated by hair follicles, complex structures specialised in making the hair shaft, and sebaceous glands (Fig 1.1B). Early studies identified cells within this niche with the ability to retain ³H-thymidine for long periods in mice (and thus were called label retaining cells or LRCs), indicating a slow cell cycle time; this is a characteristic of many stem cell populations (Cotsarelis *et al* 1990, Taylor *et al* 2000). The LRCs were found to be located in the bulge region of the hair follicle (Fig 1.1B), and could give rise to cells of the hair follicle and upper epithelial cells (Taylor *et al* 2000). Furthermore, it has been shown that LRCs, also called bulge or hair follicle stem cells (HFSCs), can migrate into the basal region of the hair follicle, as well as the upper follicle and epidermis, and give rise to a transit-amplifying population, capable of several further rounds of division before finally differentiating. These HFSCs gave rise to all cell lineages of the hair follicle and

sebaceous gland during homeostasis, as well as the interfollicular epidermis after trauma in mice, and were thus termed multipotential stem cells (Taylor *et al* 2000, Oshima *et al* 2001). Lineage-tracking studies, performed by labelling the bulge cells using label retention or GFP expression driven by a promoter that is specifically activated in bulge cells and following their fate *in vitro* or *in vivo*, allowed enrichment of the bulge cell population and again highlighted the ability of these cells to differentiate into all cell types in the hair follicle and surrounding epithelia (Tumbar *et al* 2004, Morris *et al* 2004). Recent work has succeeded in isolating clonal populations of murine bulge HFSCs using a combination of GFP-marking of the bulge cell population plus FACS sorting for CD34^{+} cells (a marker of HFSCs) (Blanplain *et al* 2004). This study demonstrated self renewal *in vitro*, as well as the formation of epidermis, hair follicles and sebaceous glands after grafting of clonal populations (cultured *in vitro*) into recipient mice, providing strong evidence of their multipotentiality. However, other studies suggest that HFSCs do not contribute to the epidermis during homeostasis, but only in response to tissue injury, and that other stem cell populations on the epidermal niche may perform this role (Claudinot *et al* 2005, Ito *et al* 2005, Levy *et al* 2005). This indicates that further definition of each cell population must be achieved; progress has been made using gene expression microarray profiling (1.7.1) of either enriched (Tumbar *et al* 2004, Morris *et al* 2004) or clonal (Blanplain *et al* 2004) populations of bulge stem cells. These were compared to profiles of non-stem cell populations from the epidermis, or other stem cell populations, and revealed numerous candidate genes that may contribute to their stem cell properties. These studies were performed using mouse models, and recently human bulge cells, of which 20% represented the label-retaining cell population, were obtained from human skin samples (Ohyama *et al* 2006). Their gene expression profile, obtained via microarray analysis, was compared with that of non-bulge cells, and a subset of genes that characterised the bulge cell population was identified, including surface markers that may allow for enrichment of this cell population (Ohyama *et al* 2006). Thus, progress is being made towards defining the molecular basis of the stem cell properties of this adult stem cell population.

1.4 Mesenchymal stem cells

1.4.1 Identification of a multipotent mesenchymal stem cell within the bone marrow

Two primary elements are typically thought to constitute the bone marrow: the haematopoietic tissue, and the associated stroma (Bonnet 2003). The stromal portion of bone marrow contains many cell types, including endothelial cells, mature cells of the mesenchymal lineage, such as osteoblasts and adipocytes, and may also contain cells of the haematopoietic lineage, for example macrophages. Bone marrow stromal cells were initially characterised as having the ability to support haematopoiesis, as well as maintain HSCs in an undifferentiated state (Pittenger and Marshak 2001). In recent years, a second bone marrow stem cell has been characterised that resides within the stromal cell compartment and can, like HSCs, give rise to multiple cell lineages.

The possibility of the existence within bone marrow of a cell with the ability to form various mesenchymal tissues first arose from experiments performed during the 1950s and 60s. The osteogenic potential of bone marrow was shown when whole bone marrow samples were transplanted at various sites into host animals. Histological analysis revealed extensive formation of tissues resembling bone and marrow stroma at the site of transplantation, which could also support haematopoiesis from invading host haematopoietic stem cells (Urist and McLean 1952, Tavassoli and Crosby 1968). Similar experiments involved the use of diffusion chambers, which consist of semi-permeable membranes separated by a plastic ring that do not allow passage of donor or host cells. Transplantation of chambers containing either fragments or cell suspensions of bone marrow resulted in formation of osteogenic tissue (alkaline phosphatase-staining) within the chamber after around day 30 (Friedenstein *et al* 1966). This indicated that, rather than the donor material attracting osteogenic host cells, there existed within the transplanted marrow stroma a progenitor cell capable of differentiating into bone.

Friedenstein and co-workers showed that the osteogenic potential of bone marrow was a feature of a small population of cells, termed fibroblast colony forming cells (FCFC) or colony forming units-fibroblastic (CFU-F). When single-cell suspensions of guinea pig bone marrow were cultured *in vitro* at low densities in serum-containing media, colonies

of fibroblasts, derived from single cells, were formed (Friedenstein *et al* 1970, Friedenstein 1976). Rodent CFU-F were shown to have a high proliferative capacity *in vitro*, but to be in a resting state *in vivo* (Friedenstein 1976); studies of the human counterpart of CFU-Fs were in strong agreement with this (Castro-Malaspina *et al* 1980). Intraperitoneal transplantation into host animals of diffusion chambers containing cultured guinea pig fibroblast colonies resulted in bone formation, which did not occur when fibroblasts cultured from the spleen were transplanted (Friedenstein *et al* 1970). This work was confirmed and extended by others, who showed extensive formation of bone, cartilage and fibrous tissue within diffusion chambers inoculated with marrow fibroblast cultures (Ashton *et al* 1980). It was proposed that an undifferentiated stromal cell precursor may exist, capable of forming both bone and cartilage; however the presence of committed progenitor cells, each giving rise to a different cell type, could not be ruled out. This question was addressed by studies of *in vivo* differentiation of single fibroblastic clones. About 15% of colonies transplanted under the renal capsule produced bone marrow organ, consisting of osteogenic tissue, typical adipose cells and marrow stromal cells (Friedenstein 1980). Thus, colonies derived from a single cell were capable of giving rise to multiple cell lineages.

The concept of a multipotent stem cell of stromal origin residing in bone marrow was formally presented by Owen in 1978. It was suggested that the differentiated cell types residing in marrow stroma may derive from a single common progenitor, or stem cell, analogous to the haematopoietic system within the marrow (Owen 1978). A more detailed model was later proposed, in which stem cells, committed progenitors and mature cells were all present in marrow stroma, with reticular, fibroblastic, adipocytic and osteogenic being the possible differentiated phenotypes due to experimental evidence showing differentiation to these lineages (Owen 1985, 1988). Caplan expanded this theory to include all mesenchymal lineages, namely myocytes, chondrocytes, tenocytes, osteocytes, adipocytes and stromal and dermal fibroblasts (because these embryonic cell types were thought to derive from the mesodermal germ layer, rather than there being experimental evidence at the time to support this differentiation potential), and was the first to use the term mesenchymal stem cell (MSC) (Caplan 1991).

The studies of differentiation of isolated clones went some way to demonstrating the multipotentiality of MSCs. Further evidence was provided from differentiation of conditionally immortalised mouse MSC clones, one of which had the ability to form four mesenchymal phenotypes: chondrocytes, osteoblasts, adipocytes and haematopoiesis-supportive cells (Dennis *et al* 1999). A key study in 1999 characterised human MSCs (hMSCs) and their differentiation potential. Human MSCs were estimated to constitute a fraction (0.001-0.01%) of nucleated cells within the bone marrow. It was shown that hMSCs isolated from three different donors could all give rise to osteocytes, chondrocytes and adipocytes. Moreover, clonal populations were established, and 2 out of 6 populations could differentiate into all three lineages, consistent with the donor hMSCs from which they were derived (Pittenger *et al* 1999). Thus, the existence of cells with *in vitro* multipotentiality was demonstrated.

1.4.2 Differentiation potential of bone marrow-derived MSCs

1.4.2.1 MSC differentiation to mesodermal lineages

Adipocytes, osteoblasts, chondrocytes and haematopoiesis-supportive stroma. As described above, MSCs were originally identified via their ability to differentiate into particular mesodermal lineages. The studies by Pittenger and Dennis definitively demonstrated their ability, at the clonal level, to differentiate *in vitro* into adipocytes, osteoblasts, chondrocytes (Fig. 1.2), and also haematopoiesis-supporting stromal cells in the case of murine MSCs (Dennis *et al* 1999, Pittenger *et al* 1999). Defined hormonal regimens have been developed that induce robust and specific *in vitro* differentiation into adipocytes (1.6.1), osteocytes and chondrocytes (Pittenger *et al* 1999, Tustsumi *et al* 2001), so the ability of MSCs (both human and rodent) to differentiate into these lineages is well established.

Tenocytes MSCs from rabbit bone marrow were shown to differentiate into tenocytes, capable of replacing severed tendon (a tissue of mesodermal origin) in an animal model (Young *et al* 1998), but this lineage is little studied, and *in vitro* differentiation of clonal hMSC populations to tenocytes has not been reported.

33
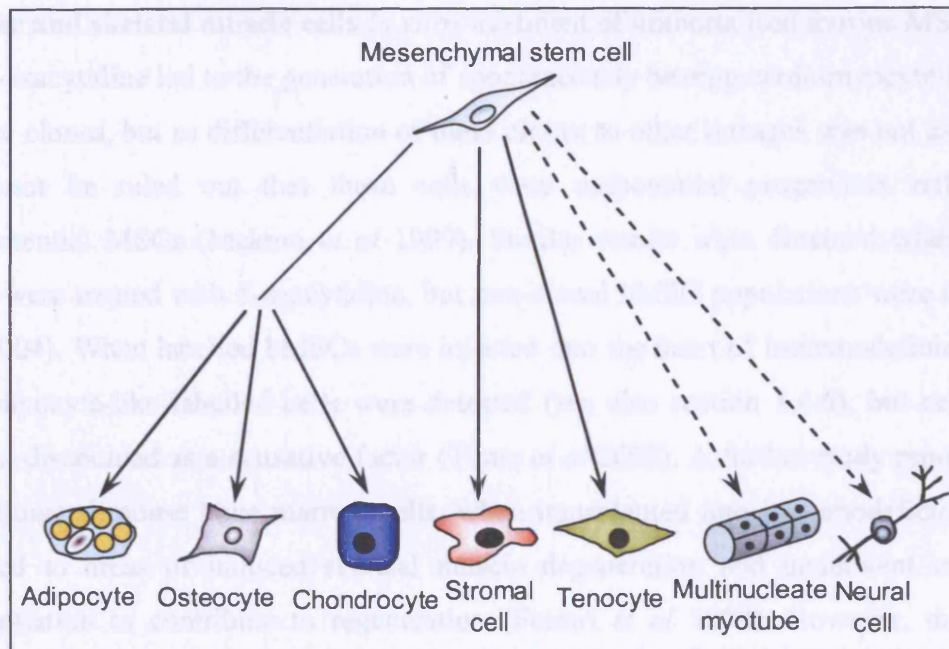
Figure 1.2. Differentiation potential of MSCs. Clonal populations of MSCs can differentiate *in vitro* into adipocytes, osteocytes and chondrocytes. They have also been shown to give rise to haematopoiesis-supportive stromal cells and tenocytes capable of repairing damaged tendon. MSCs may also be able to form cells of the muscle and neural lineages, but this capacity is under debate.

Cardiac and skeletal muscle cells *In vitro* treatment of immortalised murine MSC clones with 5-azacytidine led to the generation of spontaneously beating cardiomyocyte-like cells in some clones, but as differentiation of these clones to other lineages was not assessed it could not be ruled out that these cells were unipotential progenitors rather than multipotential MSCs (Makino *et al* 1999). Similar results were obtained when human MSCs were treated with 5-azacytidine, but non-clonal hMSC populations were used (Xu *et al* 2004). When labelled hMSCs were injected into the heart of immunodeficient mice, cardiomyocyte-like labelled cells were detected (see also section 1.4.6), but cell fusion was not discounted as a causative factor (Toma *et al* 2002). A further study reported that unfractionated mouse bone marrow cells, when transplanted into immunodeficient mice, migrated to areas of induced skeletal muscle degeneration and underwent myogenic differentiation to contribute to regeneration (Ferrari *et al* 1998). However, this result cannot be attributed directly to MSCs, as an unpurified cell population was used. Thus, there is some evidence of differentiation towards the myogenic lineage by cells that co-purify with MSCs, but further work using clonal populations and demonstrating multilineage potential is required to show that this is a property of MSCs rather than more restricted progenitor cells.

In vivo **differentiation potential** Several studies of systemic infusion of labelled murine MSCs into irradiated mice reported engraftment into several mesodermal tissues, including marrow, bone and cartilage (as well as lung, an endodermal tissue), combined with tissue-specific gene expression (Pereira *et al* 1995) and signs of osteoblast differentiation in an osteogenesis imperfecta mouse model (Pereira *et al* 1998). In a more recent study, GFP-tagged, phenotypically characterised murine MSCs were intravenously infused into minimally irradiated mice, and were found after 1 month mainly in the lungs, liver and kidney, but also in the muscle, heart, brain and spleen (Anjos-Alfonso *et al* 2004). Immunohistochemical assessment revealed that MSCs differentiated into lung epithelial cells, hepatocytes, renal tubular-like cells and myofibroblasts, albeit at a low frequency. Human MSCs that had been transplanted *in utero* into sheep engrafted and survived in multiple non-haematopoietic tissues for up to 13 months (Leichty *et al* 2000). Furthermore, hMSCs were found to differentiate into cells with morphological and phenotypic properties of chondrocytes, adipocytes, bone marrow stromal cells, thymic

epithelial cells, cardiomyocytes and skeletal myocytes. Thus, these studies provide knowledge of MSC behaviour *in vivo*, but it must again be emphasised that these populations were not clonal, and cell fusion was not investigated. However, as well as mesodermal differentiation, the cell populations used were found to differentiate into lineages of non-mesodermal origin, providing evidence that heterogeneous MSC populations may have additional differentiation properties.

1.4.2.2 MSC differentiation to non-mesodermal lineages

Hepatocytes A small study described the treatment of non-clonal hMSCs with hepatocyte growth factor (Wang *et al* 2004). After 20 days, slight expression at the mRNA and protein level of some hepatocytes markers was seen, indicating a possible hepatocyte-like phenotype, but functional studies were not performed.

Neural cells Several studies that have claimed to induce *in vitro* neural differentiation of rodent or human MSCs used β-ME or DMSO/BHA to initiate the process. This protocol was pioneered by Woodbury and co-workers: after treatment, human or rat MSCs showed expression of neurological markers and outgrowth of neuronal processes (Woodbury *et al* 2000). Similar studies reported the generation of neurons with electrophysiological properties (Hung *et al* 2002), or with characteristics similar to neurons produced after co-culture with Schwann cells (Zurita *et al* 2005). However, it has recently been reported that β-ME or DMSO/BHA treatment causes stress-induced depolymerisation of the actin cytoskeleton and retraction of the cytoplasm in numerous cell types, meaning that the "neural" morphology observed was a result of stress rather than differentiation (Lu *et al* 2004, Neuhuber *et al* 2004). Moreover, no increase in neural marker expression was seen using Western blotting, and it was concluded that the increase in immunohistochemical staining for neural markers was due to higher antigen density associated with smaller cell size.

Other *in vitro* studies have utilised growth factors known to promote neural differentiation of ESCs and neural stem cells. Incubation of human or mouse bone marrow stromal cells with differentiation medium containing such factors, or co-culture of labelled murine MSCs with rat midbrain cell suspensions, resulted in a small percentage of cells with

morphologies similar to immature neurons or glial cells and occasional expression of neural markers (Sanchez-Ramos *et al* 2000). Again however, cell fusion was not discounted from the co-culture experiments. Also, no tests of neuronal function were performed. Hermann and co-workers used suspension culture of non-clonal hMSCs followed by incubation with different neural growth factors (Hermann *et al* 2004). Cells with various phenotypic and functional properties reminiscent of neural cells were obtained, including: dopamine-releasing cells expressing some neural markers; cells with phenotypic characteristics of astrocytes, oligodendrocytes and rare immature neuron-like cells. A proportion of cells treated with a specific growth factor had voltage-dependent electrophysiological outputs in patch-clamp experiments typical of glial cells.

With regards to the *in vivo* ability of MSCs to form neural cells, Kopen *et al* injected BrdUrd-labelled murine MSCs into neonatal mouse brains (Kopen *et al* 1999). Cells double-labelled for BrdUrd and GFAP, an astrocyte marker, or neurofilament, a neuronal marker, were identified, indicating that the MSCs differentiated into these cell types, although the question of whether MSCs may have fused with recipient brain cells was not addressed.

In summary, although these studies have raised the possibility of MSC differentiation to neural and hepatic lineages, definitive evidence of this ability has not yet been provided. This will require *in vitro* studies of clonal MSC populations incorporating phenotypic and functional assays, combined with *in vivo* investigations that also address the issue of cell fusion.

1.4.2.3 Multipotent adult progenitor cells (MAPCs)

A subpopulation of bone marrow cells that co-purify with hMSCs, and that could be isolated from human or rodent marrow, has been described (Reyes *et al* 2001, Jiang *et al* 2002a, 2002b). They were termed multipotent adult progenitor cells, or MAPCs. Colonies derived from single cells were shown to be able to differentiate *in vitro* into cells with phenotypic and functional characteristics of endothelial cells (visceral mesoderm), neurons, astrocytes and oligodendrocytes (neuroectoderm), and hepatocytes (endoderm), as well as cells of mesenchymal lineages (Reyes *et al* 2001, 2002, Jiang *et al* 2002a,

2002b, Schwartz *et al* 2002). In fact when injected into an early mouse blastocyst, single mouse MAPCs contributed to most somatic tissues (Jiang *et al* 2002a). A possibility suggested to explain these findings was that MAPCs could represent a pluripotential stem cell that persists into adulthood and is capable of differentiating into cells of all three germ layers. However, it could not be ruled out that cell fusion may be responsible for the *in vivo* observations (Jiang *et al* 2002a). The relationship of MAPCs to hMSCs is also a matter for debate. The *in vitro* culture conditions of the two populations are very different; MAPCs are cultured at low densities and 2% FCS, whereas hMSCs are cultured at higher densities and 10% FCS. It is possible that this could contribute to the differences in potentiality observed, and that hMAPCs are the result of de-differentiation of hMSCs due to culture conditions (Bonnet 2003). Thus, whether MAPCs truly exist as a pluripotential stem cell population *in vivo* is as yet unknown, and whether they may be related to hMSCs is a matter for future research.

1.4.3 Characterisation of the human bone marrow MSC population

As previously stated, bone marrow hMSCs reside within a tissue consisting of numerous cell types, and it has been the aim of many investigations to characterise hMSCs in order to isolate a pure hMSC population. The classical method of hMSC isolation is via density gradient centrifugation to separate the mononuclear cell fraction, followed by adherence to plastic, thus giving some degree of purification as non-adherent haematopoietic cells are removed (Pittenger *et al* 1999). However, a lack of knowledge regarding surface markers has precluded the use of more sensitive antibody-aided selection. Studies of surface markers of cultured hMSCs have shown that they are distinct from haematopoietic cells, as they do not express CD45, CD34 and CD14, markers of the haematopoietic lineage (Pittenger *et al* 1999, Conget and Minguell 1999). Numerous other surface molecules have been identified (Table 1.2), and indicate that hMSCs do not have a unique antigenic profile, but express markers of mesenchymal (SH2, SH3 and SH4, markers of mesenchymal progenitors; ACTA1, vascular smooth muscle marker), endothelial (VCAM-1) and epithelial cells (cytokeratins 18 and 19) (Conget and Minguell 1999). It must be noted however that varied culture conditions used by different laboratories and continued time in culture mean that contradictory results regarding MSC marker profiles are frequent (Bonnet 2003).

Marker Type	Marker Name	
	Expressed	Not expressed
Specific antigens	SH2, SH3, SH4, Stro-1, ACTA1	CD133
Haematopoietic markers		CD4, CD14, CD34, CD45 (PTPRC), c-kit/SCFR (CD117)
Cytokines and growth factors	IL- 1α, 6, 7, 8, 11, 12, 14, and 15 LIF, SCF, GM-CSF, G-CSF, M-CSF	
Cytokine and growth factor receptors	IL1R, IL3R, IL4R, IL6R, IL7R, LIFR, SCFR, G-CSFR, IFNγR, TNFR1, TNFR2, TGFβR1, TGFβR2, bFGFR, PDGFR, EGFR	IL-2R (CD25)
Adhesion molecules	Integrins α1 (CD49a), α2 (CD49b), α3 (CD49c), aα (CD49e), β1 (CD29), β3 (CD61), β4 (CD104)	α4 (CD49d), αL (CD11a), Cβ2 (CD18)
Extracellular matrix molecules and receptors	ICAM-1 (CD54), ICAM-2 (CD102), VCAM-1 (CD106), ALCAM-1 (CD166), LFA3 (CD58), L-selectin (CD62L), endoglin (CD105), hyaluronate receptor (CD44) CK18, CK19	ICAM-3 (CD50), E-selectin (CD62E), P-selectin (CD62P), PECAM-1 (CD31), vWF, Cadherin 5
Other	CD9, CD13, Thy-1 (CD90), HLA-ABC (MHC I) (low)	HLA-DR (MHC II)

Table 1.2. Surface marker expression profile of MSCs. Data obtained from Pittenger *et al* 1999, Minguel *et al* 2001, Majumdar *et al* 2003 and Wagner *et al* 2005.

Unique markers are needed to definitively identify and purify non-expanded MSCs. To date, the most promising candidate has been the monoclonal antibody STRO-1. It's antigen (as yet unidentified) is expressed by a population of bone marrow stromal cells (termed CFU-Fs in these studies) that possesses the ability to proliferate extensively and differentiate *in vitro* into osteoblasts, adipocytes, chondrocytes and haematopoiesis-supportive stroma (Simmons and Torok-Storb 1991, Dennis *et al* 2002). However, around 95% of this STRO-1$^+$ population constituted glycophorin A-positive nucleated erythroid precursors, so a further glycophorin A$^-$ selection was required to enrich for CFU-Fs (Simmons and Torok-Storb 1991). Double selection for cells from whole human bone marrow expressing high levels of STRO-1 (STRO1^{BRIGHT}) and also VCAM1/CD106 resulted in 5000-fold enrichment of the CFU-F population (Gronthos *et al* 2003a). A subset of STRO1^{BRIGHT}VCAM$^+$ cells exhibited extensive *in vitro* proliferation and clonal multipotentiality; as other clones exhibited low proliferation and differentiation, it was concluded that multipotential stem cells constituted a subset of the isolated population, which also contained more restricted progenitor cells (Gronthos *et al* 2003a). CFU-Fs have also been enriched via selection of CD45$^-$,CD14$^-$,CD73$^+$/CD49a$^+$ (Boiret *et al* 2005) or CD45^{low},D7-FIB$^+$ (a fibroblast marker of unknown function, Jones *et al* 2002) populations from non-expanded bone marrow samples, but a pure CFU-F population was not obtained from either of these methods. Other attempts to define the hMSC population have included comparison of the gene expression profiles of hMSCs from different locations (1.4.5) and fibroblasts (which are morphologically similar to hMSCs), where certain genes were enriched or depleted in hMSCs (Ishii *et al* 2005). Combined with the cell surface marker studies it is clear that progress is being made towards defining the hMSC population, but there is as yet no marker that is unique to hMSCs.

1.4.4 The stem cell nature of MSCs

Although the term mesenchymal stem cell is widely used, no unequivocal evidence demonstrating stem cell properties for hMSC *in vivo* has yet been provided (Dennis and Caplan 2004, Baksh *et al* 2004). Although hMSC can be expanded for 15-25 passages *in vitro* depending on culture conditions and donor age without losing differentiation potential, this *in vitro* proliferative capacity is not indefinite (Bruder *et al* 1997, Conget and Minguell 1999,). Furthermore, it is possible that hMSC *in vitro* represent a

heterogeneous population of multipotential stem cells and lineage-restricted progenitors; evidence from this comes from the studies of Gronthos and Pittenger where only a proportion of hMSC clones could differentiate into three lineages (Pittenger *et al* 1999, Gronthos *et al* 2003a,). This could be a reflection of a stem cell hierarchy similar to HSC (Owen 1985, Minguel *et al* 2001). Furthermore, sub-optimal culturing conditions *in vitro* may also cause loss of potentiality of a proportion of multipotential cells, leading to a heterogeneous population. Definitive proof of hMSC self-renewal and multipotentiality would come from the isolation of a highly purified cell population followed by demonstration that, at the single cell level, un-manipulated hMSCs could give rise *in vivo* to diverse progeny and repopulate host animals after serial transplantation (Joseph and Morrison 2005). Currently the only adult stem cell population for which these rigorous criteria have been demonstrated is the HSC (1.3.1), and the lack of a hMSC-specific marker precludes such work in this population. However, various observations from cultured or freshly isolated hMSCs point towards a stem cell function.

Clonal populations of hMSC are capable of giving rise robustly to multiple lineages *in vitro*, arguing strongly for the presence of a multipotential stem cell (Pittenger *et al* 1999, Gronthos *et al* 2003a). In addition, studies of freshly isolated, non-expanded hMSCs revealed that they lacked detectable expression of the Ki-67 antigen, a marker of cycling cells, and were mostly in G0/G1 phase of the cell cycle, entering the cell cycle after exposure to serum in culture (Castro-Malaspina *et al* 1980, Conget and Minguell 1999, Gronthos *et al* 2003a). This indicates that hMSC *in vivo* are a resting cell population, a characteristic of other stem cell populations (Taylor *et al* 2000). Freshly isolated hMSC were also found to exhibit telomerase activity (Gronthos *et al* 2003a) which persisted in culture for at least 12 passages (Pittenger *et al* 1999); this is again a characteristic of stem cells with self-renewal capacity.

In summary, bone marrow-derived hMSCs exhibit multiple stem cell properties, but further work will be needed to definitively demonstrate their *in vivo* stem cell nature. *In vitro* studies of their stem cell characteristics, such as the ability to robustly differentiate into specific lineages, may further our understanding of stem cell function and provide mechanisms to study this population *in vivo*.

41

1.4.5 Tissue-specific MSCs

In addition to their presence in bone marrow, cells with the characteristics of hMSCs reside in other human tissues, including trabecular bone (Noth *et al* 2002), deciduous teeth (Miura *et al* 2003), skin and muscle (Young *et al* 2001). MSC-like cells from adipose tissue and umbilical cord blood have received the most attention. Fibroblastic, plastic-adherent cell populations were isolated from lipoaspirates (ADSC, Zuk *et al* 2001, 2002) and placental blood (UBSC, Erices *et al* 2000, Lee *et al* 2004a) which could be maintained *in vitro* for many population doublings. Clonal populations of both ADSCs and UBSCs showed trilineage differentiation potential to adipocytes, osteocytes and chondrocytes (Zuk *et al* 2002, Lee *et al* 2004a), although these multipotential clones were reported to be rare in ADSC and bi- and uni-potential clones were also observed. ADSC were shown to differentiate into adipocytes that functionally resembled white adipose tissue, as well as adipocytes from differentiated bone marrow hMSC (Rodriguez *et al* 2004, Dicker *et al* 2005). Clonally-derived UBSC populations underwent differentiation into hepatocyte-like and neural cells (Lee *et al* 2004a). Heterogeneous populations of ADSC and UBSC have been shown to differentiate into myogenic cells (Zuk *et al* 2001, 2002, Gang *et al* 2004), and ADSC may also be able to form hepatocyte-like (Seo *et al* 2005) and neural cell lineages (Zuk *et al* 2002, Safford *et al* 2002). However, some of the methods used to induce neurogenesis in these studies have come under question (1.4.2.2), and clonal studies will be needed to confirm findings from mixed cell populations because of a high potential for the presence of contaminating cells causing the observations.

The relationship between bone marrow hMSCs, ADSCs and UBSCs has been investigated. Surface marker profiles were found to be comparable between all populations (Zuk *et al* 2002, Lee *et al* 2004a, Wagner *et al* 2005). A microarray comparison found ADSCs and bone marrow-hMSCs to be highly similar, with only 25 genes from an array representing over 10,000 genes showing differential expression (Lee *et al* 2004b). However, a second study showed numerous genetic differences through pairwise comparisons of the three populations, but Gene Ontology analysis (1.7.1.2) of the differentially expressed genes did not reveal gene expression trends characteristic of the tissue of origin (Wagner *et al* 2005).

Interestingly, there have also been reports of the isolation of cells with a similar antigenic profile to MSCs and the ability to differentiate into adipocytes and osteocytes from normal whole blood (Zvaifler *et al* 2000, Kuznetsov *et al* 2001). An elegant study captured circulating precursor cells in rats that demonstrated adipogenic, osteogenic, chondrogenic and myogenic differentiation in culture, and when immortalised showed clonal differentiation to these lineages (Wu *et al* 2003). Furthermore, labelled clonal cells were shown to home to the bone marrow and also areas of injured cardiac tissue (Wu *et al* 2003). These studies indicate that a multipotential precursor cell, reminiscent of MSC, may circulate in the blood and have the ability to home to sites of injury or disease. This may also provide a link between the bone marrow and tissue-specific MSC-like populations, but as some studies were not able to isolate such cells from peripheral blood the ability of MSC to circulate is still controversial, and such theories require further investigation.

1.4.6 Physiological and therapeutic roles for MSCs

HSCs and HFSCs reside in tissues that constantly turnover and play an essential role in normal tissue replacement and response to injury (Joseph and Morrison 2005). It is possible that a physiological role for hMSC would also be tissue repair and renewal (Dennis and Caplan 2004). The presence of multipotential hMSC-like cells in tissues which they are capable of giving rise to, combined with a possible ability to circulate in the blood and home to a specific location, may support this theory, but this will remain as speculation until there is definitive evidence that MSCs exhibit self-renewal and multilineage differentiation at the clonal level *in vivo* (Roufosse *et al* 2004).

Although a physiological role for *in vivo* hMSCs remains hypothetical, there is clear evidence of the efficacy of hMSCs use in therapeutic repair strategies for human diseases of mesenchymal tissues. Human MSCs are easily isolated, expanded and manipulated, have the potential for allogenic transplantation, and as adult stem cells they are not surrounded by the ethical issues created by using embryonic stem cells for stem cell therapy. By way of example, MSCs from animal models have shown promise in the treatment of tendon (Young *et al* 1998) and cartilage defects (Wakitani *et al* 1994). LacZ-tagged hMSCs, when injected into the hearts of healthy mice, were shown to differentiate

into cardiomyocytes, as assessed by morphological changes and expression of myogenic markers, but cell fusion was not ruled out (Toma *et al* 2002). Such experiments involved the injection of MSCs either systemically, or into the site of injury/disease, where they engrafted and showed signs of differentiating into the desired tissue, leading to repair of damage and/or improved tissue function. In a group of clinical studies, allogenic bone marrow transplantation was used to treat osteogenesis imperfecta in children, a genetic disorder that causes severe bone fracturing and growth abnormalities. Six months after transplantation dramatic improvements in bone mineral content, body growth and fracture incidence were seen, but growth rate gradually slowed over time (Horwitz *et al* 1999). In further work, retrovirally-tagged allogenic hMSCs were infused systemically into osteogenesis imperfecta patients that had undergone this marrow transplantation, and were found to engraft, to varying degrees, in bone, skin and marrow stroma, and produced measurable growth improvements (Horwitz *et al* 2002).

An understanding of the molecular events occurring during hMSC differentiation into specific lineages could provide novel therapeutic interventions for certain conditions. It is now known that osteoblasts play a crucial role in maintaining the HSC niche within the bone marrow (reviewed in Moore and Lemischka 2006). It has also been suggested that there may be a reciprocal relationship between osteogenesis and adipogenesis within the bone marrow, and that the decrease in bone mass associated with osteoporosis and age-related osteopenia may be a consequence of the increasing marrow adipose differentiation that is also observed with age (Gimble *et al* 1996, Nuttall and Gimble 2000). Thus, controlling the balance of differentiation into these two lineages may provide a means to stabilise bone formation, and maintain a constant microenvironment within the bone marrow.

1.5 Adipose tissue

1.5.1 Structure and development of adipose tissue

Adipocytes are the main constituent of adipose tissue, and make up around one to two thirds of the total cell number (Ailhaud *et al* 1992). They are held in a network of collagen fibres, and contribute the major role of adipose tissue by storing lipid in the form of triacylglycerol (Albright and Stern 1998). Two types of adipose tissue exist: white adipose tissue (WAT) and brown adipose tissue (BAT), which have opposing physiological roles (1.5.2). Mature white adipocytes vary greatly in size, reflecting their ability to accumulate different amount of lipids, and are unilocular; they contain a single, centrally located lipid droplet that pushes the nucleus towards the plasma membrane (Fig. 1.3). Brown adipocytes are smaller than white adipocytes and have multiple, smaller fat droplets and a high number of mitochondria (the cytochromes of which result in the typical "brown" colour, Cinti 2001). Adipose tissue contains numerous other cell types. It is highly vascularised and innervated, and also contains fibroblastic connective tissue cells, macrophages and adipocyte precursor cells (or preadipocytes). These preadipocytes do not yet contain lipid, but are committed to the adipocyte lineage (Albright and Stern 1998).

The development of white adipose tissue begins prenatally in humans, probably during late gestation, but takes place to a large extent postnatally (Ailhaud 2001). The initial signs of adipogenesis (or the process of fat cell formation) are the organisation of vascular structures associated with accumulation of mesenchymal cells. It is thought that primitive mesenchymal stem cells, the embryonal equivalent of adult bone marrow MSCs, exist within the embryo and give rise to the preadipocytes and adipocytes within adipose tissue (Ailhaud *et al* 1992). Unipotential adipoblasts are the progeny of embryonic multipotential stem cells and presumably give rise to preadipocytes committed to the adipocyte lineage, but as specific markers to distinguish these populations *in vivo* do not exist, the processes of early adipose tissue development, and the events controlling each step, remain somewhat unclear (Ailhaud 2001). Preadipocyte formation is followed by formation of primitive adipocyte clusters, containing dense masses of fat cells grouped around capillaries (Ailhaud *et al* 1992). Initially, adipocytes contain few, small lipid droplets. The later phases of prenatal adipose development involve an increase in the number of
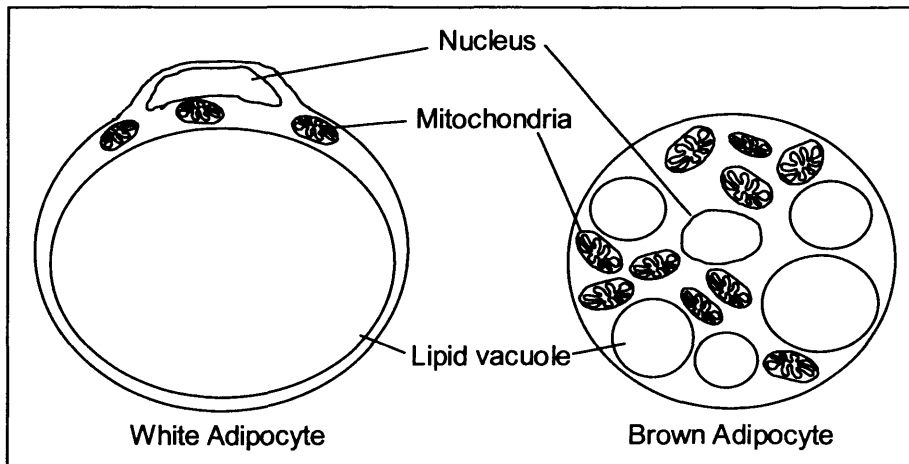
Figure 1.3. Structure of white and brown adipocytes. White adipocytes are unilocular and contain few mitochondria and an eccentric nucleus. Brown adipocytes contain multiple lipid droplets and mitochondria, and the nucleus may be centrally located. The structure of brown and white adipocytes corresponds closely to their roles in energy storage and dissipation, respectively. Figure adapted from Albright and Stern 1998.

preadipocytes, as well as adipocytes, but the lipid droplets do not expand. After birth, adipose tissue mass increases quickly, mainly due to an increase in adipocyte size from increasing volume and number of lipid droplets; the differentiation of preadipocytes also contributes to growth of adipose tissue at this time (Ailhaud *et al* 1992). ^{3}H-thymidine labelling experiments have shown that preadipocytes are highly proliferative postnatally, whereas mature adipocytes cannot proliferate or divide, so the development of adipose tissue is dependent upon *de novo* differentiation of preadipocytes (Rosen and Spiegelman 2000). Adipose tissue development occurs throughout the lifetime of an organism; preadipocytes exist in adult mammals that can differentiate in response to diet and environmental conditions (Ailhaud *et al* 1992). Mild obesity in humans is caused by increasing size of adipocytes (hypertrophy), but cells have a limit to how much lipid they can store, and when this limit is reached preadipocytes differentiate, increasing the number of fat cells (hyperplasty); this occurs in severe obesity (Grimaldi 2001).

1.5.2 Role of adipose tissue in the body

The traditional role attributed to mammalian adipose tissue was the storage and release of excess energy as lipid. Modern studies have dramatically altered this view, and it is now apparent that adipose tissue is a highly active metabolic and endocrine organ.

As discussed above, two forms of adipose tissue exist in mammals: WAT and BAT. As opposed to the energy storage role carried out by WAT, BAT serves primarily to dissipate energy in the form of heat. The BAT-specific protein UCP1 is central to this function, and uncouples energy production from fuel metabolism by dissipating the proton gradient over the inner mitochondrial membrane created by the electron transport chain. In this way, fuel metabolism can continue regardless of the cell's energy status, allowing significant heat generation (Ricquier and Bouillaud 2000). This function of BAT is crucial for non-shivering thermogenesis in hibernating mammals; accordingly they have large BAT deposits throughout life. In humans, BAT exists mainly in the neonatal period, and only in negligible amounts in adults (UCP1 mRNA can be detected in adult human WAT, indicating that small numbers of BAT cells may remain) (Champigny and Ricquier 1996). Its physiological significance in humans is unknown, although it has been suggested that BAT may protect against obesity in all species where it is present (Rosen and Spiegelman

2000). Thus, the predominant adipose species in humans is WAT, so most studies on the role and development of human adipose tissue have focused on WAT.

Adipose tissue in humans is found mainly either under the skin (subcutaneous) or surrounding major organs (visceral). Lipid storage at these sites serves multiple functions; subcutaneous adipose tissue is important for heat insulation, as it conducts energy poorly, and visceral adipose tissue provides a mechanical cushion and protects internal organs from jarring (Albright and Stern 1998). However, the main role of adipose tissue is to provide a buffering system when energy intake and output are imbalanced. In times of excess energy intake, adipocytes remove lipid from the bloodstream (found in chylomicrons or lipoproteins) as fatty acids through the action of lipoprotein lipase that is secreted by adipocytes and attached to the capillary endothelium (Albright and Stern 1998, Boschmann 2001). Fatty acids are then converted back to triacylglycerol through the action of numerous enzymes, and are stored in the cytoplasm as non-membrane-bound droplets. Triacylglycerol constitutes an ideal energy store because of it's high density and low water content relative to other energy-rich molecules such as glycogen (Klaus 2001). When energy output is greater than intake, triacylglycerol can be broken down into fatty acids and released into the bloodstream for use in energy-requiring tissues such as muscle and liver. This process is mediated mainly by the enzyme hormone-sensitive lipase (Boschmann 2001). Adipose tissue receives many signals to control this dynamic balance of fat metabolism, such as nutrient status (excess blood-borne lipid), neural signals (catecholamines from the sympathetic nervous system: increased triacylglycerol mobilisation) and hormone action (insulin: increased fatty acid uptake and decreased triacylglycerol mobilisation). It is a highly plastic organ, and can adapt swiftly to large changes in energy status, acting as a long-term energy reservoir or an immediate energy source (Spiegelman and Flier 1996, Boschmann 2001, Bays et al 2004).

Adipose tissue is now described as an endocrine organ; as well as being controlled by many signals, adipose tissue is known to express and secrete a range of molecules, or adipokines, that exert their effects locally (autocrine/paracrine) or through release into the bloodstream (endocrine) (Kershaw and Flier 2004). The discovery of the hormone leptin in 1994 led to this description (Zhang et al 1994). It is produced in adipose tissue and

secreted into the bloodstream. The major target of leptin is the hypothalamus, although its receptor can be found in other tissues of the body, indicating that is has wide-ranging effects (Trayhurn and Beattie 2001). The most well documented role of leptin is as an indicator of energy sufficiency. Leptin signals to decrease food intake and increase energy expenditure, and its levels correlate closely with extent of obesity (Spiegelman and Flier 1996). Furthermore, mutations in the genes encoding leptin and its receptor are known to cause obesity and diabetes in mice (Zhang *et al* 1994, Chen *et al* 1996a) and further work has revealed that congenital leptin deficiency can cause early onset obesity in humans (Montague *et al* 1997). Other important endocrine roles of leptin include regulation of immune function, angiogenesis and haematopoiesis (Kershaw and Flier 2004). Recent years have seen the discovery of many other adipocyte-secreted factors. Adiponectin is involved in insulin sensitisation of tissues including muscle, liver and adipose; low levels are detected in obese individuals and are also associated with insulin resistance and atherosclerosis (Fu *et al* 2005). Adiponectin is therefore considered an interesting candidate for a link between obesity and the high levels of obesity-induced insulin resistance and ensuing type II diabetes (Hutley and Prins 2005). The levels of visfatin, a newly identified adipokine, increase during the development of visceral obesity. Visfatin was found to mimic the effects of insulin (i.e. glucose uptake) on insulin-responsive tissues, and it's expression may be controlled by insulin resistance-inducing hormones, making it another potential factor involved in the link between adipose tissue, obesity and insulin resistance (Fukuhara *et al* 2005, Krailsch *et al* 2005). Adipose tissue also produces immune-related factors such as the pro-inflammatory cytokine TNFα which may also be associated with obesity-induced insulin resistance (Kershaw and Flier 2004), and factors associated with the cardiovascular system including angiotensinogen, which plays a central role in blood pressure regulation (Trayhurn and Beattie 2001).

The characterisation of these adipokines demonstrates that adipose tissue plays a major role in controlling multiple processes including glucose homeostasis, immunological functions and the cardiovascular system. Obesity, or the excess of adipose tissue, is associated with major health risks, such as insulin resistance, type II diabetes, heart disease and hypertension (collectively known as the metabolic syndrome or syndrome X), conditions that may be potentiated by the misexpression of adipose-derived adipokines

49

(Hutley and Prins 2005). Ongoing studies of the endocrine nature of adipose tissue are therefore required to combat the predominance of obesity in western society today.

1.6 Molecular control of adipogenesis

1.6.1 Cellular models of adipogenesis

The molecular events and external stimuli controlling commitment of embryonal precursors into the adipocyte lineage, and the differentiation of preadipocytes to adipocytes *in vivo*, are as yet unknown. Studying preadipocyte differentiation *in vivo* is extremely difficult as they are hard to distinguish from other fibroblastic cells or from other preadipocytes at different stages of differentiation (Ntambi and Young-Cheul 2000). Studying primary preadipocytes in culture is also problematic, largely because they have a limited lifespan (Ntambi and Young-Cheul 2000). Because of this, *in vitro* models have primarily been used to study adipogenesis. The establishment of preadipocyte cell lines in the 1970s greatly aided studies of the molecular processes involved in adipogenesis. The most commonly used models are the 3T3-L1 and 3T3-F442A cell lines, which represent preadipocytes committed to the adipocyte lineage, but are aneuploid (Green and Kehinde 1974, 1975, 1976). These were derived from disaggregated mouse Swiss embryos, through selection for cells with susceptibility to form lipid-filled adipocyte-like cells. When injected into mice 3T3-F442A cells were shown to produce fat pads indistinguishable from normal adipose tissue; such preadipocyte cell lines are therefore thought to be reliable models of *in vivo* adipogenesis (Green and Kehinde 1979). Murine multipotent cells lines also exist, such as C3H10T1/2, which can be induced to differentiate into several lineages including adipocytes, but these are less frequently used than preadipocytes (Reznikoff *et al* 1973, Taylor and Jones 1979). Differentiation of preadipocytes to adipocytes in culture will occur spontaneously over a period of 2-4 weeks. However, the addition of a specific combination of hormones and drugs enhances their rate of differentiation (Rubin *et al* 1977). Thus, after hormonal stimulation of post-confluent, growth arrested cells, a morphological change from elongated to rounded shape is observed. Markers of mature adipocytes can be seen by day 3, and by day 7 lipid-laden, rounded adipocytes are abundant (Ntambi and Young-Cheul 2000).

The classic combination of reagents used to induce adipogenesis includes insulin, dexamethasone and isobutylmethylxanthine (IBMX). Their use arose from studies investigating substances with the ability to increase the rate and/or extent of triacylglycerol accumulation in the 3T3-L1 and 3T3-F442A lines, which was at that stage the only marker known for adipogenic differentiation (Green and Kehinde 1975, Russell and Ho 1976, Rubin *et al* 1978). For the most part, the methods by which these substances stimulate adipogenesis are still relatively unknown, but several mechanisms have been described and are summarised in Figure 1.4. Insulin is a hormone secreted by the pancreas which has well known effects on metabolism, including enhancing fat accumulation (Green and Kehinde 1975). It is now thought that as well as enhancing fat accumulation *in vivo*, insulin may promote differentiation by activating signalling through kinases such as PKB/AKT and PDK1 (Magun *et al* 1996, Yin *et al* 2005). Dexamethasone, a synthetic glucocorticoid, activates the glucocorticoid receptor pathway (Rubin *et al* 1978). It stimulates expression of C/EBPδ which activates adipogenesis, and downregulates DLK1 and TGFβ which inhibit the process. In addition, Cushing's syndrome, a disorder giving rise to enhanced glucocorticoid levels in the bloodstream, is associated with visceral obesity, providing evidence for a physiological role for glucocorticoids in inducing adipocyte differentiation (Rosen and Spiegelman 2000). IBMX is a synthetic inhibitor of cAMP phosphodiesterases, and raises intracellular cAMP concentrations (Russell and Ho 1976). IBMX activates C/EBPβ and CREB, and may inhibit Wnt signalling. A modification of this cocktail was made for differentiating hMSCs into adipocytes: indomethacin, a prostaglandin synthesis inhibitor was added (Williams and Polakis 1977). It is thought that indomethacin may be a ligand for PPARγ, which could account for its adipogenic effects (Lehmann *et al* 1997).

1.6.2 Transcriptional control of adipogenesis

1.6.2.1 The core transcriptional cascade

Figure 1.4 summarises the major events in the control of adipogenesis. The first key regulators of adipogenesis to be expressed are two members of the C/EBP family of basic leucine zipper transcription factors, C/EBPβ and –δ; this occurs within a few hours of hormonal stimulation (Tang and Lane 1999). Mice lacking both the C/EBPβ and –δ genes showed abnormally small fat pads, indicating the importance of these proteins for
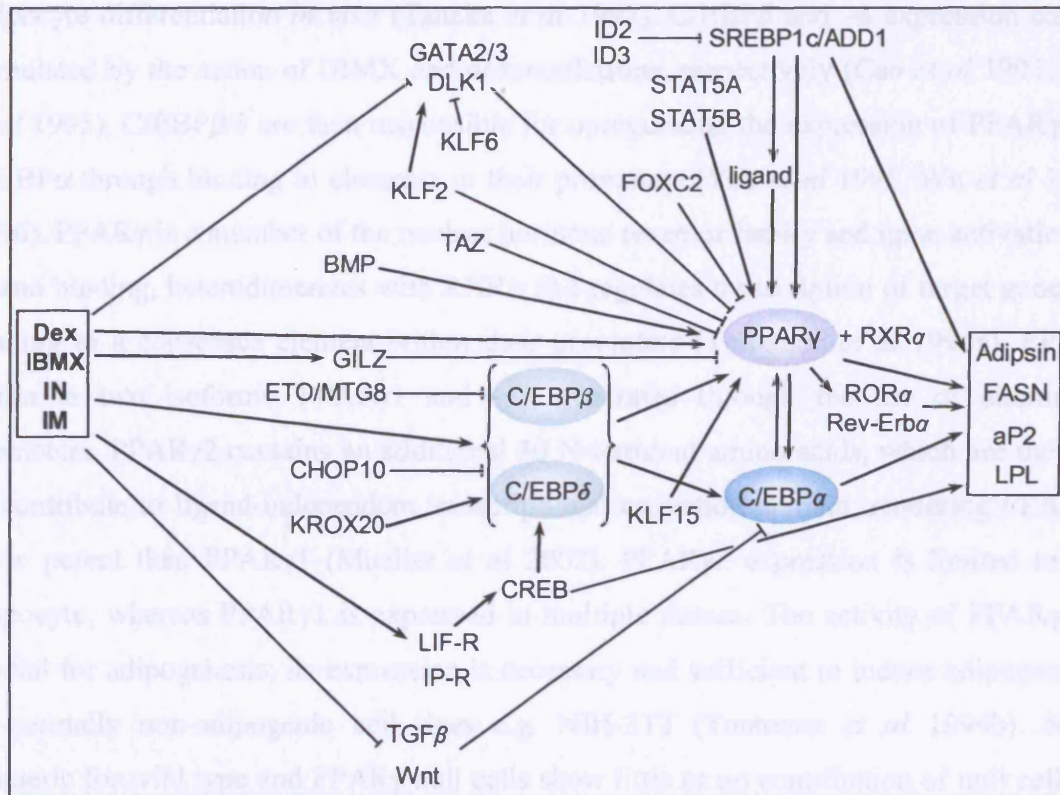
Figure 1.4. Major events of the transcriptional regulation of adipogenesis. Adipogenesis is stimulated by a specific induction cocktail which initiates a complex regulatory network leading to the upregulation of adipocyte specific genes such as adipsin, aP2, FASN and LPL. The main controller of this network is PPARγ, which is induced by C/EBPβ and –δ, and it's expression is maintained by C/EBPα. Numerous factors upstream and downstream of this core network are involved. The majority of these events are likely to represent later regulation of adipogenesis as they were shown to control differentiation of preadipocytes that are committed to the adipocyte lineage.

adipocyte differentiation *in vivo* (Tanaka *et al* 1997). C/EBPβ and –δ expression can be stimulated by the action of IBMX and dexamethasone, respectively (Cao *et al* 1991, Yeh *et al* 1995). C/EBPβ/δ are then responsible for upregulating the expression of PPARγ and C/EBPα through binding to elements in their promoters (Yeh *et al* 1995, Wu *et al* 1995, 1996). PPARγ is a member of the nuclear hormone receptor family and upon activation by ligand binding, heterodimerizes with RXRα and regulates transcription of target genes by binding to a consensus element within their promoters (Tontonoz *et al* 1994a). PPARγ exists in two isoforms, PPARγ1 and -γ2, generated through the use of alternative promoters. PPARγ2 contains an additional 30 N-terminal amino acids, which are thought to contribute to ligand-independent transcriptional activation domain, rendering PPARγ2 more potent than PPARγ1 (Mueller *et al* 2002). PPARγ2 expression is limited to the adipocyte, whereas PPARγ1 is expressed in multiple tissues. The activity of PPARγ2 is crucial for adipogenesis; its expression is necessary and sufficient to induce adipogenesis in normally non-adipogenic cell lines e.g. NIH-3T3 (Tontonoz *et al* 1994b). Mice chimeric for wild type and PPARγ null cells show little or no contribution of null cells to adipose tissue (Rosen *et al* 1999, Barak *et al* 1999). Additionally, RNAi-mediated (1.7.2) suppression of PPARγ expression in 3T3L1 cells blocked their differentiation (Hosono *et al* 2005). C/EBPα is expressed concomitantly with PPARγ during adipogenesis, and these proteins act in a positive feedback loop to maintain each other's expression throughout differentiation (Wu *et al* 1999). PPARγ is responsible for directly activating the transcription of genes characteristic of mature adipocytes, or additional transcription factors that give rise to adipocyte-specific gene expression. Mature adipocyte markers include aP2, adipsin, LPL, FASN and perilipin. Although capable of activating transcription of a subset of adipocyte-specific genes, it is thought that the main role of C/EBPα is to sustain PPARγ expression (Rosen *et al* 2002).

A key event that is absolutely required for terminal differentiation into adipocytes is growth arrest. It is thought that preadipocyte cell lines *in vitro* undergo one or two rounds of mitotic clonal expansion. This begins after an initial growth arrest induced by contact inhibition, and ceases concomitant with expression of C/EBPα and PPARγ (Rosen and Spiegelman 2000). This second growth arrest is permanent, and is likely to be brought about directly by the action of these proteins. PPARγ can induce expression of the cyclin-

dependent kinase inhibitors p18 and p21 (Morrison and Farmer1999), and also inhibits E2F/DP-DNA binding activity through downregulation of PP2A (Altiok *et al* 1997). C/EBPα may also play a role in inducing p21 (Rosen and Spiegelman 2000). Additionally, the retinoblastoma protein Rb regulates adipogenesis; it interacts with C/EBPβ during the early stages and enhances DNA binding activity (Chen *et al* 1996b, Cole *et al* 2004), but is later inhibitory to PPARγ's transactivation function (Fajas *et al* 2002).

1.6.2.2 Activators of adipogenesis

A vast array of factors are known to play a role in promoting adipogenesis; these act at different stages of the process, mostly feeding into or out of the core transcriptional network; here, a summary of regulatory mechanisms is provided. Autocrine and/or paracrine signalling mediated by LIFR and IPR positively regulates adipogenesis (Aubert *et al* 1999, 2000). Stimulation of these receptors activates the ERK/MAPK signalling cascade, leading to phosphorylation and activation of the transcription factor CREB, which then activates transcription of C/EBPβ and -δ (Belmonte *et al* 2001). Additionally, IPR stimulates PKA activity by increasing cAMP concentrations, and this again leads to C/EBP expression through phosphorylation and activation of CREB (Belmonte *et al* 2001). CREB can also directly activate the transcription of several adipocyte specific genes, including aP2 and FASN (Reusch *et al* 2000). KROX20 and KLF6 have recently been identified as regulating early events of adipogenesis; KROX20 promotes 3T3-L1 adipogenesis through C/EBPβ dependent and independent mechanisms (Chen *et al* 2005), and KLF6 represses the inhibitory protein DLK1 that is expressed in preadipocytes (Li *et al* 2005). Many positive adipogenic regulators act by increasing PPARγ activity. Members of the STAT family of transcription factors, specifically STAT5A and STAT5B, are expressed just prior to PPARγ during 3T3L1 adipogenesis and may be involved in the production of PPARγ ligands (Floyd and Stephens 2003). Additionally, the helix-loop-helix transcription factor SREBP1c is induced during adipogenesis at around the same time as PPARγ, and can activate its transcription as well as the transcription of adipocyte-specific genes through promoter E-box motifs (Fajas *et al* 1999). It is also likely that SREBP1c can further activate PPARγ by production of an endogenous ligand, thus enhancing adipogenesis (Kim *et al* 1998). KLF15 acts synergistically with C/EBPα to activate PPARγ expression during preadipocyte differentiation (Mori *et al* 2005a). Late

adipogenic activators include RORγ and Rev-Erbα, which are targets of PPARγ and mediate its effects by activating the transcription of subsets of adipocyte-specific genes (Austin *et al* 1998, Fontaine *et al* 2003).

Several signalling pathways have been shown to positively regulate adipogenesis, including the BMP signalling pathway. BMP ligands bind a subset of receptors which transduce signals through SMAD proteins 1, 5 and 8 to control transcription in the nucleus. The effects of BMPs 2, 4 and 7 on adipogenesis have been studied in mice, where addition of these ligands in low doses to the multipotent C3H10T1/2 cell line induced adipogenesis; this effect was however dose-dependent, as higher doses led to osteogenic and chondrogenic differentiation (Ahrens *et al* 1993, Wang *et al* 1993, Gimble *et al* 1995, Asahina *et al* 1996, Tang *et al* 2004). The effect of BMP2 may be mediated through BMPR-1A rather than BMPR-1B, as a constitutively active BMPR-1A led to an increase in PPARγ levels (Chen *et al* 1998).

1.6.2.3 Inhibitors of adipogenesis

Many inhibitors of adipogenesis are expressed in preadipocytes and must be downregulated for differentiation to proceed. These include the transmembrane protein DLK1 and transcription factors GATA2 and -3, which have been shown by overexpression studies to block adipogenesis (Tong *et al* 2000, Lee *et al* 2003); GATA2 and -3 may perform this function by suppressing transcription from the PPARγ promoter, as well as forming repressive protein complexes with C/EBPα and -β (Tong *et al* 2005). GILZ was recently identified as an inhibitor of adipogenesis in C3H10T1/2 cells, and also acts through binding and repression of the PPARγ promoter (Shi *et al* 2003). KLF2 inhibits preadipocyte differentiation via repression of PPARγ expression, and also partly through upregulation of DLK1, but has no effect on the formation of preadipocytes from murine precursor cells (Banerjee *et al* 2003, Wu *et al* 2005). The transcription factor FOXC2 blocks adipogenesis, likely through inhibiting PPARγ transactivation of selected promoters (Davis *et al* 2004). Protein-protein interactions are a common mechanism of adipogenic inhibition. CHOP10 is an inhibitory member of the C/EBP family that can form dominant negative dimers with other C/EBP proteins and inhibit their transcriptional activity (Tang and Lane 2000, Pereira *et al* 2004). Interestingly, the translocation typically

found in myxoid and round cell liposarcomas involves the fusion of the basic leucine zipper domain of CHOP10 to FUS or EWS transcription factors (Crozat et al 1993). The inhibitor of DNA binding (ID) family of nuclear proteins also have the helix-loop-helix structure of some transcription factors, but lack the basic DNA binding domain, and form non-functional heterodimers with other helix-loop-helix proteins, for ID2 and -3 this includes SREBP1c (Moldes et al 1997, 1999).

The Wnt and TGFβ signalling pathways have long been known as crucial regulators of the development of mesenchymal tissues (Roelen and Dijke 2003), and their roles in development of adipose tissue has been investigated. Wnt signalling is inhibitory to adipogenesis in 3T3-L1 preadipocytes; this is due to suppression of PPARγ and C/EBPα, rather than C/EBPβ and -δ, but the exact mechanism by which this occurs is unknown (Ross et al 2000, Bennett et al 2002). Transgenic mice expressing Wnt10b under the control of an adipose-specific promoter showed a significant reduction in adipose tissue mass (Longo et al 2004), and when β-catenin was deleted in the mouse uterus, a switch from smooth muscle to adipose tissue was seen (Arango et al 2005). Human MSCs express various Wnt signalling components (Etheridge et al 2004), and an inhibitory effect of Wnt signalling on hMSC adipogenesis has also been shown (De Boer et al 2004). The TGFβ pathway inhibits adipogenesis of human and rodent MSCs as well as preadipocytes (Ignotz and Massague 1985, Locklin et al 1999), and TGFβ overexpression in transgenic mice prevents adipose differentiation (Clouthier et al 1997). Inhibition of 3T3L1 adipogenesis may be mediated by activin A, a TGFβ-family ligand (Hirai et al 2005), amongst others. The TGFβ pathway inhibits adipogenesis in a similar manner to CHOP10, in that SMAD3 binds to and sequesters the function of the C/EBP proteins (Choy et al 2000, 2003).

1.6.2.4 The earliest events of adipogenesis

The vast majority of this research was carried out using the preadipocyte models discussed in section 1.6.1. As implied by their name, these cells are committed to the adipocyte lineage, so the transcriptional networks described control the later, post-commitment stages of adipogenesis. Thus, there is little information on the earliest mechanisms involved in stem cell differentiation into adipocytes prior to commitment to this lineage,

and how this relates to the later regulatory networks described. The earliest events in the transcriptional cascade are of crucial importance, as they set in motion the multitude of events that lead to a specific change in phenotype. A detailed knowledge of the earliest mechanisms by which adipogenesis is brought about may be extremely valuable in manipulating the differentiation of hMSCs, as well as providing a more detailed knowledge of the events involved in stem cell differentiation.

The knowledge that hMSCs differentiate into adipocytes, and the ability to do this in culture, is fairly new relative to the preadipocyte models. However, some detail is emerging about events and factors that may regulate early differentiation events, in particular those that control the decision to differentiate along a specific lineage. TAZ, a transcriptional coactivator, was recently found to enhance osteogenesis but inhibit adipogenesis through modulation of transcription by the main osteogenic (CBFA1) and adipogenic factors (PPARγ) (Hong *et al* 2005). Interestingly, cell shape and cytoskeletal tension was found to be crucial for hMSC lineage commitment choice (McBeath *et al* 2004). A less spread shape with lower cytoskeletal tension favoured adipogenesis, whereas cells that were more spread out so had higher cytoskeletal tension underwent osteogenesis. These events were mediated by the Rho GTPase RhoA and it's effector Rho kinase; Rho activity induced osteogenesis and a dominant negative RhoA induced adipogenesis (McBeath *et al* 2004). In a study using MEFs, Rho activity was shown to control the decision between adipogenesis and myogenesis by mediating lineage-specific responses to IGF-1 signalling rather than in response to cytoskeletal tension. Similar to the McBeath study, a constitutively active form of Rho blocked adipogenesis whereas dominant-negative Rho blocked myogenesis (Sordella *et al* 2003). Finally, Jakkaraju *et al* identified three isoforms of a novel gene which were regulated by mechanical stretch; a condition that induces myogenic differentiation in embryonic mesenchymal cells (Jakkaru *et al* 2005). TIP-1, induced by stretch, has pro-myogenic activity, whereas stretch inhibited TIP-3, a factor that promotes adipogenesis. These proteins had chromatin remodelling activity, and it was shown that remodelling at the promoters of lineage-specific genes contributed to the observations of lineage choice (Jakkaru *et al* 2005).

1.7 Techniques for functional genomics

Functional genomics has been described as the study of gene function through parallel expression measurements of a genome (Butte *et al* 2002). Although sequencing of the human genome has provided a wealth of information regarding the sequence and structure of human genes, knowledge of the function of these genes remains far from complete (Lander *et al* 2001, Venter *et al* 2001). Techniques have been developed that can assess the expression and/or function of many genes in a single assay; these include gene expression microarrays and RNA interference.

1.7.1 Microarray analysis of gene expression

The concept of microarray technology was developed when it was noted that the use of fluorescent labels allowed sensitive detection of a specific target analyte. This meant that "microspots" of "capture agent" (now known as probe), only a few μm in diameter were sufficient so multiple microspots detecting different analytes could be grouped together on a microarray and analysed simultaneously (Ekins and Chu 1999). Although this technology was originally developed for antibody analysis in immunodiagnostics, the authors speculated that it could be used for many other ligand binding assays (Ekins and Chu 1991). Indeed, in the mid-1990s the production of DNA microarrays, and their use to monitor gene expression in *Arabidopsis thaliana*, was described (Schena *et al* 1995). This microarray contained 45 cDNAs; the continuing development of this technology, combined with the sequencing of entire genomes, now allows the simultaneous analysis of the expression of every known gene in an organism.

1.7.1.1 Microarray design

The fundamental basis of microarray technology is the hybridisation of labelled RNA or DNA targets derived from biological samples to a collection of DNA molecules that are immobilised on a solid surface. These immobilised DNA molecules, or probes, each represent a specific gene or transcript and the fluorescence signal created by hybridisation of the complementary target molecule gives an expression readout for each gene. Two types of DNA microarray exist: cDNA arrays and oligonucleotide arrays. cDNA arrays are produced by the automated spotting onto glass slides of cDNA molecules, produced from

PCR amplification of specific targets or from library clones, which are typically hundreds of base pairs in length. They have the advantage that they can be custom-designed, but tend to contain fewer probes than oligonucleotide arrays so permit analysis of expression of fewer genes; nevertheless, this number can still be in the thousands (Choudhuri 2004). Oligonucleotide array technology is used for Affymetrix Genechips (www.affymetrix.com). These arrays are synthesised via the controlled sequential addition of nucleotides to the surface of glass slides. This process builds up 25mer oligonucleotide probes designed to be complementary to a region of a specific gene transcript (Affymetrix 2004). Where possible, probe design is biased toward the 3' region of each transcript, as the method used to synthesise target molecules occurs from this end, so uneven amplification from transcripts of different lengths would give misrepresentative results if many 5' probes were included. Eleven probe sequences are designed for each transcript. Each probe sequence is represented by a perfect match (PM) and mismatch (MM) probe, where the PM probe is exactly complementary to the target transcript but the MM probe contains one mismatched base at position 13. The 11 pairs of PM and MM probes for a single transcript are collectively called a probeset (Fig. 1.5). Inclusion of PM and MM probes allows for calculation of background and non-specific hybridisation. Each probe is present in multiple copies located in a specific cell on the array, thus allowing for the sensitive analysis of expression of each gene (Knudsen 2004).

Target RNA molecules can be synthesised and labelled in either one or two colour format. In two-colour format labelling, RNA is extracted from two biological samples and used as a template for cDNA synthesis. Each cDNA sample is labelled with different coloured dyes, normally Cy3 and Cy5 as with two-colour arrays produced by Agilent Technologies (www.agilent.com), and a comparison is made by hybridising the samples to the same array and measuring colour change to detect over- or under-expressed genes in each sample. This format is mostly applicable to spotted cDNA arrays (Stears et al 2003). Oligonucleotide arrays, more specifically Affymetrix Genechips, generally use one-colour target labelling. RNA is extracted from each biological sample, reverse transcribed, then cRNA synthesis is performed using biotin-labelled ribonucleotides. Biotin-labelled cRNA probes are hybridised to the arrays, which are then washed and stained with a streptavidin-phycoerythrin conjugate and scanned. The level of fluorescence emitted by a particular
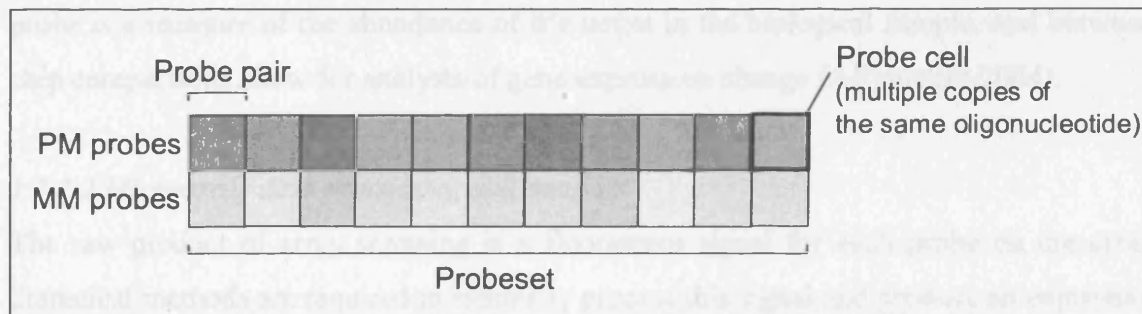
Figure 1.5. Design of an Affymetrix probeset on the U133 Plus 2.0 GeneChip Array. A probeset consists of 11 probe pairs. Each probe pair contains a perfect match (PM) probe (25 bases in length) that exactly matches the target sequence, and a mismatch (MM) probe, which contains one mutation at base 13. Each probe is present in multiple copies within a probe cell. This design results in sensitive quantification of target abundance, represented by the different levels of shading in each probe cell (dark shading represents high levels of target hybridisation to its probe – note that some level of hybridisation can also occur on the MM probes).

probe is a measure of the abundance of it's target in the biological sample, and between-chip comparisons allow for analysis of gene expression change (Affymetrix 2004).

1.7.1.2 Microarray data processing and analysis

The raw product of array scanning is a fluorescent signal for each probe on the array. Statistical methods are required to rationally process this signal and produce an expression value for each transcript that truly reflects it's abundance in the original sample. Further techniques are employed to identify trends in the data through comparison of expression values across arrays.

Pre-processing of the raw signal image to produce expression values is performed in three steps: background correction (or image processing), normalisation and summarisation of the expression values for each probe in a probeset to give a single value. Background correction is designed to minimise the effects of variation caused by technical artefacts such as uneven hybridisation across an array. Normalisation is performed to take into account variation across multiple arrays caused by experimental variation (i.e. variation that was not present in the biological sample), and is often performed by fitting the data from each array to a single statistical model. This allows comparison of expression values between arrays. Finally, expression values for each probeset on each array are estimated, often using $\log_2$ scale values (Allison *et al* 2006). Several methods have been developed to perform these functions on Affymetrix microarray data, including MAS5 (Affymetrix 2001), MBEI (Li and Wong 2001) RMA (Irizarry *et al* 2003a), and GC-RMA (Wu and Irizarry 2004).

Once data pre-processing is complete, analytical methods can be employed to identify differential expression of genes between biological samples. Typically, modified t-tests or ANOVA are used to calculate significant changes in expression of a gene between two conditions, given that replicates of each condition have been performed (Knudsen 2004). This will produce a fold-change value with an associated p-value for a gene between conditions, and allows a cut-off to be set in order to define significantly changed genes. Significance at the 0.05 level, and/or a fold change greater than 2, are frequently used. Importantly, when analysing microarray data thousands of hypothesis tests will be

performed as each array represents thousands of genes. Multiple-testing correction methods have been developed that decrease the likelihood of false positive results, including the highly stringent Bonferroni correction, and False Discovery Rate correction, which is less severe and more biologically relevant (Benjamini and Hochberg 1995).

Classification of genes or arrays in terms of similarity in expression profiles is a commonly used method for identifying biological trends and relationships within a microarray dataset. Supervised classification entails designing algorithms that best assign objects (i.e. genes) to pre-defined categories, and assumes prior knowledge of these categories. Unsupervised classification methods assume no prior knowledge and categories are developed during the analysis (Allison *et al* 2006). Unsupervised methods include hierarchical cluster analysis, which forms groups of genes based on similarity of expression profiles and represents these relationships in a dendrogram (Eisen *et al* 1998). The Euclidean distance is often used to measure distance (or degree of dissimilarity) between genes. Other cluster analysis methods include self-organising maps and k-means clustering (Knudsen 2004).

Gene annotation and pathway analysis are increasingly popular methods used to relate gene expression change to biological responses. This form of analysis is becoming possible due to the development of microarrays that assay whole transcriptomes, allowing a comprehensive view of entire pathways or functional classes. Gene Ontology is a structured vocabulary which attempts to unify gene product descriptions across species and databases (Ashburner *et al* 2000). It can be used to extract information regarding the function of a group of co-expressed genes in a microarray experiment. Similarly, up- or downregulation of genes in a specific pathway can be investigated using searchable pathway maps, and statistical analysis of these annotations can attach a significance value to the findings (Curtis *et al* 2005).

1.7.2 The use of microarrays to study adipogenesis

As well as a gene-by-gene approach to investigating adipogenic regulation (1.6.2), microarrays have been used to simultaneously analyse the expression of thousands of genes at specific stages of differentiation. These analyses identified many potential novel

regulators of differentiation, and in some cases provided a global characterisation of the processes involved in adipogenesis.

The first such study used cDNA microarrays representing around 5000 cDNA clones to compare the expression profiles of primary human preadipocytes before and after differentiation *in vitro* (Zhou *et al* 1999). A similar study was performed but in greater detail, as it utilised a larger sample size and arrays containing 17000 cDNAs, and also used mature primary adipocytes for comparison rather than *in vitro*-differentiated preadipocytes (Urs *et al* 2004). Both groups identified a list of genes that were differentially expressed between conditions, including some novel genes which were concluded to be potentially important for preadipocyte differentiation. In 2001 Soukas *et al* published a comprehensive study in which Affymetrix murine microarrays representing 11,000 transcripts were used to investigate the expression profiles of primary preadipocytes and adipocytes and compare them to undifferentiated and differentiated 3T3-L1 cells. It was found that while many of the gene expression changes associated with *in vivo* and *in vitro* differentiation overlapped, distinct subsets of genes were only expressed in one or the other condition, revealing that 3T3-L1 cells do not completely mimic the fully differentiated adipocyte phenotype (Soukas *et al* 2001).

Microarray studies, mainly using 3T3-L1 cells, have been used obtain temporal expression profiles of adipogenesis in order to elucidate molecular regulation over the differentiation time-course. The study by Soukas *et al* also followed a differentiation time-course, and included 10 timepoints ranging from pre-confluent to 28 day differentiated cells (Soukas *et al* 2001). Other studies utilised the same murine microarrays and focused on the first day of 3T3-L1 differentiation (Burton *et al* 2002, 2004). The earliest timepoint studied was 2h; in one experiment an inhibitor of adipogenesis was used to identify and eliminate gene expression responses to the differentiation induction cocktail that were not specific to differentiation (Burton *et al* 2004). In each of these studies, different arrays, data normalisation, clustering methods and significance cut-offs were used. Although in every study the main adipogenic regulators were identified, the overall lists of significantly changed genes differed. However, each investigation contributed towards the understanding of preadipocyte differentiation, and achieved their aim of identifying targets

for future research. Gene expression changes caused by known adipogenic regulators, such as Wnt and PPARγ ligands, have also been investigated in 3T3-L1s in order to identify downstream targets of these factors (Ross *et al* 2002, Gerhold *et al* 2002). A comprehensive investigation of 3T3L1 differentiation has recently been published, which used large arrays and advanced analysis techniques to provide a global overview of the molecular processes of adipogenesis, as well as identifying interesting candidates for further work (Hackl *et al* 2005). This study highlighted the ability of microarray analysis combined with the latest annotation methods to provide a detailed insight into differentiation.

In addition to the 3T3L1 work, microarray analyses of hMSC adipogenesis have been performed. The earliest study, performed in 2003, used cDNA arrays representing 3400 clones to study adipogenesis over a 14-day time-course, with the earliest timepoint during differentiation at day 1 (Nakamura *et al* 2003). Sekiya *et al* used Affymetrix Genechips containing around 12,000 genes, and similarly chose the day 1 timepoint as the earliest point of study, but carried differentiation through until day 21 (Sekiya *et al* 2004). Both studies identified genes upregulated at day 1 that had been shown to be repressed during 3T3-L1 differentiation, leading to the conclusion that they may be involved in stages of differentiation occurring prior to the committed stage that preadipocytes represent (Burton and McGehee 2004). A more detailed analysis of the data from Sekiya *et al* has been performed (Ylostalo *et al* 2006). Hierarchical clustering was used to group genes according to their expression pattern during differentiation, and novel downstream targets of PPARγ and C/EBPα were identified via selection of genes that were co-expressed in clusters with known targets. Their target status was then confirmed via identification of PPARγ or C/EBPα binding sites in their promoter (Ylostalo *et al* 2006). Work published by Hung *et al* used cDNA arrays comprising 7,500 cDNAs/ESTs to compare undifferentiated hMSCs with day 3-differentiated cells (Hung *et al* 2004a), and identified numerous genes which had not before been associated with adipogenesis. A recent study of bovine MSC adipogenesis also identified numerous differentially expressed genes across a time-course; gene expression was analysed after 6h of differentiation, but some discrepancies such as the lack of upregulation of PPARγ were noted (Tan *et al* 2006). These studies highlighted the importance of studying adipogenesis from the uncommitted

hMSC stage, and revealed their potential to increase the understanding of events controlling early adipogenesis. However, no study of human adipogenesis has analysed gene expression during the first day of differentiation. Furthermore, the use of small arrays and few functional annotation methods means that much detail of the earliest events of hMSC adipogenesis has yet to be described.

1.7.3 RNAi and it's application to study gene function

1.7.3.1 Mechanism of RNAi

The term RNA interference (RNAi) was first used by Fire and co-workers, who noted that the introduction of dsRNA into *Caenorhabditis elegans* elicited a potent sequence-specific inhibition of gene expression (Fire *et al* 1998). In fact, this phenomenon had previously been noted in plants when introduction of exogenous pigmentation genes resulted in their silencing as well as occasional silencing of the endogenous gene (van der Krol *et al* 1990, Napoli *et al* 1990). It is now known that RNAi is an evolutionarily conserved mechanism that operates in eukaryotes, possibly as a defence mechanism against transposons or replicating viruses (Shi 2003). A similar pathway controls RNAi in all organisms. Entry of dsRNA into a cell, experimentally or as a consequence of natural phenomena such as viral infection, results in cleavage by an RNase III-like enzyme termed Dicer (Figure 1.6). The products are 21-23nt small interfering dsRNAs (siRNA) that have a 2-nucleotide 3' overhang. These are incorporated into the RNA-induced silencing complex (RISC), composed of multiple protein subunits. The helicase activity of RISC results in unwinding of the RNA duplex, after which the antisense strand is retained within RISC and guides the protein complex to the target mRNA. Through RISC endonuclease activity, the target mRNA is cleaved and degraded, resulting in silencing of the gene with homologous sequence to the siRNA (Zou and Yoder 2005, Sledz and Williams 2005).

1.7.3.2 Application of RNAi in mammalian cells

Although long dsRNA can be used effectively to silence gene expression in organisms such as *C. elegans* and *Drosophila*, dsRNAs longer than 30nt trigger an antiviral interferon response mechanism in mammalian cells that results in non-specific degradation of transcripts and shutdown of protein synthesis in the cell (Williams and Haque 1997). This was overcome by the production of synthesised versions of the Dicer
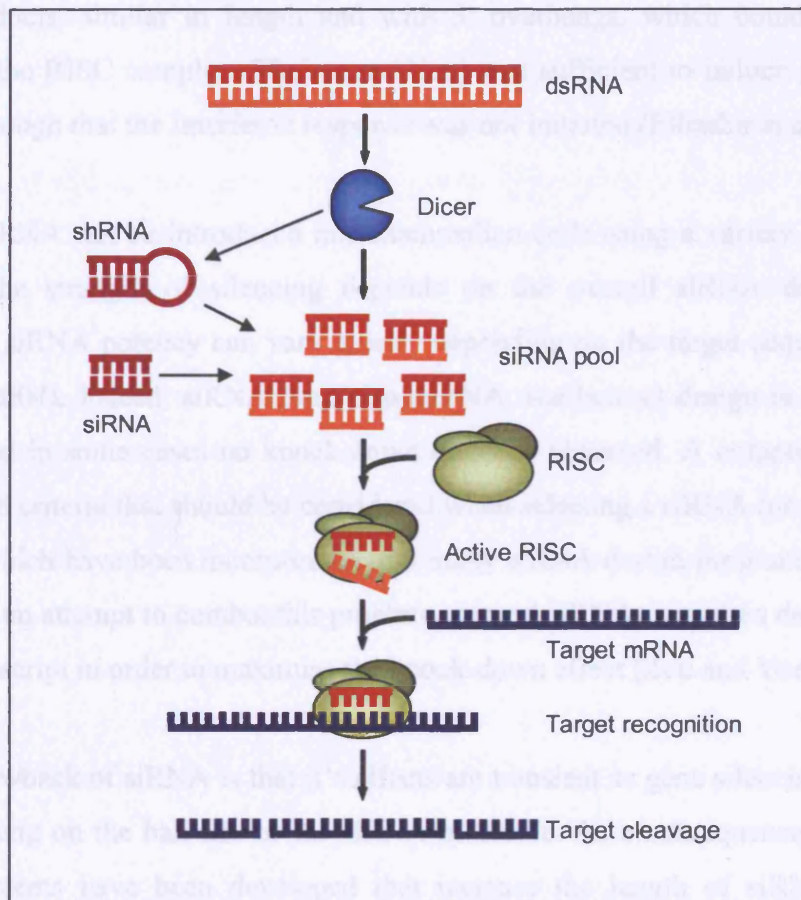
Figure 1.6. Mechanism of RNA interference. Long dsRNA are processed by Dicer into 21-23nt siRNAs. Synthetic siRNAs and shRNAs, which are probably nicked by Dicer in the loop structure to produce double stranded RNA, mimic these siRNAs and can substitute as substrates for RISC. RISC binds the siRNAs and after removal of one RNA strand, converts into it's active form. The ssRNA guides active RISC to the target mRNA, to which it binds in a sequence-specific manner and induces target mRNA cleavage by RISC endonuclease activity. Figure adapted from www.cenix-bioscience.com.

siRNA products, similar in length and with 3' overhangs, which could be processed directly by the RISC complex. Their size (21nt) was sufficient to induce gene silencing, but small enough that the interferon response was not initiated (Elbashir *et al* 2001).

Synthetic siRNA can be introduced into mammalian cells using a variety of transfection methods. The strength of silencing depends on the overall siRNA dose, but more importantly siRNA potency can vary greatly depending on the target sequence (Hannon and Rossi 2004). Indeed, siRNA (and also shRNA, see below) design is still not 100% efficient, and in some cases no knock-down effect is observed. A comprehensive study determined 8 criteria that should be considered when selecting a siRNA (or shRNA) target sequence, which have been incorporated into many siRNA design programs (Reynolds *et al* 2004). In an attempt to combat this problem, several siRNAs are often designed against a single transcript in order to maximise the knock-down effect (Zou and Yoder 2005).

A major drawback of siRNA is that it's effects are transient so gene silencing is gradually lost, depending on the half-life of the siRNA molecule. As a consequence, vector-based delivery systems have been developed that increase the length of siRNA expression (Brummelkamp *et al* 2002a, 2002b). An RNA polymerase III promoter, such as H1 or U6, drives expression of a small hairpin RNA molecule (shRNA) that mediates RNAi. shRNAs are created by cloning a 19-29nt sense strand, followed by a spacer sequence, followed by the 19-29nt antisense strand, downstream of the polymerase III promoter. The spacer allows base-pairing within the transcribed shRNA molecule and consequently forms a hairpin loop which is probably nicked by Dicer to form a siRNA-like structure. Polymerase III promoters are suitable for driving shRNA expression as they generate short messages that terminate precisely and leave a 3' overhang, a feature required for shRNA function (Brummelkamp *et al* 2002a). Viral-mediated shRNA delivery results in longer-term expression due to viral integration and stable expression of shRNA from the host genome. Retroviral, adenoviral and lentiviral vector systems are now available containing shRNA expression cassettes combined with selectable markers that allow for long-term shRNA expression and selection of cells containing the viral construct (Brummelkamp *et al* 2002b, Abbas-Terki *et al* 2002, Shen *et al* 2003, Clements *et al* 2006).

1.7.3.3 RNAi libraries

Genome sequencing in various organisms followed by identification of the encoded genes has allowed RNAi to be used on a large scale to investigate gene function. Simultaneous knock down of the expression of multiple genes followed by assay for loss of function provides a high throughput method to identify genes involved in a specific process. The creation of RNAi libraries has augmented this approach to functional genomics.

An early high throughput screen in *C. elegans* involved feeding the nematode worms with different bacterial strains, each engineered to express a single dsRNA oligonucleotide, which constitute a library targeting 86% of all *C. elegans* genes (Ashrafi *et al* 2003). Visualisation of changes in fat storage by staining with the lipophilic dye Nile Red identified genes which either stimulate (0.7% of all genes) or inhibit (1.8%) the process, some of which had mammalian homologues (Ashrafi *et al* 2003).

Mammalian oligonucleotide siRNA libraries have been developed, and one study demonstrated their utility to identify genes involved in TRAIL-induced apoptosis (Aza-Blanc *et al* 2003) using a high-throughput screening strategy. A siRNA library directed against 510 genes, focusing on kinases, was designed and arrayed in microplate format so that each well contained a single siRNA oligonucleotide species. Reverse transfection of HeLa cells was performed and transfected cells were incubated with or without TRAIL to assay for increased or decreased cell death induced by knock-down of a TRAIL-associated gene. Both TRAIL-sensitisers (less apoptosis) and TRAIL inhibitors (more apoptosis) were identified (Aza-Blanc *et al* 2003).

As mentioned previously, siRNA-mediated gene knockdown has the drawback that it only produces a transient effect on gene expression. In the context of high-throughput screening, it is also an extremely expensive approach, as many siRNAs are synthesised and experiments must be replicated. Accordingly, in 2004 two groups described the creation of retrovirally delivered shRNA libraries containing shRNAs designed against human, or both human and mouse genes (Berns *et al* 2004, Paddison *et al* 2004). Paddison *et al* developed a shRNA library covering around 10000 human and 5000 mouse genes, with 2-9 shRNAs for each gene (Paddison *et al* 2004). Each 29nt shRNA was linked with

a 60nt "molecular barcode", which could be used to follow that fate of particular shRNAs in mixed populations using microarrays containing each barcode sequence as a probe. The efficacy of this library has been demonstrated in several studies; it was used to recover shRNAs involved in proteasome function by assaying for loss of function after shRNA introduction using a fluorescent reporter coupled to a degradation signal (Paddison *et al* 2004). In a mammalian screen for genes preventing epithelial cell transformation, a cell line susceptible to transformation by a single genetic event (through prior addition of transgenes, section 1.8.3) were transduced with the library and assayed for anchorage independent growth (Westbrook *et al* 2005). Anchorage-independent clones were pooled and enriched barcodes linked to shRNAs causing this colony outgrowth were identified. These shRNAs targeted several genes implicated in cancer pathogenesis, as well as a novel tumour suppressor, REST (Westbrook *et al* 2005). Recently, a shRNA library constructed using a doxycycline-inducible vector was used in a similar screening strategy, and identified a novel gene involved in the survival of a specific cancer subtype (Ngo *et al* 2006).

The shRNA library developed by Bernards and co-workers consists of three 19nt shRNAs targeted against each of 7914 human genes, cloned individually into a retroviral vector (Berns *et al* 2004). An early version of this library targeting 50 deubiquitination enzymes was shown to be efficient in a simple screening strategy; each shRNA was individually transfected into HEK-293 cells along with a NF-κB luciferase reporter construct. Knock-down of CYLD expression led to higher NF-κB transcription and therefore luciferase activity, identifying this gene as a novel NF-κB regulator (Brummelkamp *et al* 2003). The full retroviral library has been used in the more complicated selection strategy, where shRNA pools are used to transduce a cell population, and after exposure to a certain condition, shRNAs conferring a survival advantage in that condition permit cell outgrowth, allowing subsequent shRNA isolation by PCR, sub-cloning and re-screening, or by the use of the shRNA sequence as a molecular barcode (Berns *et al* 2004). This strategy led to the identification of novel regulators of p53-mediated proliferation arrest (Berns *et al* 2004), and has recently demonstrated that PITX1 is a novel suppressor of Ras activity and therefore tumourigenicity (Kolfschoten *et al* 2005). It is therefore clear that shRNA library screens can be used in a variety of experimental strategies, and are

extremely effective at identifying novel genes with a particular function or involved in a specific process.

1.8 Stem cells and cancer

1.8.1 Cancer stem cells

Cancer stem cells are currently defined in terms of functional attributes; thus, a cancer stem cell is a tumour cell that has the ability, after transplantation into immunodeficient mice, to form tumours containing all cell types of the tumour of origin, including further cancer stem cells (Polyak and Hahn 2005, Bjerkvig *et al* 2005).

It has been observed that only a small number of cancer cells are capable of extensive proliferation *in vitro* or *in vivo*, and large numbers of cancer cells must be transplanted into an animal model in order to initiate new tumours (Al-Hajj and Clarke 2004). These findings could be explained by every cancer cell having the ability to proliferate and form new tumour cells, with the probability of them doing so being small. An alternative explanation was the existence of a small subset of cancer cells with stem cell-like properties including an extensive self-renewal capacity, which is exclusive to this cell population (Reya *et al* 2001). The existence of cancer stem cells was definitively demonstrated in acute myeloid leukaemia (AML, Bonnet and Dick 1997). A rare subset of human AML cells, termed SCID leukaemia-initiating cells (SL-IC), was capable of proliferating and sustaining human AML after transplantation into NOD/SCID mice. Furthermore, SL-IC were capable of self-renewal in serial transplantation experiments. Crucially, these cells could be prospectively identified and purified as $CD34^+CD38^-$; only this population had self-renewal capability during serial transplantation so exhibited the defining characteristics of a cancer stem cell (Bonnet and Dick 1997).

Cancer stem cells have also been identified in tumours of breast and brain cancer. Al-Hajj *et al* found that a small population of human breast cancer cells were capable of forming tumours in NOD/SCID mice (Al-Hajj *et al* 2003), which were purified as $CD44^+CD24^{-/low}$. These cells formed tumours 6 months after injection, whereas $CD44^+CD24^+$ cells

formed no detectable tumours. Furthermore, as few as 200 CD44$^+$CD24$^{-/low}$ cells were capable of tumour formation, but larger numbers of other cells were not. Lastly, it was shown that these cells could give rise to further CD44$^+$CD24$^{-/low}$ cells that formed tumours after serial transplantation, as well as non-tumourigenic cells with diverse phenotypes that recapitulated the complexity of the tumour from which they were derived (Al-Hajj *et al* 2003). When dissociated paediatric brain tumours were cultured *in vitro* using conditions that favoured neural stem cell growth, a small fraction of cells demonstrated growth into clonally derived neurospheres termed tumour spheres, whereas other tumour cells failed to proliferate (Singh *et al* 2003). All neurosphere-like cells stained positive for CD133, a neural stem cell marker, but did not express markers of differentiated cell types. Tumour spheres were capable of *in vitro* self-renewal, with cells from more aggressive tumour types showing higher self-renewal capacity, and differentiated preferentially into the cell type predominant on the tumour of origin (Singh *et al* 2003). Self-renewal capacity was limited to the CD133$^+$ (and not the CD133$^-$) population. It was unlikely that these findings were due to contamination by normal neural stem cells as the population studied exhibited an abnormal karyotype (Singh *et al* 2003). A further study by the group of Dirks demonstrated that injection of as few as 100 CD133$^+$ cells into NOD/SCID mice initiated tumour formation, whereas injection of 100000 CD133$^-$ cells resulted in no tumour formation (Singh *et al* 2004). Serial transplantation of tumours from CD133$^+$ cells was possible and resulted in tumours phenotypically resembling the human tumour of origin, thus demonstrating *in vivo* cancer stem cell properties of this subset of brain cancer cells (Singh *et al* 2004). Thus, defined cellular hierarchies have been described in several human cancers, where a small cell population with stem cell-like properties gives rise to the heterogeneous phenotypes present in the tumour.

1.8.2 The relationship between normal stem cells and cancer

The cellular origin of cancer, or cancer stem cells, remains to be determined. However, the biology of cancer often closely resembles the biology of normal stem cells, leading to the suggestion that normal stem cells may be the source of cancer stem cells. Normal stem cells have the ability to self renew, as do cancer stem cells, but it is poorly controlled in this instance. Normal stem cells have the ability to differentiate; so do cancer stem cells, but it is aberrant and does not produce mature cells with normal function (Pardal *et al*

2003). Crucially, and perhaps understandably from these observations, many pathways involved in stem cell function are deregulated during cancer progression, further posing the possibility that normal stem cells are the source of cancer stem cells. By way of example, the Wnt signalling cascade promotes self-renewal ability (but can also control lineage-specific differentiation in some contexts, section 1.6.2.3) in several adult stem cell populations including HSC, skin stem cells and intestinal epithelial stem cells (reviewed in Reya and Clevers 2005). Significantly, alterations that cause constitutive activation of this pathway are associated with cancers of these lineages, classically intestinal tumours such as adenocarcinomas (Reya and Clevers 2005, Radtke and Clevers 2005). In addition, Rb family members are known to play roles in differentiation of mesenchymal tissues (see 1.6.2.3, for example); perhaps related to this, a 500-fold greater occurrence of osteosarcoma is exhibited by individuals with inherited heterozygous loss of the *RB* gene than the general population (Thomas and Kansara 2006). Further, as previously mentioned (1.2), it has been reported that prolonged culture of hESC lines can lead to chromosomal abnormalities that strongly resemble those observed in their cancer-derived counterparts, embryonic carcinoma cells (Draper *et al* 2004b).

Given the close relationship between normal stem cells and cancer cells, it could be hypothesised that cancer-causing mutations would take place in the normal stem cell population, leading to their transformation into cancer stem cells (Marx 2003, Passague *et al* 2003). The self-renewal ability of normal stem cells, combined with their long life span, could render stem cells more likely to accumulate mutations than progenitor cells or mature cell types; progenitor cells would have to regain this self-renewal ability (Pardal *et al* 2003, Marx 2003). In addition, it has recently been found that both embryonic and adult stem cells lack a specific cell cycle checkpoint that renders them more likely to gain chromosomal aberrations, and the efficiency of this checkpoint increases as cells differentiate (Damelin *et al* 2005). In most AML subtypes, the only cells capable of initiating AML in NOD/SCID mice (as described above) have a $CD34^+CD38^-$ phenotype, similar to normal HSCs (Bonnet and Dick 1997). Furthermore, one of the most common translocations in AML results in AML1-ETO fusion transcripts in leukaemic cells. One study found this translocation in a subset of normal HSCs that did not exhibit leukaemic properties (Miyamoto *et al* 2000). This indicated that this translocation occurred originally

in the HSC population, but further mutations that led to cancer progression occurred in an HSC subset or in their downstream progeny. Alternatively, cancer stem cells may arise from mutations in progenitor or mature cells that allow re-acquisition of self-renewal ability, and subsequent accumulation of further mutations. Myeloid leukaemias have been shown to arise when transgenes are expressed in a haematopoietic progenitor cell population (Krivtsov *et al* 2006). Thus, it appears that mutations leading to cancer may appear in both stem and progenitor cells, depending on the cancer type. It will be of great interest to investigate further the molecular transformation of stem cell populations, including the effects on self-renewal and differentiation.

1.8.3 Molecular pathways of cancer progression

Cancers are believed to arise from a series of stepwise mutations that accumulate as a result of environmental challenges or genetic instability; early mutations give rise to cells with increased propensity to accumulate further tumourigenic mutations (Hanahan and Weinberg 2000, Bjerkvig *et al* 2005). The progression of a normal cell to a cancer cell capable of forming tumours is termed neoplastic transformation. Much work has focused on the molecular pathways that underlie transformation in all cancers. Although over 100 distinct types of cancer exist that contain mutations in a wide variety and combination of genes, disruption of a defined set of regulatory pathways may be sufficient to induce transformation of many, or all, types of human cell (Hahn and Weinberg 2002). Broadly, it has been proposed that normal cells undergoing transformation acquire the capabilities to: generate their own growth signals, become refractory to growth-inhibition signals, evade apoptosis, proliferate without limit (become immortalised), induce angiogenesis in their local environment and lastly to metastasise and invade other tissues. Genome instability of cancer cells may contribute to the attainment of these attributes (Hanahan and Weinberg 2000, Hahn and Weinberg 2002).

These conclusions were drawn largely from studies aimed at elucidating the type and number of mutations required to transform a specific cell type *in vitro*, focusing mainly on human and murine cells. Properties of *in vitro*-transformed cells include density- and anchorage-independent growth, a lack of requirement for growth factors and tumour formation after transplantation into animal models. It has been shown that whilst murine

73

cells require only two mutations, or hits, to become transformed, human somatic cells require 4-6 hits depending on the cell type (Hahn *et al* 1999, Rangaranjan *et al* 2004). Stepwise introduction of distinct viral oncogenes or mutated proto-oncogenes, or inhibition of tumour suppressor gene activity, leads to *in vitro* cellular transformation (Fig. 1.7). These steps can include the introduction of hTERT, the rate-limiting subunit of the telomerase enzyme required for immortalisation (bypass of senescence and crisis), followed by introduction of the human papillomavirus E6 and E7 proteins. E6 binds to and inhibits the function of p53, which normally senses DNA damage and can inhibit the cell cycle and induce apoptosis. E7 sequesters the function of pRb, an important regulator of the G1/S phase of the cell cycle which normally responds to (anti)growth signals. Alternatively, the SV40 large T antigen (LT) can be employed which sequesters both pRb and p53. Further, introduction of the SV40 small T antigen (ST) inhibits dephosphorylation of Myc, resulting in increased Myc activity thus leading to increased cell proliferation and impaired differentiation. Finally, introduction of an oncogenic version of the signalling protein Ras results in uncontrolled pro-growth signals. Ras normally transmits pro-growth signals from growth factor signalling through three pathways, mediated by PI3K, Raf, and RAL-GEFs. It has been shown that at least two of these pathways must be perturbed for transformation of human cells, but the pathway combinations and numbers required differ between cell types. Thus, combinations of these introduced genes can transform a wide variety of normal human somatic cell types (Hahn *et al* 1999, 2002, Hahn and Weinberg 2002, Rangaranjan *et al* 2004, Zhao *et al* 2004).

As mentioned above, human somatic cells require 4-6 hits to become transformed. In contrast to this, when embryonic stem cells are implanted into immunodeficient mice, teratomas are formed, showing that these cells have tumourigenic capability without the need for transforming mutations (Thomson *et al* 1998). Transplantation of minimally cultured normal adult stem cells into recipient animals does not result in tumour formation, but their stem cell properties such as self-renewal may imply that some, but not all, tumourigenic properties are already possessed by adult stem cells (Polyak and Hahn 2006). Elucidating the number of hits required to transform adult human stem cells may be performed by determining the effects of increasing hits on their self-renewal and

74
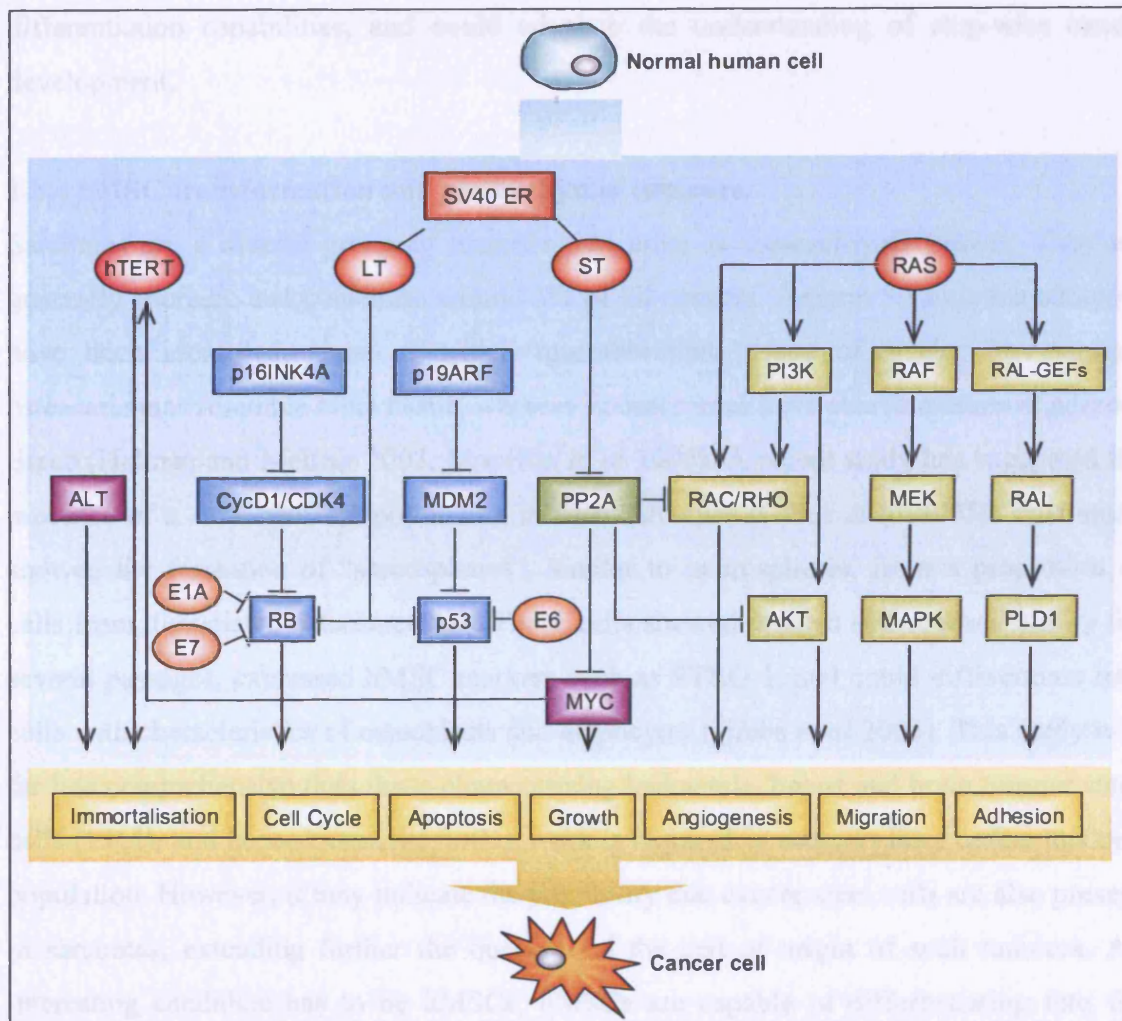
Figure 1.7. Pathways sufficient for transformation of human somatic cells. hTERT (or alternative lengthening of telomeres, ALT) lead to telomere stabilisation. pRB and p53 are inactivated by viral oncoproteins such as E1A, human papilloma virus subtype 16 E7 and E6; the SV40 LT antigen encoded by the SV40 early region (ER) can inactivate both pRB and p53. The SV40 ST antigen inactivates PP2A, leading to stabilisation of MYC and upregulated PI3K-AKT and RAC/RHO small GTPase function. Oncogenic RAS transmits constitutive signals through several pathways. The concerted action of these disruptive changes results in deregulation of processes such as the cell cycle, growth and migration. Figure adapted from Zhao *et al* 2004.

differentiation capabilities, and could advance the understanding of step-wise cancer development.

1.8.4 hMSC transformation and mesenchymal tumours

Sarcomas are a diverse group of tumours that arise in mesenchymal tissues. They are generally sporadic and constitute around 1% of all cancers. Around 50 sarcoma subtypes have been identified, some of which resemble their tissue of origin; for example osteosarcomas resemble bone tissue, whereas liposarcomas have characteristics of adipose tissue (Helman and Meltzer 2003, Mocellin *et al* 2006). A recent study has suggested the presence of a stem cell-like population in osteosarcomas (Gibbs *et al* 2005). This study showed the formation of "sarcospheres", similar to neurospheres, from a proportion of cells from dissociated osteosarcomas. These cells showed *in vitro* self-renewal ability for several passages, expressed hMSC markers such as STRO-1, and could differentiate into cells with characteristics of osteoblasts and adipocytes (Gibbs *et al* 2005). This study was far less comprehensive than those characterising leukaemia, breast and brain tumour stem cells (1.8.1), and hence extensive further work is required in order to fully define this cell population. However, it may indicate the possibility that cancer stem cells are also present in sarcomas, extending further the question of the cell of origin of such tumours. An interesting candidate has to be hMSCs; hMSCs are capable of differentiating into the tissue types from which such tumours originate, and previous evidence has suggested that cancer stem cells may arise from tumourigenic mutations in the stem cell population that lead to aberrant self-renewal and differentiation properties (1.8.2).

Various studies have investigated the potential of MSCs to become transformed. Human adipose tissue derived-MSCs cultured for 4-5 months gave rise to tumours in immunodeficient mice so were considered to have been spontaneously transformed, whereas hMSCs grown for ~2 months did not, and hence had only reached the stage of spontaneous immortalisation (Rubio *et al* 2005). This spontaneous transformation of hMSCs highlighted the fact that mutated stem cells may be able to initiate tumour formation (Rubio *et al* 2005). In a further study, murine MSCs cultured for 29 passages or more showed spontaneous transformation associated with chromosomal abnormalities including elevated c-Myc and telomerase expression (Miura *et al* 2006). These cells

formed tumours resembling fibrosarcomas in immunodeficient mice. The authors were not able, however, to transform human MSCs under the same conditions, perhaps reflecting experimental differences between the two studies (Miura *et al* 2006). Human MSCs artificially immortalised by retroviral introduction of the hTERT gene have also showed spontaneous transformation and tumour formation in immunodeficient mice (Serakinci *et al* 2004). Transformation was associated with acquired genetic changes, including deletion of the INK4A/ARF locus and an activating mutation in the K-Ras gene (Serakinci *et al* 2004). When single-cell clones from a cell line of these transformed hMSCs was analysed for their tumourigenicity, all clones were found to form tumours, but to a widely varying extent; it was therefore concluded that the cell line represented a cancer stem cell model that was hierarchical in nature (Burns *et al* 2005).

Some researchers have also applied the stepwise transformation model described in the previous section to investigate the effects on hMSCs. Two studies introduced the hTERT gene into hMSCs, and showed that the cells became immortalised, as demonstrated by a greatly extended lifespan in culture (Shi *et al* 2002, Simonsen *et al* 2002). Furthermore, it was shown that their differentiation potential was not diminished at high passage numbers (Shi *et al* 2002, Simonsen *et al* 2002). Hung and co-workers transduced hMSCs with a retrovirus expressing the E6 and E7 proteins, and found that the cells became immortalised, as assessed by unlimited passaging, but were not transformed as growth in soft agarose or tumour formation in mice were not observed (Hung *et al* 2004b). The hMSCs also retained multilineage differentiation potential. In other work, retroviral introduction of hTERT did not immortalise hMSCs; addition of E6 and E7 was required for immortalisation but this did not affect differentiation ability (Okamoto *et al* 2002). Thus, several studies have investigated the potential of MSCs to spontaneously transform, and have raised the intriguing possibility that MSCs may be the target of mutations leading to tumour formation. It has also been shown that the introduction of up to three hits does not induce full transformation of hMSCs or affect differentiation capabilities. Therefore, a comprehensive analysis of hMSC transformation using the full range of hits required to transform human somatic cells, and the corresponding effects on differentiation, has yet to be undertaken.

1.9 Thesis hypothesis and aims

The overall hypothesis of this thesis is that crucial molecular events regulate the commitment and differentiation of hMSCs towards adipocytes, but are as yet uncharacterised. The application of functional genomics during the early stages of hMSC adipogenesis will enable the identification of such events.

In order to investigate this hypothesis, this thesis aims to achieve the following:

Establish optimal hMSC differentiation to adipocytes. Human MSCs can be isolated from human bone marrow samples, and protocols for their growth and differentiation to adipocytes are well established. However, hMSC isolation, growth and adipogenic differentiation will be optimised in order to achieve the robust and reproducible differentiation required for functional genomics.

Define the time-course of molecular events during adipogenesis in order to identify the earliest stages of differentiation. Much detail is known regarding the later molecular regulation of adipogenesis. This knowledge will be related to hMSC adipogenesis in order to define the point during differentiation at which crucial early events are likely to be occurring.

Identify and validate the gene expression patterns that characterise early differentiation using microarray analysis. Microarrays covering all known human genes will be used to obtain transcriptional profiles during a time-course study of the early stages of hMSC adipogenesis. Data processing and analytical techniques will be employed to define gene expression patterns during early differentiation, and the accuracy and reproducibility of the data will be assessed.

Characterise the nature of early transcriptional events through global annotation techniques, and identify potential candidates for novel regulators of adipogenesis. The use of genome-wide arrays in this study, combined with automated annotation

methods, will allow comprehensive annotation of the molecular processes and pathways involved in early differentiation. Furthermore, manual investigation of gene function coupled with stringent selection criteria will be used to identify genes showing early regulation that may play an essential role during early stem cell differentiation.

Develop and apply techniques allowing the use of RNAi screening to functionally investigate the role of candidate genes during hMSC adipogenesis. Several methods for using retroviral RNAi libraries to screen for genes with a particular function have been described, but have not been applied to hMSCs. Therefore, a technique for using an RNAi library, plus an additional screening method, will be developed and applied to hMSC adipogenesis to investigate the role of candidate genes during the process.

Investigate the effects of specific mutations involved in cellular transformation on hMSC differentiation. It has been proposed that cancer stem cells may arise from tumourigenic mutations in the normal stem cell population, which would be expected to affect intrinsic stem cell properties such as differentiation. Human MSCs will therefore be investigated for their differentiation capabilities after the introduction of progressive numbers of transforming mutations.

CHAPTER 2. MATERIALS AND METHODS.

2.1 Materials

All reagents, histochemical stains and general chemicals were from Sigma-Aldrich (Dorset, UK), unless otherwise stated. All tissue culture media and other tissue culture reagents (trypsin/EDTA, PBS) were from Gibco (Invitrogen, Paisley, UK), unless otherwise stated. Viral vectors and antibody sources are listed in Tables 2.1 and 2.2, respectively. The source of all other materials is specified in the relevant section.

LiSa2 cells (Wabitsch *et al* 2000) were a gift from Dr. D. Guiliano (UCL, UK). The 293gp cells were from Clontech (Saint-Germain-en-Laye, France), and 293T cells were a gift from Dr. D. Trono (Geneva University, Switzerland)

2.2 Cells

2.2.1 MSC isolation and culture

Bone marrow aspirates were taken from the iliac crest of healthy transplantation donors after informed consent and with prior approval from University College London Hospital ethics of human research committee, in association with the Molecular Haematology Department, UCL. The average bone marrow aspirate volume was 10ml. Aspirates were passed through a cell strainer (pore size 40μm) to remove tissue fragments, and washed with an equal volume of PBS. Aspirates were carefully loaded onto 20ml Ficoll-Paque (1.077g/ml, Amersham Biosciences, Bucks, UK) in a 50ml Falcon tube and centrifuged (1800rpm, 30min, PK120 centrifuge, DJB Labcare, Bucks, UK) at room temperature (20-25°C). The cell layer containing mononuclear cells was isolated, washed with PBS and resuspended in Mesencult medium containing 10% human serum (both from Stem Cell Technologies, Vancouver, Canada) with or without 1ng/ml bFGF supplement (R&D Systems, Oxon, UK). Cells were plated onto 6-well plates at a density of 2x10^5 bone marrow mononuclear cells/cm^2, and incubated at 37°C and 5% CO_2 in a humidified atmosphere. Two days after plating, non-adherent cells were removed with a PBS wash and fresh media was added. Human MSC colonies formed were designated as passage 0, and were maintained in the appropriate medium, with the first passage occurring approximately 7 days after initial isolation. Cells were kept at sub-confluent levels during culture, and were passaged at a ratio of 1:4 on reaching 70% confluence, which occurred

Vector construct	Map (Figure)	Source
pMSCV-Hyg	2.1	
		Purchased from Clontech (Saint-Germain-en-Laye, France)
pMSCV-Hyg-PPARγ1	2.1	Constructed as described
pMSCV-Hyg-PPARγ2	2.1	Constructed as described
pMSCV-Puro	2.1	Purchased from Clontech
pMSCV-Puro-WPRE-eGFP	2.1	Gift from E. Crick (UCL,UK)
pRetroSuper-Puro(-shRNA)	2.1 and 5.1	NKi RNAi library
pRS-Neo-GFP(-shRNA)	2.1 and 5.5	Purchased from OligoEngine (Seattle, WA, USA) (and shRNA constructs cloned as described)
pSilR-H1-Scrambed	2.1 and 5.6	Purchased from Ambion (Cambs, UK)
pSilR-H1(-shRNA)	2.1 and 5.6	Purchased from Ambion (and shRNA constructs cloned as described)
pCSGW(-shRNA)	5.11 and 2.2	Gift from A. Godfrey (UCL, UK) (and shRNA constructs cloned as described)
pCSGW-v-cyclin	5.11 and 2.2	Gift from A. Godfrey (UCL, UK)
pVSVG	2.2	Purchased from Invitrogen (Paisley, UK)
pLP1	2.2	Purchased from Invitrogen
pLP2	2.2	Purchased from Invitrogen

Table 2.1 Plasmids used in this thesis and their source.

Antibody	Conjugation		Dilution		Incubation time (min)		Details	
	1° Antibody	2° Antibody	1°	2°	1°	2°	1°	2°
CD13	PE		1:10		30		BD Biosciences	
CD34	FITC		1:10		30		BD Biosciences	
CD44	PE		1:10		30		BD Biosciences	
CD45	PECy7		1:30		20		BD Biosciences	
CD49b	FITC		1:10		30		BD Biosciences	
CD49d	PE		1:10		30		BD Biosciences	
CD117/C-kit	-	FITC	1:10	1:400	20	30	Chemicon	DakoCytomation
CD133*	PE		1:40		20		BD Biosciences	
HLA-ABC	FITC		1:10		30		BD Biosciences	
HLA-DR	APC		1:10		30		BD Biosciences	
SH2	FITC		1:50		20		Stem Cell Technologies	
SH4	FITC		1:50		20		Stem Cell Technologies	
Stro-1	-	PE	1:5	1:8	40	30	R&D Systems	DakoCytomation

Table 2.2 Antibodies used to define the cell surface marker profile of hMSCs. The dilution at which each antibody was used and the incubation time are indicated. *co-incubate with Fc Receptor block whilst staining. All primary antibodies were monoclonal and raised in mouse; all secondary antibodies were polyclonal goat anti-mouse immunoglobulins. BD Biosciences (Oxon, UK); Chemicon (Hampshire, UK); DakoCytomation (Glostrup, Denmark); R&D Systems (Oxon, UK); Stem Cell Technologies (Vancouver, Canada).

generally every 3 days and corresponded to re-plating the cells at a density of approximately $4 \times 10^3 \text{cells/cm}^2$. For all hMSC experiments, cells were used at passage 4-6, and were from a healthy 31 year old male donor, unless otherwise indicated.

2.2.2 hMSC differentiation

2.2.2.1 Adipogenic differentiation

For adipogenic differentiation, cells were grown to confluence in media containing bFGF (unless otherwise indicated). One day post-confluence, medium was changed to Mesencult +10% serum without bFGF (unless otherwise indicated), supplemented with $1\mu M$ dexamethasone, $500\mu M$ IBMX, $200\mu M$ indomethacin and $10\mu g/ml$ human insulin (Roche, Sussex, UK). Medium was changed every 3 days (unless otherwise indicated) and the day of addition of differentiation media to cells was designated as day 0.

2.2.2.2 Osteogenic differentiation

For osteogenic differentiation, cells were grown to confluence as for adipogenic differentiation. One day post-confluence, medium was changed to Mesencult +10% serum without bFGF (unless otherwise indicated), supplemented with 100nM dexamethasone, $50\mu M$ ascorbic acid phosphate and 10mM β-glycerophosphate. Medium was changed every 3 days (unless otherwise indicated) and the day of addition of differentiation media to cells was designated as day 0.

2.2.2.3 Chondrogenic differentiation (performed by S. Vujovic)

Chondrocyte differentiation media was composed of high glucose DMEM with 1% penicillin/streptomycin (Invitrogen), $1\mu M$ dexamethasone, $50\mu g/ml$ ascorbate-2-phosphate, $40\mu g/ml$ 1-proline, $100\mu g/ml$ sodium pyruvate, 50 mg/ml insulin/ transferrin/selenium + Premix ($6.25\mu g/ml$ insulin, $6.25\mu g/ml$ transferrin, 6.25ng/ml selenious acid, 1.25mg/ml BSA, 5.35mg/ml linoleic acid, BD Biosciences, Franklin Lakes, NJ, USA), 500ng/ml BMP6 (R&D Systems), 10ng/ml TGFβ3 (R&D Systems). To initiate chondrogenesis, 2.5×10^5 hMSCs were suspended in $500\mu l$ chondrocyte differentiation media, placed in a 15 ml Falcon tube and centrifuged (600g, 30min) to form a loose cell pellet. Pellets were then incubated at 37°C and 5% CO_2 in a humidified

atmosphere. The first media change occurred 4 days later, after which the differentiation media was changed every 3 days (protocol adapted from Sekiya *et al* 2002).

2.2.3 Staining protocols for assessment of differentiation

2.2.3.1 Oil Red O: adipogenesis

Lipid accumulation was assessed by Oil Red O staining (Koopman *et al* 2001) at the timepoint indicated. Cells were fixed in 4% formaldehyde and 1.5% methanol (in sodium phosphate buffer, final concentration 75mM) for 30min, washed twice with PBS and stained for 2h with 0.04% Oil Red O dissolved in 40% triethyl phosphate. Excess dye was removed with two PBS washes and cells were visualised by phase contrast microscopy using an Axiovert 100 inverted microscope with Axiocam (Zeiss, Herts, UK).

2.2.3.2 Alkaline phosphatase activity: osteogenesis

Alkaline phosphatase activity was assayed as a marker of osteogenesis using Sigma Aldrich Kit 85L-2, following manufacturer's instructions. Fast Blue RR salts were used (Sigma-Aldrich), resulting in a blue colour where alkaline phosphatase activity was present.

2.2.3.3 Alizarin Red: osteogenesis

Calcium deposition during osteogenic differentiation was assessed with Alizarin Red staining at the timepoint indicated. Cells were fixed with ice-cold 100% methanol for 15 min, and then washed 3 times with dH$_2$O and stained for 10 min with 1.5% w/v Alizarin red (pH 4.2 with ammonium hydroxide). Excess stain was removed with further dH$_2$O washes and cells were left to air-dry before being visualised.

2.2.3.4 Toluidine Blue: chondrogenesis (in collaboration with S. Vujovic)

Wax embedding and sectioning of chondrogenic pellets Chondrocyte pellets were fixed in 4% paraformaldehyde in 2x PBS overnight at 4°C. Pellets were then placed in cellsafe biopsy capsules and subjected to the following dehydration steps: 30% ethanol, 10min; 80% ethanol, 10min; 95% ethanol, 10min; 100% ethanol, 30min; xylene, 20min. Pellets were then embedded in paraffin and left overnight to set at 4°C. Wax sections were cut

$5\mu m$ thick and floated in a water bath at 40°C before being applied to polysine coated slides, and allowed to dry on a hotplate overnight.

Staining of wax sections with toluidine blue Slides were dewaxed in xylene for 10min, rehydrated in 90% ethanol 5min, 70% ethanol 5min and washed in distilled water. The toluidine blue stain (0.1% w/v toluidine blue dissolved in 0.1M sodium phosphate buffer – 94% 0.1M NaH_2PO_4, 6% 0.1M Na_2HPO_4) was applied for 45sec and washed with distilled water. The slides were then dehydrated by consecutive ethanol washes (70%, 90%, 100%), and dipped in xylene for 5min. Slides were air dried and mounted with DPX.

2.2.4 Other cell lines

LiSa-2

LiSa2 cells were grown in DMEM/Ham's F12 (1:1) media with 10% FCS (Gibco), and were incubated at 37°C and 5% CO_2 in a humidified atmosphere. On reaching confluence, cells were passaged at a ratio of 1:3.

293gp and 293T

293gp (retroviral packaging cell line containing Moloney murine leukaemia virus gag and pol genes integrated in the genome) and 293T (human renal epithelial cell line transduced with SV40 LT antigen) cells were grown in DMEM + Glutamax media supplemented with 10% FCS (Perbio Science, Northumberland, UK), and were incubated at 37°C and 5% CO_2 in a humidified atmosphere. At confluence, cells were passaged at a ratio of 1:5-1:10.

2.2.5 Splitting cells

All cell lines were passaged using the same protocol. Cell monolayers were washed with 1x PBS then incubated for 2-5min at 37°C with trypsin/EDTA (0.25%/1% v/v). Detached cells were harvested and trypsin was inactivated via a wash with serum-containing medium. Cells were collected via centrifugation (1200rpm, 5min, PK120 centrifuge). Media was then removed and the cell pellet was resuspended at the required dilution in fresh media.

2.2.6 Counting cells

At the passage number specified, hMSCs were harvested as above (2.2.5). After collecting cells via centrifugation (1200rpm, 5min, PK120 centrifuge), the pellet was resuspended in a known volume of media, and the cell number in $10\mu l$ was quantified using a haemocytometer (counting at least 100 cells).

For growth curves, the total number of cells was then calculated and cells were re-seeded at the specified density. This process was repeated after the interval indicated; the total cell number before passaging was again quantified and cells were reseeded at the same density at every passage. Total cell numbers were used to calculate population doublings using the formula:

Population doublings = $\log_2$(total cell number obtained/total cell number seeded)

Cumulative population doublings were calculated via addition of population doublings at each passage.

2.3 FACS analysis

2.3.1 Antibody staining for surface marker expression

Subconfluent hMSCs at passage 4 were harvested (2.2.5) and washed in MACS buffer (PBS containing 2 mM EDTA and 0.5% w/v BSA, pH7.2) via centrifugation (1200rpm, 5min, PK120 centrifuge) to remove traces of media. Aliquots of 5×10^5 cells per sample were incubated with the appropriate primary antibody (all monoclonal and raised in mice) for the specific time (see Table 2.2 for antibody information, dilutions and incubation time) in MACS buffer (final volume $100\mu l$) on ice and in the dark. Cells incubated with the anti-CD133 antibody were simultaneously incubated with human Fc receptor blocking reagent (Miltenyi Biotech, Surrey, UK) to avoid cross-reaction of the antibody with this receptor. After removal of excess antibody via 3x MACS buffer washes (3 min, 2500rpm, Heraeus Biofuge Pico, DJB Labcare), secondary antibody (all polyclonal goat anti-mouse immunoglobulins) incubations were carried out as before if required, followed by further washes. Control hMSCs were stained with FITC-secondary antibody only for FITC-

conjugated antibodies, PE-secondary antibody only for PE-conjugated antibodies, and control cells were unstained for PE-Cy7 or APC-conjugated antibodies. Surface marker expression was then determined by counting 10000 events per sample on a FACScalibur flow cytometer (BD Biosciences, Oxon, UK).

Cell size and shape were analysed and gates were set to exclude debris. Control cells (either unstained, or stained with the appropriate control antibody) were analysed on the wavelength corresponding to the fluorescent signal being detected, and a gate was set to contain >99% of this population. Positive events in specific antibody-stained cells were then scored as those that were outside of this gate (i.e. above the background level).

2.3.2 GFP expression

To analyse GFP expression after transduction with GFP-containing viral vectors, cells were harvested (2.2.5) and washed once with 1x PBS to remove excess media. Cells were resuspended in PBS + 1% v/v FCS and then fixed with CellFix (10x stock, BD Biosciences). GFP expression was determined by counting 10000 events per sample. FACS gating was set up as described above, with untransduced cells as a control.

2.4 Cloning

2.4.1 Vectors used

Table 2.1 lists all vectors used in this thesis and their source, and vector maps are shown in Figures 2.1 and 2.2.

2.4.2 General cloning protocols

The following protocols were used for all cloning work; any modifications to these protocols are indicated in the specific section.

2.4.2.1 Restriction digestion

Restriction digestions of plasmid DNA or PCR fragments were set up in the volumes indicated. Final concentrations in a $20\mu l$ reaction were: 5 units each enzyme (specified, all from Promega, Southampton, UK); 1x Promega buffer (specified); DNA: $1\mu g$ DNA was
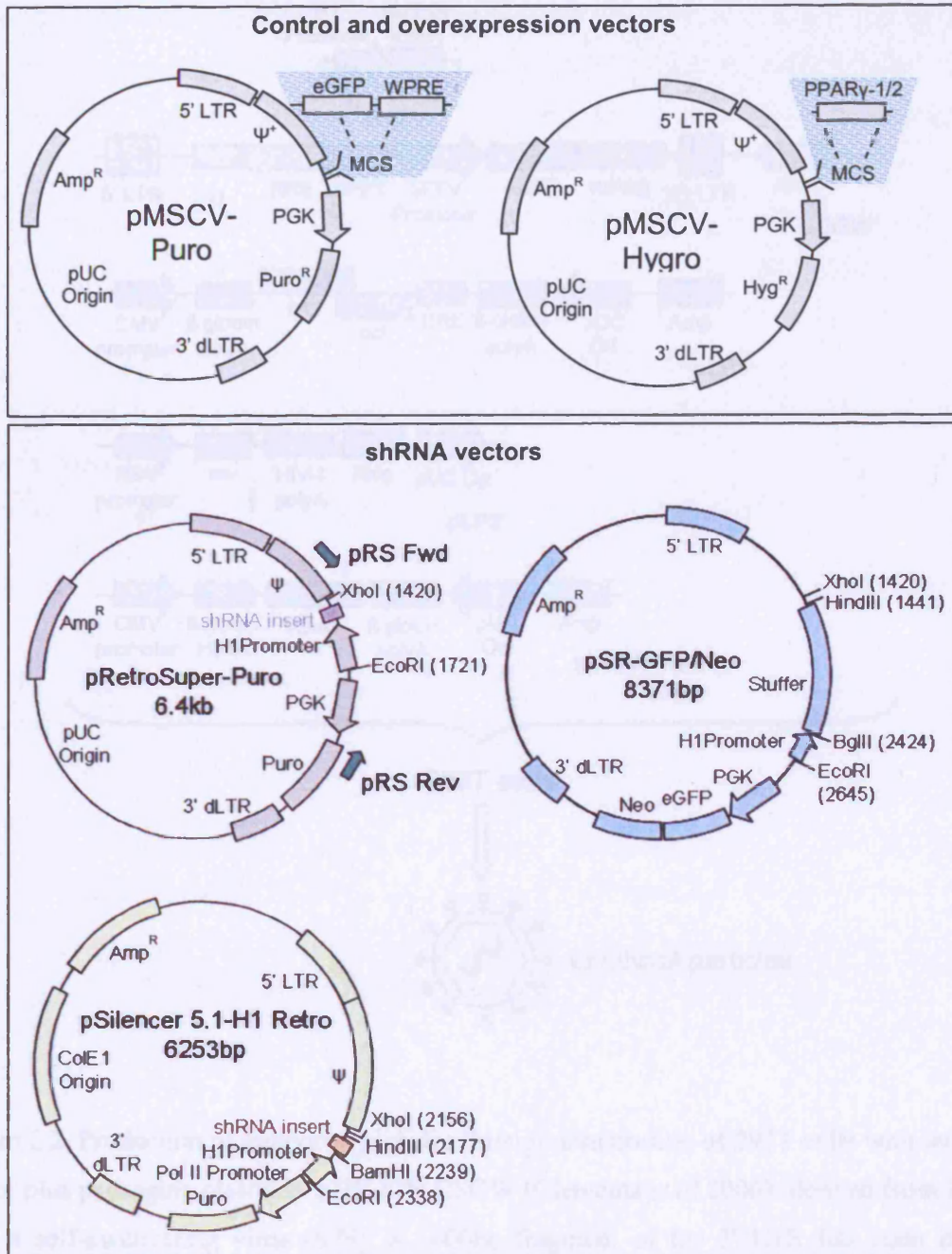
Figure 2.1. Schematic illustration of the retroviral vectors used in this thesis and their purpose. Blue shaded sections indicate cloning procedures that were used to create vectors containing the indicated genes; during experimentation, the empty vectors (name indicated, not containing blue shaded sections) were used as controls.
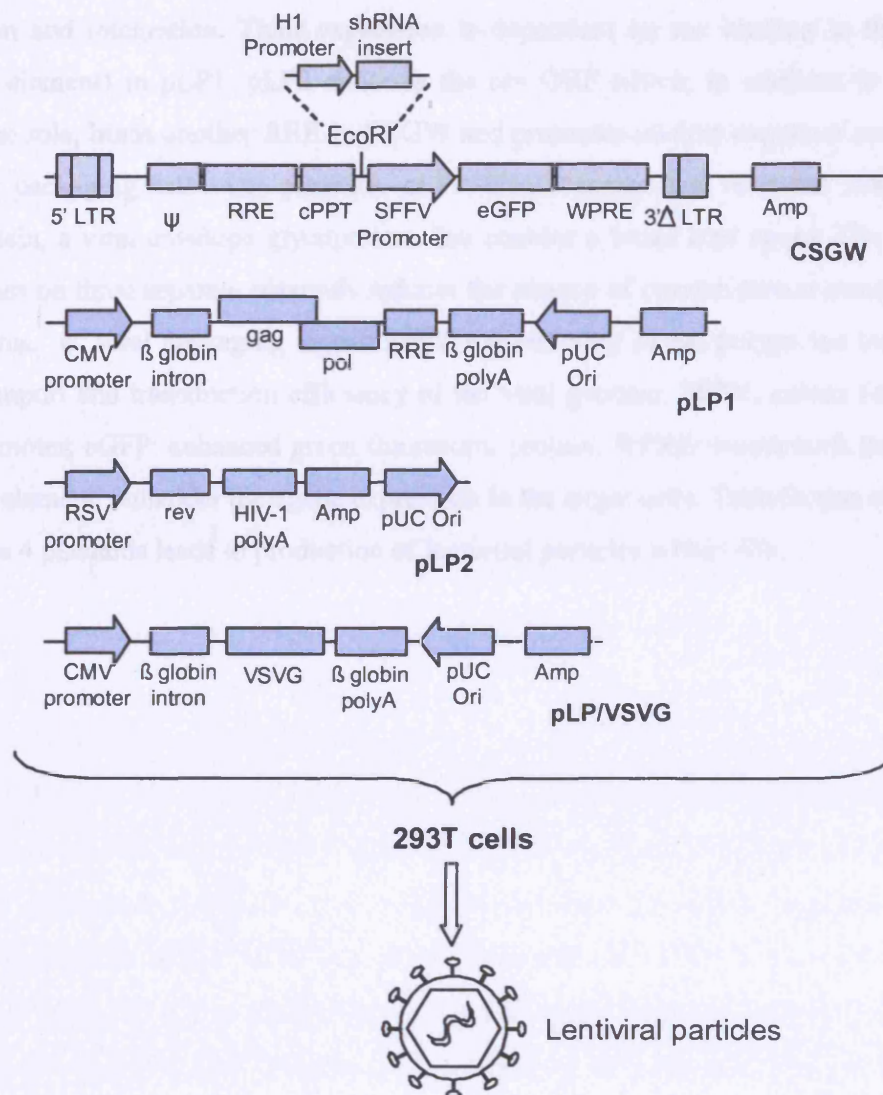
Figure 2.2. Production of lentivirus particles through transfection of 293T cells with lentiviral vector plus packaging plasmids. pHR-SIN CSGW (Clements *et al* 2006): derived from HIV-1. It is a self-inactivating virus (SIN): a ~400bp fragment of the 3' LTR has been deleted, including the TATA box, thus abolishing the promoter activity of the LTR. This is transferred to the 5' LTR during reverse transcription. Thus, after viral integration into the host genome, viruses are replication-deficient (Zufferey *et al* 1998). CSGW also lacks the viral proteins required for assembly of viral particles within the 293T cells (gag, pol, env, rev). With the Invitrogen ViraPower^{TM} Packaging Mix, these genes are provided by three packaging plasmids, pLP1, pLP2 and pLP/VSVG. pLP1: contains gag and pol ORFs; gag encodes core viral proteins for the structure of the viral particle and pol encodes enzymes required for viral

replication and integration. Their expression is dependent on rev binding to the RRE (rev response element) in pLP1. pLP2 contains the rev ORF which, in addition to it's gag/pol expression role, binds another RRE in CSGW and promotes nuclear export of unspliced viral RNA for packaging into viral particles. pLP/VSVG encodes the vesicular stomatitis virus glycoprotein, a viral envelope glycoprotein that enables a broad host range. The presence of these genes on three separate plasmids reduces the chance of recombination events producing active virus. ψ: viral packaging signal; cPPT: central copy of the polypurine tract enhances nuclear import and transduction efficiency of the viral genome; SFFV: spleen focus forming virus promoter; eGFP: enhanced green fluorescent protein; WPRE: woodchuck hepatitis virus response element: enhances transgene expression in the target cells. Transfection of 293T cells with these 4 plasmids leads to production of lentiviral particles within 48h.

used in 20μl diagnostic restriction digests, and 3μg was used in initial cloning restriction digests (50μl). Subsequent cloning steps used all products from the first restriction digestion. Reactions were carried out for 1h (unless otherwise specified) at 37°C. DNA fragments were either purified using a Gel Extraction Kit (Qiagen, Sussex, UK), according to manufacturer's protocols, and eluted in 30μl dH$_2$O, or separated on agarose gels (2.4.2.3).

2.4.2.2 Shrimp alkaline phosphatase (SAP) treatment

This was performed to remove free 5' phosphate groups from linearised vector to prevent re-ligation. Reactions were set up as follows in a 50μl volume: 1unit SAP (Promega) per 1μg DNA; 1x SAP buffer (10x stock, Promega). Reactions were incubated at 37°C for 15min, then 65°C for 15min to inactivate the reaction.

2.4.2.3 Gel separation of DNA for cloning

To isolate PCR or restriction digestion products, DNA was separated on agarose gels, density as specified, composed of x% w/v agarose in TAE buffer (tris-acetate 0.4M, EDTA 0.01M) plus 0.00002% v/v ethidium bromide for DNA visualisation, using autoclaved DNA loading buffer (6x stock concentration: 15% ficoll, 0.125% xylene-cyanol in dH$_2$O). DNA was excised from gels via visualisation on a UV block. DNA was then isolated and purified using a Gel Extraction kit, according to manufacturer's protocols, and eluted in 30μl dH$_2$O.

2.4.2.4 Ligation

Insert and vector were ligated using two different protocols:

Protocol 1: Ligation reactions (10μl final volume) were set up as follows: 3 units T4 DNA ligase (Promega), 1x ligase buffer, 50ng insert DNA, 750ng vector DNA. Ligations were carried out overnight at 16°C.

Protocol 2: Ligation reactions (10μl final volume) were set up as follows: 400 units T4 DNA ligase (New England Biolabs, Herts, UK); 1x ligase buffer, 50ng insert DNA, 700ng vector DNA. Ligations were carried out for 1hr at room temperature (20-25°C).

For both ligation protocols and in all cloning procedures, control ligations were set up by replacing insert DNA with an equal volume of dH$_2$O.

2.4.2.5 Transformation and isolation of bacterial clones

One μl of each ligation reaction was transformed into 50μl CaCl$_2$-competent TOP10 *E. coli* (prepared in the laboratory), always including the control ligation so that the level of re-ligation of the empty vector backbone could be assessed. This was accomplished by incubating DNA and bacteria on ice for 30min, followed by a 90sec incubation at 42°C. Following a further 2min incubation on ice, 950μl LB (Lennox L Broth® base, Invitrogen) was added and cells were expanded at 37°C in a shaking incubator for 30-45min. Next, 50-100μl of culture was plated using sterile glass beads onto LB-Agar (Lennox L Agar®, Invitrogen) plates containing 50μg/ml ampicillin (unless otherwise stated), which were incubated overnight at 37°C. Bacterial colonies formed overnight were picked individually into 2-4ml LB + 50μg/ml ampicillin. These were cultured overnight at 37°C in a shaking incubator. The following morning, a Spin Miniprep Kit (Qiagen) was used to isolate plasmid DNA.

2.4.2.6 Plasmid DNA isolation

Plasmid DNA was isolated using Spin Miniprep Kit, Midiprep Kit or Maxiprep Kits from Qiagen, according to manufacturer's protocols. Plasmid preparations were quantified using a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies, Montchanin, DE, USA).

2.4.2.7 DNA sequencing (performed by DNA sequencing service, WIBR)

All DNA sequencing was performed using a Beckman Coulter CEQ 8000 genetic analysis system (Beckman Coulter, Bucks, UK). Normal sequencing was performed using a Quickstart Kit DTCS, CEQ™ (Beckman Coulter), and shRNA hairpins were sequenced using a Methods Development Kit, GenomeLab™ (Beckman Coulter).

2.4.3 PPARγ isoforms: specific protocols

RT-PCR primers were designed that would amplify the human PPARγ1 and -2 isoforms separately (Table 2.3). Restriction sites for *Xho*I and *Sma*I were included on the 5' and 3' primers, respectively (plus additional bases to allow optimal efficiency of the restriction enzymes), in order to clone the PCR fragments into pMSCV-Hygro (Table 2.1). Total RNA was extracted (2.6.1) from 1.5 x10^6 hMSCs at day 4 of adipogenesis (2.2.2.1), and

Gene	GenBank ID	Fwd	Rev	TaqMan Probe	Product size (bp)
Cloning Primers					
PPARγ1	NM_138712	CCGCTCGAGCGGATGGTTGACAC AGAGATGCC	TCCCCCGGGGGATGCTAGTACA AGTCCTTGTAG		~1.4kb
PPARγ2	NM_015869	CCGCTCGAGCGGATGGGTGAAA CTCTGGAGA	TCCCCCGGGGGATGCTAGTACA AGTCCTTGTAG		~1.5kb
pRS	-	CCCTTGAACCTCCTCGTTCGACC	GAGACGTGCTACTTCCATTTGTC		N/A
Sequencing primers					
pRS Seq	-	GCTGACGTCATCAACCCGCT	N/A		N/A
pSilR-H1 Seq	-	TTGTACACCCTAAGCCTC	N/A		N/A
QPCR Primers					
GAPDH	Primers designed against a region common to all GAPDH sequences	GGAGTCAACGGATTTGGTCGTA	GGCAACAATATCCACTTTACCA GAGT	CGCCTGGTCACCAGGGCTGC	78
CSGW packaging signal (ψ)	-	GCACGGCAAGAGGCGA	CGCACCCATCTCTCTCCTTCTA	CGGCGACTGGTGAGTACGCCAA AAAT	84

Table 2.3. Primers used for gene expression studies, cloning techniques or sequencing (continued overleaf).

Gene	GenBank ID	Fwd	Rev	Product size (bp)
RT-PCR primers				
ACTA2	NM_001613	ACTGGGACGACATGGAAAAG	GAAGGAATAGCCACGCTCAG	368
Adipsin	NM_001928	GGTCACCCAAGCAACAAAGT	CCACCTCAGCCTCCTGAGTA	248
ALP	NM_000478	GGACATGCAGTACGAGCTGA	AGACTGCGCCTGGTAGTTGT	457
aP2	NM_001442	TGCAGCTTCCTTCTCACCTT	CATCCCCATTCACACTGATG	192
β-actin	NM_001101	CCAAGGCCAACCGCGAGAAGATGAC	AGGGTACATGGTGGTGCCGCCAGAC	587
CBFA1	NM_004348	CGCATTCCTCATCCCAGTAT	TATGGAGTGCTGCTGGTCTG	581
C/EBPα	NM_004364	GAGGAGGGGAGAATTCTTGG	CCCTATGTTTCCACCCCTTT	302
CLP1	NM_001884	TGCTCAGATTGCAAAAGTGG	TATCTGGGAAACCCACGAAG	170
COL1A1	NM_000088	TCGTGGAAATGATGGTGCTA	ACCAGGTTCACCGCTGTTAC	349
DACT1	NM_016651	AAGAAGTGTCGCTTCCCAGA	GGGCCTCTTCGTAGGAAATC	244
DLX2	NM_004405	AGCAGCTATGACCTGGGCTA	GAAGCACAAGGTGGAGAAGC	380
DNCH1	NM_001376	GGAAAGTGGCTGGTGTTGTT	GGAAGTGCTTCAGAGCAACC	610
FAT	NM_005245	GCACAGTCCTGCTTTCCTTC	GTGATTACGCCGGACAGTTT	504
GAPDH	NM_002046	GATCATCAGCAATGCCTCCT	TGTGGTCATGAGTCCTTCCA	97
HES1	NM_005524	CTACCCCAGCCAGTGTCAAC	CATTGATCTGGGTCATGCAG	408
HOXA7	NM_006896	TCAACAGCCCCCTTTATCAG	CAGTCGGACCTTCGTCCTTA	417
HOXB2	NM_002145	AGCTGCTGGAACTGGAGAAG	CTCCGGATAGCTGGAGACAG	501
INSIG1	NM_198336	ACCGATTCTGAGAGCAAGGA	ACCTACCTCCTTTGGGCACT	375
LHFPL2	NM_005779	CCTTGCCCTGAGTGAATCAT	CCTTGTGACTGTTCCCAGGT	620
LIFR	NM_002310	GAACAGGCCGTGGTACTGAT	TTCCAAGGGCATATCTGAGG	422
MEOX2	NM_005924	AGAAGTGCACCGCTATTGCT	TTCCTGGGAGTCTGAGCTGT	657
NEDD9	NM_006403	TGTATGACATCCCGCCTACA	GACATGTTACTCCGGGTGCT	506

95

Gene	GenBank ID	Fwd	Rev	Product size (bp)
RT-PCR primers continued				
Osteocalcin	NM_199173	GGCAGCGAGGTAGTGAAGAG	CTGGAGAGGAGCAGAACTGG	230
PLZF	NM_006006	CAAGGCTGACGCTGTATTGA	TTTGTGGCTCTTGAGTGTGC	676
PPARγ	NM_138712 NM_015869	TTCAGAAATGCCTTGCAGTG	CACCTCTTTGCTCTGCTCCT	332
SIX2	NM_016932	CTGGAGAGCCACCAGTTCTC	CGGGTTGTGGCTGTTAGAAT	393
SNAI1	NM_005985	ACCCCACATCCTTCTCACTG	CCGACAAGTGACAGCCATTA	552
SNAI2	NM_003068	GAGCATACAGCCCCATCACT	CTCCCCCGTGTGAGTTCTAA	479
SOCS3	NM_003955	CCACCTGAGTCTCCAGCTTC	GAGGAGCATGTCACCAGGAT	332
TBX2	NM_005994	TATCAGATCCCGGTCACCAT	CCAGTTTTATCACGCGGTCT	684
TCF8	NM_030751	GCACCTGAAGAGGACCAGAG	TGGTGATGCTGAAAGAGACG	537
TGFBR3	NM_003243	CAATGGCTCCTGTGGCTAAT	CCATCTGGCCAACCACTACT	589
TGIF	NM_003244	GAAAGGATGGCAAAGATCCA	TTTAGCAATCCCATGTGCAA	639

was quantified using a Nanodrop ND-1000 spectrophotometer. cDNA was synthesised from 1μg total RNA (2.6.2) and used as a template (0.25μg) in 100μl PCR reactions with isoform-specific primers (Table 2.3) to amplify the cDNAs of interest (2.6.3.1). PCR fragments were separated on 1% agarose gels then isolated and purified (2.4.2.3). PCR fragments were digested with *Sma*I in 50μl reaction volumes using Multicore buffer (Promega) to produce blunt ends (2.4.2.1). Simultaneously, 5μg pMSCV-Hygro (Fig. 2.1 and Table 2.1) was digested with *Hpa*I in Multicore buffer to produce a blunt end in the multiple cloning site. DNA was isolated and purified from both reactions as described (2.4.2.3). The products of these reactions were next digested with *Xho*I in Promega buffer D in 50μl reactions; the pMSCV-Hygro restriction digest was carried out for 3hrs due to the close proximity of the *Xho*I site to the already cut *Hpa*I site rendering the reaction less efficient. The digested fragments and vector were separated on 1% agarose gels then DNA was isolated and purified as before. The PCR fragments and linearised vector were ligated overnight (2.4.2.4, protocol 1). Ligation products were transformed into bacteria and plated onto LB-agar containing 50μg/ml ampicillin (2.4.2.5). Bacterial colonies formed overnight were picked (10 for each isoform), suspended in 25μl dH₂O, and 2.5μl of each was used as a template in a PCR screen (20μl reaction volume, using the PPARγ isoform-specific primers, 2.6.3.1) to identify positive clones. Positive clones were grown in LB + ampicillin (50μg/ml) overnight, and then DNA was isolated using a Spin Miniprep Kit. Plasmids were sequenced (2.4.2.7) using the cloning primers to ensure that the coding sequences did not contain nonsense or missense mutations.

2.4.4 Retroviral shRNA constructs: specific protocols

2.4.4.1 PCR-based shRNA insert isolation from RNAi library

Bacterial cultures (in LB + 100μg/ml ampicillin, containing a mix of the 3 shRNA constructs for each gene) were inoculated directly from the RNAi library glycerol stocks and grown overnight. If bacteria did not grow, glycerol stocks were picked and bacteria was suspended directly in 25μl dH₂O. To amplify the shRNA inserts, 2μl bacterial culture (or suspension) was used in 50μl PCR reactions (2.6.3.1) with the pRS primers (Table 2.3). PCR products were separated on 1.5% agarose gels and DNA was isolated and purified as described. Both PCR products and pRetroSuper-Puro vector were digested simultaneously with *Xho*I and *Eco*RI in Promega buffer H in 50μl reactions, and then

DNA was separated on 2% agarose gels. Linearised vector and digested PCR products were isolated and purified from the gel, and ligated overnight (2.4.2.4, protocol 1). Ligations were transformed into bacteria and plated onto LB-agar containing 100μg/ml ampicillin. Bacterial colonies formed overnight were picked (10 for each shRNA), suspended in 25μl dH$_2$O, and 1μl of this was used as a template in a PCR screen (20μl reaction volume, using the pRS primers) to identify positive clones (2.6.3.1). Positive clones were grown in LB + 100μg/ml ampicillin overnight, and then plasmids were isolated using a Spin Miniprep Kit. Plasmids were sequenced (2.4.2.7) using the pRS Seq primer to ensure that the shRNA inserts did not contain mutations.

2.4.4.2 Sub-cloning of shRNA inserts into different retroviral vectors

When it was found that the pRetroSuper-Puro plasmid used above had recombined, shRNA inserts were sub-cloned into pRS-Neo/GFP and then pSilR-H1. pSuperRetro-Puro shRNA constructs were digested simultaneously with *Xho*I and *Eco*RI in Promega buffer H in 50μl reactions. Additionally, pRS-Neo/GFP or pSilR-H1 was digested in the same manner. Restriction digest products were separated on a 1.5% agarose gel, and the shRNA insert and linearised vector were isolated, purified and ligated (2.4.2.4, protocol 2). Plasmid DNA was isolated from a selection of bacterial colonies (5 for each shRNA) as described (2.4.2.5). A diagnostic *Xho*I/*Eco*RI restriction digestion in 20μl reactions was used to confirm the presence of shRNA inserts. The shRNA insert was sequenced (2.4.2.7) using the pSilR-H1 Seq primer (Table 2.3) to ensure that no further mutations had occurred.

2.4.5 Lentiviral shRNA constructs: specific protocols

2.4.5.1 shRNA oligonucleotide design

For two genes, 19nt shRNA sequences were obtained from the literature. For two further genes, 19nt sequences were designed using OligoEngine design software (available at www.oligoengine.com). Oligonucleotides were then ordered which contained the sense and antisense 19nt sequence separated by a hairpin sequences (Table 2.4 illustrates shRNA and hairpin sequences and an example of the full oligonucleotide sequence). Additionally, *Bgl*II and *Hin*dIII overhangs at the 5' and 3' ends, respectively, plus an additional *Eco*RI site downstream of the shRNA were included.

Gene (insert name)	shRNA sequence	Hairpin
Nki RNAi library		
HOXA7 X	CTCACAGAGAAGCGGCTAC	TTCAAGAGA
HOXA7 Y	GCGTCTTTAAGAGACTCAC	unless
TCF8 X	GCTGTTGTTCTGCCAACAG	otherwise
TCF8 Y	CAAGGTGGCCATTCTGTTA	stated
TCF8 Z	GGTCAACTATCACTAGTGT	
HOXB2 X	GGAGTCGACATTAATTCCT	
HOXB2 Y	TTTGAGAGGGAGATTGGGT	
TBX2 X	CTGCTGAGCATCCAGAGAC	
MEOX2 X	CCAGAGCCAGTTTCTTGTC	
TGIF X	GGGGCAACCTACCCAAGGA	
TGIF Y	CATGAGGAATCTGGCTGCT	
FAT X	CTACGACAGTCACTTCGAT	
FAT Y	CTCTGCAGCTAAGACTTAT	
LHFPL2 X	TTCTGGCAGGCCACAGCTA	
LHFPL2 Y	GCAGAAATTGCAACCTCTA	
LHFPL2 Z	GGAATTGCAGGTCTATTCC	
SIX2 X	CGAGAACTCCAATTCTAAC	
SIX2 Y	CATGTCAGCCAACCTCGTG	
SNAI2 X	GATGCATATTCGGACCCAC	
PPARγ X	GAGCCTTCCAACTCCCTCA	
Control hairpins		
Scrambled	ACTACCGTTGTTATAGGTG	
KSHV viral cyclin	GTTCCTGCCAACGTCATTCCGCAG	TTCG
Published sequences		
DACT1	GAATTCGAAACAGGCGTGT	
SNAI1	GCGAGCTGCAGGACTCTAA	GAAGCTTG
Custom designed (using OligoEngine software)		
SIX2 a	CAAGGCACACTACATCGA	
SIX2 b	AACGAGAACTCCAATTCTA	
TCF8 a	AAGCAGGATGTACAGTAA	
TCF8 b	CTAGTCAGCCACCTTTAAA	

Bglll   19nt   hairpin   19nt   EcoRI   HindIII

GATCCCC   TTTTTGAATTC
GGG   AAAAACTTAAGTCGA
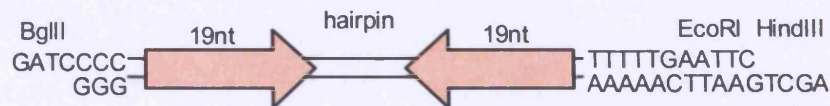
Table 2.4. shRNAs sequences and hairpins used in this thesis, and design of oligonucleotide used for cloning shRNAs into lentiviral vectors.

2.4.5.2 Annealing oligonucleotides

To anneal each complementary pair of oligonucleotides, each individual oligo was diluted to 3mg/ml using nuclease-free dH₂O, then 1μl of each oligo was added to 48μl annealing buffer (100 mM NaCl, 50 mM HEPES, pH 7.4). The reaction was incubated at 94°C for 4min, then cooled using a thermal cycler in 2-3 degrees steps until 75°C, incubating at each temperature for 4min. The reaction was then cooled in 5 degree steps until 50°C, incubating at each temperature for 4min. The reaction was then cooled in 10 degree steps until 20°C, incubating at each temperature for 10min. Annealed oligos were stored at -20°C for a short time before ligation.

2.4.5.3 Ligation into retroviral shuttle vector

The retroviral shuttle vector used was pSilR-H1, which contains the H1 promoter plus *Bam*HI and *Hind*III restriction sites in the desired locations (Fig. 2.1, this vector is purchased linearised with *Hind*III and *Bam*HI, which produces sticky ends compatible with *Bgl*II ends). Oligonucleotides and linearised vector were ligated (2.4.2.4, protocol 2), and then transformed into bacteria (2.4.2.5). Plasmid DNA was isolated from a selection of bacterial colonies (for each shRNA, 2.4.2.5) and was then digested with *Eco*RI (20μl reaction, buffer H) to confirm the presence of shRNA inserts.

As the DACT1 shRNA sequence contained an *Eco*RI site, this oligonucleotide was directly ligated (as described for other shRNA oligonucleotides) into pSilR-H1. Positive clones were identified as described above, and this vector was then used for experimentation. This construct was sequenced (2.4.2.7) with the pSilR-H1 Seq primer (Table 2.3).

2.4.5.4 Transferring the H1 promoter and shRNA into CSGW

pSilR-H1-shRNA constructs were digested with *Eco*RI to excise the H1 promoter and shRNA insert, which was then separated on a 1.5% agarose gel and purified as before. Next, 3μg CSGW was digested with 3 units *Eco*RI (Multicore buffer, 10μl reaction volume), for 30min at 37°C (due to two *Eco*RI star sites elsewhere in the vector). The DNA was purified as before and SAP-treated (2.4.2.2). Next, the restriction digest mixture was separated overnight at low voltage on a 1% agarose gel to ensure removal of

100

undigested vector. The band corresponding to linearised vector (as compared to undigested vector control run simultaneously on the gel) was excised and DNA was purified as before. Ligations were performed as described (2.4.2.4, protocol 2), and transformed into bacteria. The *Eco*RI diagnostic restriction digest described above was used to identify shRNA-positive vectors. Constructs were sequenced using the pSilR-H1 Seq primer (2.4.2.7).

2.5 Cell manipulation: transfection and infection

2.5.1 Transfection: all retroviral vectors

A common protocol was used to transfect 293gp cells with all retroviral vectors. Cells were at 75% confluence at the time of transfection. Equal ratios of vector and pVSVG helper plasmid were combined in serum-free DMEM. The amount of DNA used varied according to the scale of the transfection, and was determined according to the manufacturer's protocols for the Lipofectamine reagent (Invitrogen), but typically for one well of a 6-well plate, 2μg of each plasmid were combined in 250μl media. Simultaneously, the corresponding amount of Lipofectamine was diluted in serum-free media; 6μl was used for each well of a 6-well plate and added to 250μl media (this was scaled up if necessary). Diluted DNA and Lipofectamine were then mixed and incubated at room temperature (20-25°C) for 20min. Following this, DNA-Lipofectamine complexes were added drop-wise to cells. After 16h, media was changed to Mesencult +10% human serum supplement if hMSCs were to be infected, or to fresh DMEM +10% FCS if LiSa2 cells were to be infected. Sixty hours after the media change, target cells were infected with the fresh viral supernatant produced.

2.5.2 Transfection: all lentiviral vectors

Figure 2.2 illustrates how viral particles are produced from CGSW and the helper plasmids. For this transfection, 293T cells (at around 75% confluence) were used as they do not have stably integrated gag or pol genes (like 293gps) which would compete with the helper plasmids. DNA was mixed in serum-free DMEM at a ratio of CSGW:pLP1:pLP2:pVSVG 3:2:2:2. For a 10cm dish, a total of 4.5μg DNA was used in 50μl media. Simultaneously, Fugene transfection reagent (Roche) was diluted in serum-

free DMEM; for a 10cm dish, 13.5μl Fugene (3x amount of DNA w/v) was diluted in 200μl media. The two solutions were then mixed and incubated at room temperature (20-25°C) for 15min. Following this, DNA-Fugene complexes were added drop-wise to cells. After 16h, media was changed to Mesencult +10% human serum supplement for hMSC infection. Sixty hours after the media change, viral supernatant was collected, filtered through a 0.45μm filter, and frozen in 1ml aliquots at -80°C.

2.5.3 Infection: all retroviral vectors

Fresh retroviral supernatant was used to infect hMSCs or LiSa2 cells, which were at 30-50% confluence at the time of infection. Viral supernatant was collected, filtered through a 0.45μm filter, diluted 1:2 in fresh media (corresponding to the target cell's normal growth media), and supplemented with polybrene (final concentration 4μg/ml). Target cell media was removed and diluted viral supernatant was added. Fresh media (corresponding to the target cell's normal growth media) was added to the transfected 293gps. This was performed in the evening; the following morning, media on target cells was changed, and in the evening the infection process was repeated, as this was found to increase infection rates. The next morning, media was again changed, and if cells were reaching confluence they were split 1:2 to prepare them for drug selection.

2.5.4 Infection: all lentiviral vectors

Frozen viral supernatant was thawed slowly at 4°C, according to requirements. Simultaneously, target hMSCs were harvested and counted (2.2.6), and divided into aliquots containing 5x10^4 cells. For neat infections, these cells were collected via centrifugation (1200rpm, 5min, PK120 centrifuge), resuspended in undiluted viral supernatant, and seeded at 2.5x10^4 cells/cm^2. For all other infections, hMSC growth media was added to achieve the dilution indicated. Cells were then resuspended in this and seeded at 2.5x10^4 cells/cm^2. After 16hr, cells had adhered and media was changed. Cells were either harvested ~48h after infection for GFP expression studies, or allowed to reach confluence before experimentation.

2.5.5 Drug selection (retroviral vectors only)

Drug selection was commenced 2 days after the final infection (or when cells were around 70-80% confluent). Each drug was added at the final concentrations indicated to cells

containing vectors expressing the corresponding resistance cassette. Drug selection was continued for the duration indicated, or until 100% cell death was observed in uninfected cells incubated with the specific drug. When selection was complete, media was changed and if cells were sparse, they were passaged and reseeded at higher densities. Hygromycin (Calbiochem, Nottingham, UK): $50\mu g/ml$ for hMSCs, $200\mu g/ml$ for LiSa2s, added for 4-6 days. Puromycin: $750ng/ml$ for hMSCs, added for 3-4 days. Neomycin: $500\mu g/ml$ for hMSCs, added for 4-6 days.

2.5.6 Titering lentivirus (in collaboration with R. Vart)

Genomic DNA was extracted from 1×10^5 hMSCs 48h after infection with a 1:2 dilution of viral supernatant for each lentiviral shRNA construct (plus empty CSGW lentivirus). This was performed in duplicate for each construct, using the QIAamp DNA Mini Kit (Qiagen) and following the protocol for cultured cells. DNA was eluted in $70\mu l$ AE buffer (10 mM Tris-Cl; 0.5 mM EDTA; pH 9.0, Qiagen) and quantified using a Nanodrop ND-1000 spectrophotometer. TaqMan QPCR was used to determine the average number of viral copies per cell for each construct, using an ABI Prism 7700 Sequence Detection System® (Applied Biosystems, Warrington, UK). Gene-specific primers and TaqMan probes (conjugated with the 6-FAM™ fluorescent dye and TAMRA™ fluorescent quencher) were designed using Primer Express Software (Applied Biosystems) against the CSGW packaging signal (ψ), and an endogenous reference gene, GAPDH (Table 2.3). Q-PCR reactions were carried out in $50\mu l$ volumes, using PCR mastermix from Applied Biosystems, and $10\mu l$ template from each sample. PCR cycle conditions used were: 50°C for 2min, 95°C for 15min, then 40 cycles of 95°C for 15sec then 60°C for 1min. In 96-well plate format, standard curves for ψ and GAPDH were constructed via determination of the threshold cycle number obtained from duplicate serial dilutions of DNA in which the copy number of GAPDH and ψ is known. The threshold cycle number, Ct, is the number of PCR cycles required to reach a threshold fluorescence value (that is constant across the entire experiment), where fluorescence is emitted from TaqMan probes bound to their target gene undergoing amplification, and hence is proportional to target abundance. On the same reaction plate, the Ct for GAPDH and ψ in 1:1 and 1:10 dilutions of each duplicate unknown sample (genomic DNA from hMSCs transduced with each lentiviral construct) was determined. The standard curves were used to determine the number of

103

cells (from GAPDH standard curve) or viral copies (from ψ standard curve) that corresponded to the average Ct value from each unknown sample. Average viral copy number per cell was then calculated by dividing the number of cells in each sample by the number of viral copies.

2.6 Gene expression analysis

2.6.1 RNA extraction (including microarrays)

Total RNA was extracted using RNeasy Mini Kit (Qiagen, day 0 – 1 of differentiation), following the protocol for cultured cells, or RNeasy Lipid Mini Kit (Qiagen, day 2 – 7 of differentiation) following manufacturers protocols, and eluted in nuclease-free dH_2O. RNA was quantified using a Nanodrop ND-1000 spectrophotometer and was stored at -80°C.

2.6.2 cDNA synthesis (excluding microarrays)

cDNA was synthesised from $0.1\mu g$ total RNA (unless otherwise stated) using Superscript II reverse transcriptase (Invitrogen), following manufacturer's protocols. RNA was diluted to a volume of $10\mu l$ in nuclease-free dH_2O, and $1\mu l$ each of dNTP mix and oligo(dT)$_{15-18}$ primer were added (stock concentrations 10mM and $500\mu g/ml$, respectively). The mixture was incubated at 65°C for 5min, then $4\mu l$ first strand buffer (Invitrogen, 5x stock), $2\mu l$ DTT (stock concentration 0.1M), and $1\mu l$ dH_2O were added. This was incubated at 42°C for 2min, and then 200 units of Superscript II were added. First strand cDNA synthesis was carried out at 42°C for 50min, and the reaction was inactivated at 70°C for 15min.

2.6.3 PCR

Two PCR protocols were used during this study; the protocol was changed for gene expression studies in Chapter 5 in an attempt to increase sensitivity and maximise detection of small changes in gene expression. The primers used in each of the PCR studies are listed in Table 2.3. Both PCR protocols were executed using either DNA Engine DYAD® (BioRad, Herts, UK) or Primus 96^{PLUS} (MWG-Biotech, Ebersberg, Germany) thermal cyclers.

104

2.6.3.1 Protocol 1: used for PPARγ cloning and gene expression studies, and for shRNA cloning in Chapter 5

Gene expression studies and PCR screens were carried out in 20μl reactions, and PCR amplification of cloning fragments was carried out in 50μl or 100μl reactions. A typical 20μl PCR reaction was set up with the final concentration of reagents as follows: 250μM dNTP mix (Promega), 1x DyNAzyme™ EXT Buffer containing 1.5mM $MgCl_2$ (unless otherwise stated, Finnzymes, Herts, UK), 1μM each primer, 5% DMSO, 0.2 units DyNAzyme EXT DNA polymerase (Finnzymes), 50ng template. Larger reactions were scaled up following these proportions. PCR conditions used were:

PPARγ cloning and PCR screens: (3mM $MgCl_2$) 94°C for 5min, 30 cycles of (94°C for 30sec, 55°C for 1min, 72°C for 1.5min), then 72°C for 10min.

shRNA cloning and PCR screens: (1.5mM $MgCl_2$) 94°C for 10min, 30 cycles of (94°C for 20sec, 62°C for 30sec, 68°C for 1.5min), then 68°C for 10min.

Gene expression studies in Chapters 3, 4 and 6: (1.5mM $MgCl_2$) 94°C for 2min, X cycles (indicated in experimental section) of (94°C for 30sec, 55°C for 1min, 72°C for 45sec), then 72°C for 10min.

2.6.3.2 Protocol 2: used for gene expression studies in Chapter 5

Gene expression studies in Chapter 5 were carried out in 20μl reactions, with reagent final concentrations as follows: 200μM dNTP mix, 1x green GoTaq® Flexi Buffer (Promega), 1.5mM $MgCl_2$, 1μM each primer, 5% DMSO, 0.2 units GoTaq® Flexi DNA polymerase (Promega), 50ng template. PCR conditions used were: (1.5mM $MgCl_2$) 94°C for 2min, X cycles (indicated in experimental section) of (94°C for 30sec, 55°C for 1min, 72°C for 45sec), then 72°C for 10min.

2.6.3.3 Gel separation for gene expression

For both PCR protocols, reactions were separated on agarose gels (density as specified, composed of x% w/v agarose (typically 1.5% for gene expression studies) in TAE buffer plus 0.00002% v/v ethidium bromide for DNA visualisation), with protocol 1 using DNA loading buffer (6x stock concentration: 15% ficoll, 0.125% xylene-cyanol), and then visualised on a VersaDoc 3000 Imaging System (BioRad, Herts, UK).

2.7 Microarray gene expression studies

2.7.1 Synthesis of cRNA probes

Total RNA was extracted and quantified as described (2.6.1) from differentiating hMSCs at the timepoints indicated. First and second strand cDNA synthesis and preparation of biotin-labelled cRNA probes were performed using the One Cycle Target Labelling Kit (Affymetrix, Bucks, UK), following manufacturers instructions (Affymetrix (2004), section II, chapter I). The following details or modifications were employed. For the first strand synthesis reaction: the standard protocol was employed, and $10\mu g$ total RNA and 400units Superscript II reverse transcriptase were used. PolyA RNA controls were not added. The 2nd strand cDNA synthesis reaction was performed immediately after the 1st strand reaction, and EDTA was not added after the reaction was complete. cDNA was next cleaned up and eluted in $14\mu l$ cDNA elution buffer. The *in vitro* transcription reaction to produce biotin-labelled cDNA was performed immediately after the 2nd strand reaction, using the entire volume of eluted cDNA. The reaction was performed overnight at 37°C, then cleaned up and eluted as per manufacturer's protocols. Probe quantity was then determined using a Nanodrop ND-1000 spectrophotometer, and probe quality was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Cheshire, UK). Probes were stored at -80°C prior to fragmentation, when $20\mu g$ of cRNA probe was fragmented as per manufacturer's protocols.

2.7.2 Array hybridisation, washing, staining and scanning (performed by Affymetrix microarray service, WIBR)

Hybridisation of fragmented cDNA probes to U133 Plus 2.0 Affymetrix GeneChips was performed according to manufacturer's protocols for the 49 array format (standard) (Affymetrix (2004) section II, chapter II). Array staining with streptavidin-phycoerythrin conjugate, washes and array scanning were performed according to manufacturer's protocols for the 49 format using the Fluidics Station 450, GeneChip Scanner 3000 with autoloader, and GeneChip Operating Software, all from Affymetrix (Affymetrix (2004) section II, chapter III).

2.7.3 Data processing and analysis (in collaboration with Dr M Trotter for Chapter 3, and Dr S Henderson for Chapter 4)

Array scanning produces .cel files (Appendix Folder 1) which contain fluorescent signal intensity readings for each probe on the array. To obtain expression values for each gene, the robust multi-array analysis method of Irizzary *et al* (2003a, 2003b) was used to process these files. The R statistical environment and software packages therein were employed to accomplish this. Specifically, the "affy" package and the "rma" algorithm from Bioconductor were used (Bolstad *et al* 2003, Gentleman *et al* 2004). RMA performs three steps: probe-level background correction, quantile normalisation and summarisation of all the signal values in a probeset to give a single expression value. Firstly, background correction algorithms remove "obscuring variation" from the data caused by factors such as uneven washing of the microarrays causing patches of uneven signal intensity within the array, with the aim that "interesting variation", namely changes in gene expression across arrays, is more accurately represented. Only the PM probes (1.7.1.1) are included in RMA processing; MM probes are considered ineffectual in this method (Irizarry *et al* 2003a). Probe-level $\log_2$-transformed expression signals are obtained, allowing up- and downregulation to be evenly distributed over a numerical scale and to reduce skew caused by outliers. Normalisation attempts to equalise expression profiles of the arrays present, and is thus essential if values across arrays are to be compared. The quantile normalisation method is used, which normalises data quantile by quantile, starting with the highest expression intensities (Bolstad *et al* 2003). Finally, an additive linear model is fit to the normalized data to obtain an expression measure for each probe on each array. Expression measures are reported in RMA expression units ($\log_2$), which range from 4-16.

To calculate statistically significant gene expression changes, expression values for each gene were compared between consecutive timepoints (i.e. 0-1h, 1-3h, 3-9h, 9h-7d, named "phases" of differentiation) using the LIMMA package in R (Smyth 2004). This method uses robust estimation of variance combined with a standard t-test to increase the accuracy of the statistical result, which could be affected by the low number of replicates used in this study (n=3). Robust estimation of variance entails using an estimated variance based not only on the probeset in question but on the variance of other probesets with similar expression intensities on the same array.

To correct for false positives (probesets classified as significant by chance) caused as a result of multiple testing, the false discovery method (FDR) of Benjamini and Hochberg was employed (Benjamini and Hochberg 1995). Q values, the result of FDR, are obtained by multiplying the p value for a particular probeset by the number of probesets that have already been tested, if genes are ranked according to their p-value and those with the smallest p-value are tested first. This results in q values for each probeset that are higher than their p values; thus, if a particular significance cut-off is chosen, less genes will be classified as significant but there is less likelihood of identifying false positives. The significance cut-off of q<0.05 was chosen to select genes with significant expression changes in at least one phase of differentiation.

Boxplots were created using the "affy" package in R to demonstrate the expression value data for each array before normalisation. Pearson correlation coefficients (PCC) were also calculated in R. The PCC is a measure the strength of a linear relationship between two variables. Thus, PCCs close to 1 indicate a strong positive correlation. Each permutation of two replicates at each timepoint (i.e. 0hA-0hB, 0hA-0hC, 0hB-0hC) was compared in this manner and an average PCC for each timepoint was calculated from the mean average of the three PCCs obtained.

Average linkage clustering using Euclidean distances was performed using entire gene expression profiles from each array using the "affy" package in R to examine the relationships between the expression profiles of each array. Euclidian distances allow a method of measuring the dissimilarity between arrays, or probesets. Agglomerative hierarchical clustering using Euclidean distances was performed using the 6319 probesets (selected as having at least one significant expression change during adipogenesis according to the significance cut-off used) to group probesets according to their expression pattern during differentiation, and results were graphically represented on a heatmap using the Cluster and Treeview software of Eisen (Eisen et al 1998).

2.7.4 Global annotation of adipogenesis

2.7.4.1 Automated annotation

Automated annotation of global molecular processes during adipogenesis was achieved using the Gene Ontology (GO) classification system (1.7.1.2) and WebGestalt (WEB-based GEne SeT AnaLysis Toolkit, Zhang *et al* 2005). Briefly, this analysis tool retrieves annotations for each input ID (here, Affymetrix probeset IDs were used) in a given list, including GO categories or KEGG metabolic pathways. To calculate statistically significant enrichment, the number of genes in a specific GO category or KEGG pathway is compared in two conditions: the sample (here, the gene list for each expression cluster, uploaded to WebGestalt) and a reference (here, the gene list for the entire Affymetrix U133 Plus 2.0 GeneChip, available in WebGestalt). As these lists are not independent of each other, the hypergeometric test (Zhang *et al* 2005) was used to identify GO categories/KEGG pathways that were significantly enriched in the sample list, with the significance cut-off or $p<0.05$ used (unless otherwise stated). WebGestalt does not provide the facility to correct for multiple testing.

Analysis of the temporal expression of signalling intermediates in the Wnt, TGFβ and BMP pathways during adipogenesis was performed using a free trial of the Ingenuity Pathway Analysis software (Ingenuity® Systems, www.ingenuity.com). Briefly, data files containing the $\log_2$-fold change values, plus their corresponding significance (q) value, for each gene in each "phase" of differentiation (Appendix Table 1), were uploaded to the website. The software represented these changes on interactive pathway "maps", with red shades indicating degrees of upregulation during a differentiation phase (defined by the fold-change value), and green shades indicating downregulation.

2.7.4.2 Manual annotation

Manual annotation of gene function was performed systematically for each expression cluster indicated. Manual searches of the GeneCards (www.genecards.org), NetAffix (www.affymetrix.com/analysis/index.affx) OMIM, EntrezGene and PubMed (all at www.ncbi.nlm.nih.gov) databases were performed for each gene. Data was compiled in a single file for each expression cluster and summarised to provide an overview of gene function.

CHAPTER 3. OPTIMISATION OF hMSC ISOLATION AND
ADIPOGENESIS, AND DEFINITION OF THE
DIFFERENTIATION TIME-COURSE.

3.1 Introduction and aims

Human MSCs possess multiple stem cell properties (1.4.4), including the *in vitro* capacity to differentiate clonally into adipocytes, osteocytes and chondrocytes (Pittenger *et al* 1999). The ability to achieve robust adipogenesis in culture provides the opportunity to study the molecular events underlying differentiation. However, many different hMSC isolation, growth and differentiation protocols have been used by other laboratories, and it is therefore necessary to determine the optimal conditions for experimentation in this system. Additionally, the fact that different protocols exist makes defining the particular conditions used crucial, so that comparisons between this thesis and other work can be made.

Human MSC-like cells can be isolated from numerous locations in the body (1.4.5), including umbilical cord blood (Erices *et al* 2000, Lee *et al* 2004a) and adipose tissue (Zuk *et al* 2001, 2002), but the most well characterised population are bone marrow-derived. The classic method of hMSC isolation from this location is via plastic adherence (Pittenger *et al* 1999), but other methods including antibody-aided selection are also being pioneered (Gronthos *et al* 2003a, 1.4.3). Cell surface marker profiles and differentiation assays are frequently used to define these cell populations as hMSCs. The expression of a range of antigens (Table 1.2), the most specific being STRO-1 (Simmons and Torok-Storb 1991, Dennis *et al* 2002, Gronthos *et al* 2003a), is currently used to define hMSCs. Various methods of hMSC culture have been described, involving the use of different growth media, serum concentrations and supplementation of cytokines or growth factors (for example, Pittenger *et al* 1999, Conget and Minguell 1999, Wagner *et al* 2005). These differences between research groups will inevitably lead to alterations, however slight, in the biology of the cells. It is therefore necessary to define the properties of the cell population under investigation, as well as the effects of any supplements used. The addition bFGF to hMSC growth media during expansion has been employed in several studies, and has been reported to increase the growth rate and life-span of hMSCs without diminishing their multilineage potential (Tsutsumi *et al* 2001). In other work, bFGF has been reported to enhance osteogenic (Hanada *et al* 1997), chondrogenic (Solchaga *et al* 2005) and adipogenic (Neubauer *et al* 2004) potential of hMSCs from a variety of sources.

Thus, the use of bFGF may have advantages, but its effects should be fully described in order to relate the current work to other research.

A cocktail of chemical reagents including insulin, IBMX and dexamethasone has been developed that maximises the efficiency of preadipocyte differentiation (1.6.1). Modifications of this cocktail, such as the addition of indomethacin (Williams and Polakis 1977, Lehmann et al 1997) have been made for inducing adipogenesis in other cell types including hMSCs. However, the direct targets of these inducers have on the whole not been identified (1.6.1 and Fig 1.4), and therefore the relative contribution of each individual substance to hMSC adipogenesis is unknown. Through studies of adipogenesis in murine preadipocyte cell lines such as 3T3L1 and 3T3-F442A, extensive detail has been revealed regarding the molecular control of adipogenesis (1.6.2). PPARγ lies at the heart of the process, and regulation both up- and downstream of this factor acts in a cascade manner to bring about the eventual expression of adipocyte marker genes. The adipogenic role of PPARγ was definitively demonstrated by it's ectopic expression in NIH-3T3 fibroblasts, which do not normally differentiate into adipocytes (Tontonoz et al 1994b). PPARγ over-expression led to adipocyte formation, but required the additional presence of dexamethasone and a PPAR ligand. The effect of PPARγ over-expression in hMSCs has not before been tested, and it would be interesting to reveal whether PPARγ alone is sufficient to induce adipogenesis, or whether additional inducers are required. Ectopic expression of other genes in hMSCs has been achieved (for example, Shi et al 2002, McBeath et al 2004), with the aim of investigating the role of such genes in these cells. It would be advantageous to establish such a technique in this system, which could subsequently be used to assess the role of interesting genes identified in later work.

As preadipocytes are committed to the adipocyte lineage, the regulatory factors involved in their differentiation are likely to control post-commitment events, which could therefore be considered to occur relatively "late" during differentiation. In contrast, hMSCs are uncommitted progenitors of this lineage, so their differentiation would involve additional regulation that occurs prior to, or during, commitment. A major step towards investigating these "early" differentiation events is to define the point during the hMSC differentiation time-course at which they occur. However, this is difficult as few factors that control early

hMSC differentiation are known. One technique to achieve this could be to identify the time at which "late" regulators, such as PPARγ, are expressed during adipogenesis in the system used in this study. Furthermore, several time-course studies of hMSC adipogenesis have been performed, and have described genes that are expressed at specific timepoints during differentiation, the earliest of those being day 1 (Nakamura *et al* 2003, Sekiya *et al* 2004). Studying the relationship of this work to the hMSC differentiation system used in this thesis would be instrumental in defining the time-frame of the cascade of transcriptional regulation involved in hMSC adipogenesis. Importantly, this would also allow identification of the specific phase of differentiation that may represent the early stages of interest, so should be the focus of further work in this thesis.

Aims

The major goal of these experiments is to isolate hMSCs and define their time-course of differentiation to adipocytes; adipogenic differentiation of hMSCs had not previously been established in this laboratory. This would provide a background on which future work will be based. Specific aims required to achieve this are:

- To isolate hMSCs from bone marrow aspirates and characterise the cell population that will be used for subsequent studies
- To optimise the growth and differentiation protocols to be used
- To establish adipogenic differentiation of hMSCs and investigate the contribution of specific factors to this process
- Finally, to investigate the time-course of hMSC differentiation and determine the timepoints which represent the earliest regulatory stages of adipogenesis.

3.2 Isolation of hMSCs and optimisation of growth protocols

Human MSCs are most often isolated from bone marrow aspirates. Some variability between research groups has been reported in terms of cell surface antigens and growth properties of the isolated cell populations. It was therefore the aim of these experiments to isolate hMSCs from bone marrow and characterise their properties in detail, in order to fully define the cell population that would be used in future experiments.

3.2.1 Human MSC isolation

Fresh bone marrow samples from healthy transplant donors who had given research consent were used for the isolation of hMSCs. It was chosen to isolate hMSCs via the plastic adherent method as it is efficient and does not require the specialised equipment used in some antibody-mediated cell sorting techniques. Additionally, there is currently no agreement regarding the particular antibodies that definitely define hMSCs (1.4.3). Bone marrow aspirates were processed as described in section 2.2.1 to isolate the bone marrow mononuclear cell fraction. These cells were plated onto plastic dishes in hMSC growth media supplemented with 1ng/ml bFGF (see section 3.2.3) and 10% human serum supplement. After two days the growth media was changed, and contaminating non-adherent cells, which may include populations such as HSCs (Pittenger *et al* 1999), were mostly removed by this process. By this stage individual (or small clusters of) adherent, spindle-shaped fibroblastic cells were visible (Fig. 3.1A). At day 7 after isolation, clusters of spindle-shaped cells had formed (Fig. 3.1B). When the cell density of these clusters was nearing confluence (this occurred on average at day 7), cells were detached using trypsin and passaged at a ratio of 1:4. Cells were then re-plated at approximately 4×10^3 cells/cm^2 in growth media containing bFGF; this passage was designated passage 1. For expansion, cells were maintained at sub-confluent levels and passaged at 70% confluence. Sub-confluent cells exhibited an elongated, spindle-like morphology (Fig. 3.1C), and on reaching confluence retained this morphology and exhibited a distinct swirling arrangement (Fig. 3.1D). The morphology and growth characteristics of the isolated cells were consistent with previous reports for hMSCs (Pittenger *et al* 1999, Conget and Minguell 1999).

114
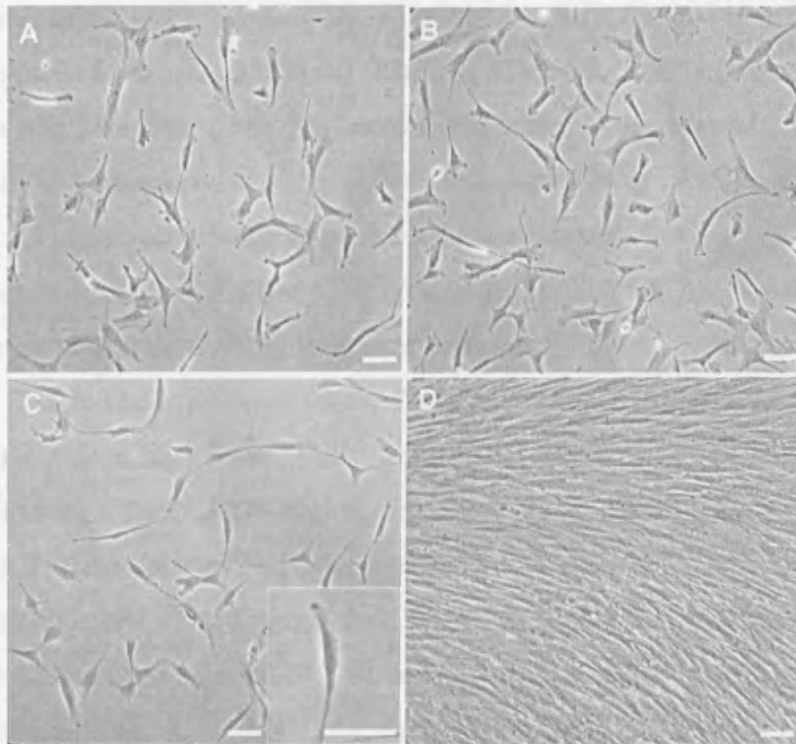
Figure 3.1 Representative illustration of the stages during isolation and expansion of hMSCs from bone marrow aspirates. Bone marrow mononuclear cells were isolated from fresh bone marrow via density gradient centrifugation, washed and plated at a density of approximately 2×10^5 cells/cm^2 in medium containing 1ng/ml –bFGF (2.2.1); A, Day 2 after isolation: adherent cells are visible. Non-adherent cells were removed and fresh media added; B, Day 7 after isolation: clusters of hMSCs have formed and cells were passaged; C, cells were maintained at subconfluent levels during expansion; D, For differentiation, cell monolayers were grown to confluence. Scale bars represent 150μm. Panels A and B are the work of S. Elliman.

3.2.2 Properties of the hMSC population

MSCs are reported to express a defined selection of cell surface markers (Pittenger *et al* 1999, Conget and Minguell 1999, Majumdar *et al* 2003, Wagner *et al* 2005). Therefore, adherent cells at passage 4 was harvested at sub-confluent levels and stained with antibodies (2.3.1) to detect a range of markers that define hMSCs (Fig. 3.2 and Table 1.2). The population was positive for the STRO-1 antigen, which has previously been used to isolate mesenchymal cells with multilineage potential (Simmons and Torok-Storb 1991, Dennis *et al* 2002). The expression level of this marker was variable, with a minority of cell expressing it to a high level. Cells were also positive for SH2 and SH4, and the adhesion molecules CD49b, CD49d (integrins α2 and α4), CD44 (hyaluronate receptor), and CD13. All of these markers have previously been used to positively distinguish hMSCs (Pittenger *et al* 1999, Minguel *et al* 2001). Human MSCs have been reported to be both positive (Conget and Minguell 1999) and negative (Pittenger *et al* 1999) for CD49d; this could reflect some variability between isolation/culture methods or length of cell culture time before marker expression was assessed. Cells were negative for markers of the haematopoietic lineage including c-Kit (CD117), CD45, and CD34, indicating that this cell population was not haematopoietic in origin, and also that there was no contamination from haematopoietic cells. CD133, a stem cell-specific marker, was also not expressed, but this has been previously reported for hMSCs (Wagner *et al* 2005). Finally, expression of HLA-ABC (MHC class I) but not HLA-DR (MHC class II) was detected, which is in line with other work characterising hMSCs (Majumdar *et al* 2003, Wagner *et al* 2005). Thus, the expression marker profile detected in this study was comparable to previous reports for hMSCs, so from hereon it is considered that the adherent, fibroblastic cell population represents hMSCs.

To investigate their differentiation ability, hMSCs isolated from a single bone marrow were induced to differentiate into adipocytes, osteocytes and chondrocytes using specific induction protocols (2.2.2). Many reports of multilineage differentiation use heterogeneous hMSC populations (Tsutsumi *et al* 2001), although it has also been shown that clonal hMSC populations possess this ability (Pittenger *et al* 1999). Here, it was shown that the heterogeneous hMSC population could differentiate robustly to cells of the adipogenic (Fig. 3.3A, as assessed by Oil Red O staining of lipid deposition), osteogenic
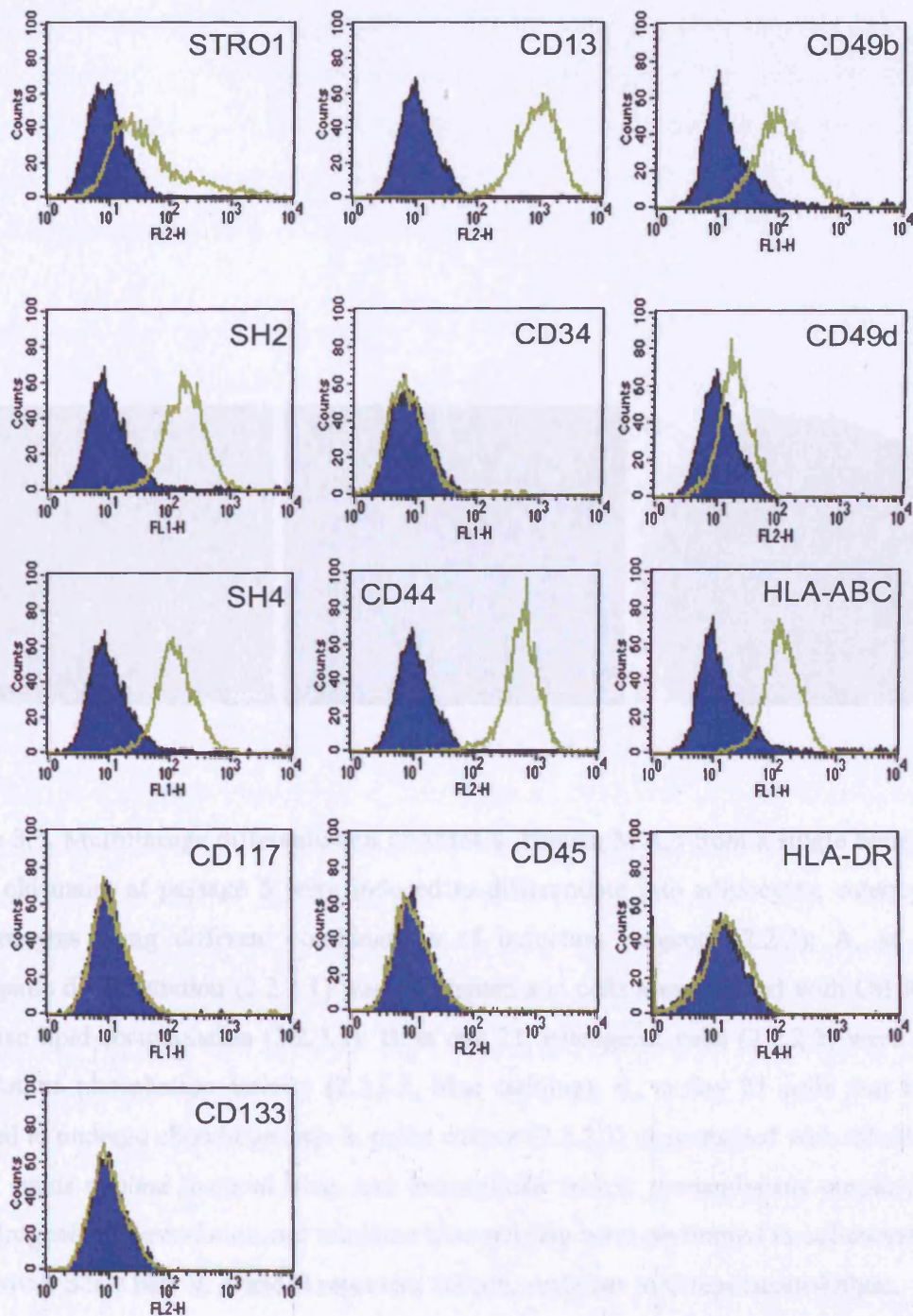
Figure 3.2. Flow cytometry analysis of cell surface markers of hMSCs. Subconfluent, passage 4 hMSCs were harvested, washed and stained with the appropriate antibody (5×10^5 cells per sample), and then secondary antibody if required (2.3.1). Expression was determined by counting 10^4 events per sample. Shaded histograms (blue) indicate secondary antibody controls, and green lines indicate cells stained with the marker-specific antibodies.
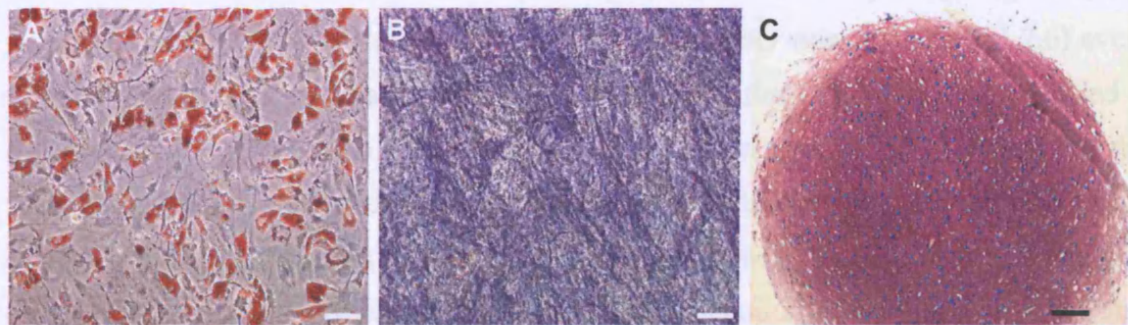
Figure 3.3. Multilineage differentiation of hMSCs. Human MSCs from a single bone marrow (24yr old male) at passage 5 were induced to differentiate into adipocytes, osteocytes and chondrocytes using different combinations of induction reagents (2.2.2); A, at day 14, adipogenic differentiation (2.2.2.1) was terminated and cells were stained with Oil Red O to visualise lipid accumulation (2.2.3.1); B, at day 21, osteogenic cells (2.2.2.2) were assessed for alkaline phosphatase activity (2.2.3.2, blue staining); C, at day 21 cells that had been induced to undergo chondrogenesis in pellet culture (2.2.2.3) were stained with toluidine blue, which stains nuclear material blue, and extracellular matrix proteoglycans purple (2.2.3.4). Chondrogenic differentiation and toluidine blue staining were performed in collaboration with S. Vujovic. Scale bars in A and B represent 200μm, scale bar in C represents 400μm.

(Fig. 3.3B, as assessed by alkaline phosphatase activity) and chondrogenic (Fig. 3.3C, as assessed by toluidine blue staining of extracellular matrix proteoglycans) lineages (2.2.3). Human MSC populations isolated from four different bone marrows that were assessed in this manner had similar differentiation capabilities. Thus, the hMSCs described in this thesis exhibited multilineage differentiation properties.

Finally, the *in vitro* growth potential of five hMSC populations was determined. Human MSCs isolated from different bone marrow donors were seeded at passage 4 (p4) - p6 at 6250 cells/cm^2 in media containing bFGF. Total cell number was quantified (2.2.6) every three days when cells were passaged (before contact inhibition), and cells were reseeded at 6250 cells/cm^2 every passage. This was continued for 9 passages, and it was found that the majority of bone marrows exhibited comparable growth kinetics and achieved an average of 15.8 population doublings in this time (Fig. 3.4). One bone marrow (55yr old female) exhibited substantially slower growth, ceased to proliferate after around 5 passages and achieved a total of only 6.7 population doublings. This could possibly be due to the age of the donor, and suggests that hMSCs from different bone marrows are heterogeneous in their expansion potential. Further, these data reveal that the majority of hMSC populations can be expanded for at least 15 population doublings with no significant loss of proliferation, although population doubling time increased slightly every passage. When the ability of hMSCs from a single bone marrow to undergo adipogenesis at different passages was assessed, it was found that adipogenic differentiation potential at passage 9 was less than that at passage 5, and by passage 15 hMSCs had virtually lost their adipogenic differentiation potential (Fig. 3.5). As with the surface marker profile and multilineage differentiation capability, these data are in line with most other studies of hMSC growth and differentiation. Taken together, these results demonstrate that the cells used in subsequent experiments in this thesis are analogous to hMSCs described in the literature (Bruder *et al* 1997, Pittenger *et al* 1999, Conget and Minguell 1999).

3.2.3 The effect of bFGF on hMSC growth and differentiation

Several methods have been employed to maximise hMSC outgrowth and expansion potential in culture. These include selection of serum batches with the highest ability to promote outgrowth of hMSCs with differentiation potential (Pittenger *et al* 1999), or

119
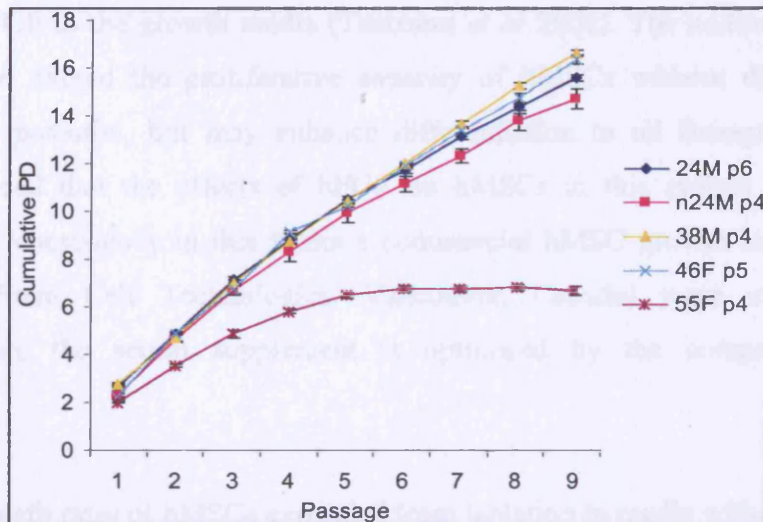
Figure 3.4. Growth rates of hMSCs isolated from different donors. Frozen hMSCs isolated from several bone marrows (age and gender indicated) were expanded in bFGF-supplemented media. At the passage indicated for each bone marrow, cells were seeded at 6250 cells/cm^2 in triplicate for each growth curve. Three days later cells were counted and reseeded at this density; this process was repeated for a total of 9 passages (2.2.6). Data represent cumulative population doublings (PD) and are mean ± standard deviation (SD). Experiments performed by S. Vujovic, data analysis performed by C. Westwood.
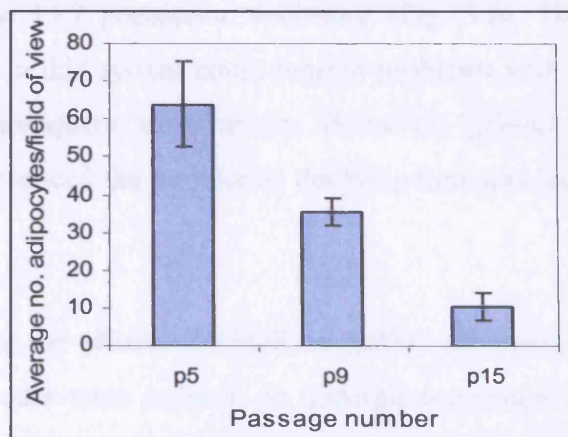
Figure 3.5. Adipogenic ability of hMSCs after continued passaging. MSCs from a single bone marrow (age/gender unknown) were expanded for the indicated number of passages, and then grown to confluence in media supplemented with bFGF. Adipogenic differentiation was induced as described (2.2.3.1). At day 9, differentiation was terminated and quantified by counting the number of adipocytes in two fields of view for each condition. Data represent number of adipocytes per field of view and are mean ± SD.

120

addition of bFGF to the growth media (Tsutsumi *et al* 2001). The addition of bFGF has been shown to extend the proliferative capacity of hMSCs without diminishing their differentiation potential, but may enhance differentiation to all lineages (3.1). It was therefore decided that the effects of bFGF on hMSCs in this system would be fully described. For consistency in this thesis a commercial hMSC growth media and serum supplement (Stem Cell Technologies, Vancouver, Canada) were used throughout experimentation; the serum supplement is optimised by the company for hMSC expansion.

Firstly, the growth rates of hMSCs expanded from isolation in media with/without 1ng/ml bFGF supplement were investigated. At passage 4, cells from both conditions were seeded at 5200 cells/cm^2. Three days later, +bFGF hMSCs were at ~70% confluence and needed passaging, whereas by day 6 after seeding, -bFGF cells were near to, but had not reached, this level of confluence. Therefore, +bFGF cells were passaged every 3 days for a total of 10 times, whereas -bFGF cells were passaged every 6 days, for a total of 5 times. At each round of expansion, the total number of cells was counted. By day 30, -bFGF cells had undergone a total of 5.0 population doublings on average, whereas +bFGF cells had undergone on average 15.7 population doublings (Fig. 3.6). The slow growth rate of -bFGF cells obtained in this system could present problems with obtaining enough early passage cells for subsequent experiments. However, growth of hMSCs in bFGF-supplemented media reduced the population doubling time and led to an abundant supply of low passage cells.

In order to investigate the effects of bFGF on hMSC differentiation, cells expanded in +bFGF or –bFGF media were induced to undergo osteogenesis or adipogenesis with lineage-specific induction cocktails. In addition, the direct effect of bFGF on differentiation was assessed by including it in the differentiation media. After 14 days of differentiation in the absence of bFGF, the extent of adipogenesis was assessed by Oil Red O staining and it was found that hMSCs expanded both with and without bFGF differentiated to a similar extent (Fig. 3.7A). The rate of differentiation was also similar with and without bFGF, as assessed by the timepoint that lipid accumulation was first observed (data not shown). Interestingly, it seemed that the presence of bFGF in the
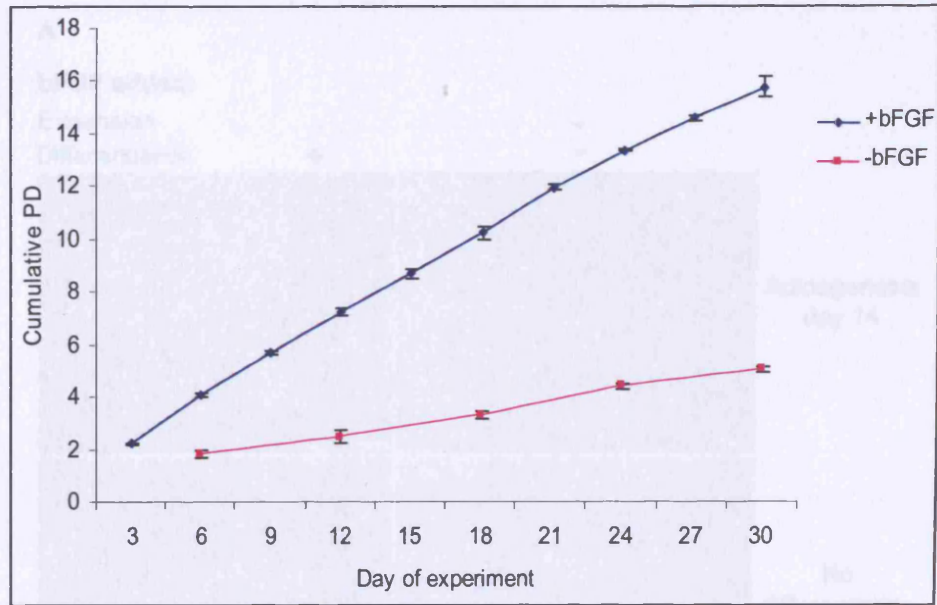
Figure 3.6. The effect of bFGF on hMSC growth rate. hMSCs were isolated from a single bone marrow (29yr old male), and immediately suspended in media with or without 1ng/ml bFGF. Cells were maintained in these conditions throughout the experiment. At passage 4, cells were seeded at 5200 cells/cm^2 in triplicate for each growth curve. Three (+bFGF) or 6 (-bFGF) days later cells were counted and reseeded at this density; this process was repeated for a total of 10 (+bFGF) or 5 (-bFGF) rounds of expansion. Cumulative population doublings (PD) were calculated from cell counts (2.2.6). Data represent cumulative PD and are mean ± SD.

A

bFGF added:

Expansion	−	−
Differentiation	+	−
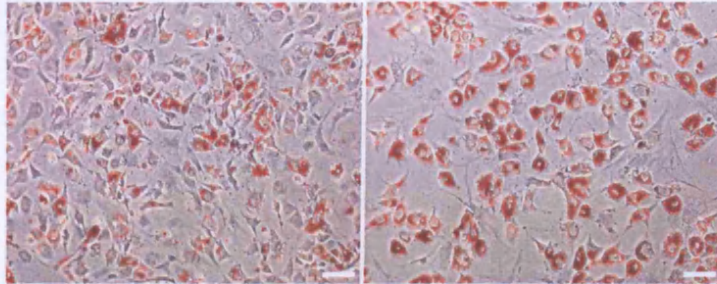
Adipogenesis
day 14
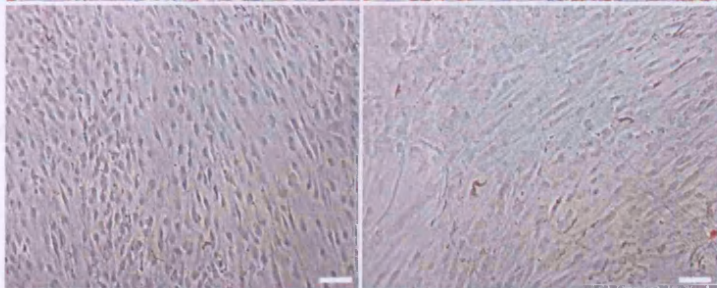
No
differentiation

bFGF added:

Expansion	+	+
Differentiation	+	−
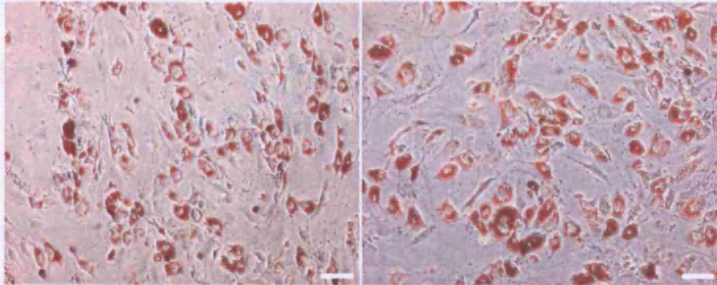
Adipogenesis
day 14
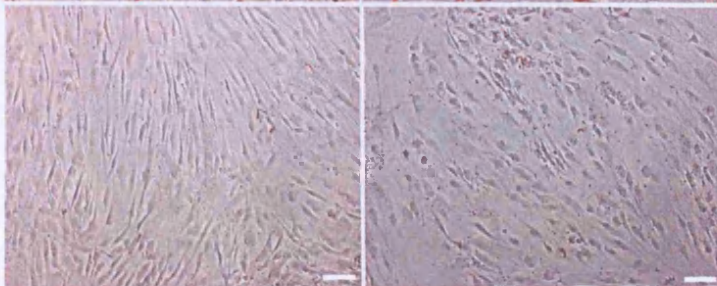
No
differentiation

Figure 3.7. Continued overleaf.

123
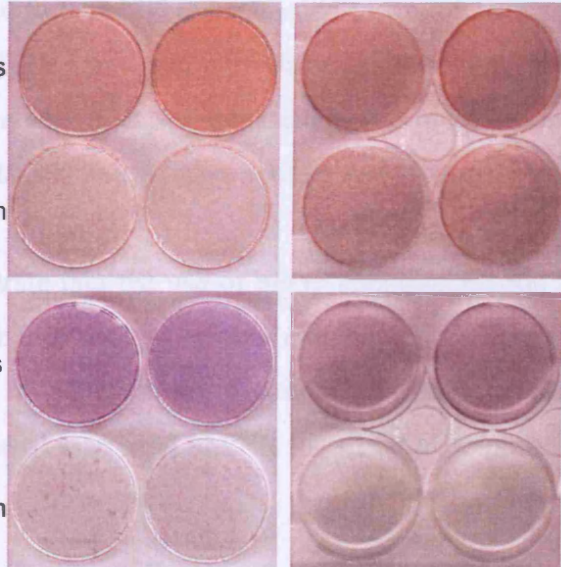
Figure 3.7. The effect of bFGF on hMSC adipogenesis and osteogenesis. Passage 5 hMSCs grown with/without bFGF were induced to differentiate into adipocytes (A, 2.2.2.1) or osteoblasts (B, 2.2.2.2) in the presence/absence bFGF; A, At day 14 of adipogenesis, differentiated and control cells were fixed and stained with Oil Red O (2.2.3.1); B, At day 21 of osteogenesis, cells were fixed and stained with Alizarin Red (2.2.3.3), or assayed for alkaline phosphatase activity. Scale bars represent $200\mu m$.

differentiation media decreased the number of differentiated cells. When osteogenesis was assessed at day 21, a similar effect was observed; there was little difference in the extent of osteogenesis between cells grown with or without bFGF as assessed by staining calcium deposition (using alizarin red) or alkaline phosphatase activity (Fig. 3.7B). However, there seemed to be a higher background level of alizarin red staining when cells were expanded in bFGF but were not differentiated. The highest level of calcium deposition was seen in cells that were expanded and differentiated without bFGF, but a similar pattern with alkaline phosphatase activity was not observed.

These results demonstrate that the use of bFGF supplementation in hMSC growth media allowed a higher number of cells to be obtained than in the absence of bFGF, but without affecting their ability to differentiate to either adipocytes or osteoblasts. Hence, bFGF supplementation was used to expand hMSCs for all future experiments.

3.3 Factors affecting hMSC adipogenesis

3.3.1 Contribution of induction cocktail constituents to hMSC adipogenesis

Various induction protocols for initiating adipogenesis *in vitro* have been reported (for example, Gimble *et al* 1995, Pittenger *et al* 1999, Neubauer *et al* 2004). MSC differentiation is commonly initiated using a modified cocktail originally designed for preadipocyte differentiation, but how each reagent contributes to adipogenic induction is not fully understood. The aim of this experiment was therefore to investigate the effect of individual reagents on adipogenesis, and to identify the optimal induction cocktail for initiating differentiation in this system.

Confluent monolayers of hMSCs were incubated with individual components of the induction cocktail, or combinations thereof. Assessment of lipid accumulation was performed at day 14 (Fig. 3.8). The only substance capable of inducing differentiation on its own, albeit to a low level, was dexamethasone. Differentiation was frequently higher with reagent combinations where dexamethasone was present. Indomethacin also played a significant role in the induction cocktail, and the combination of these two substances led
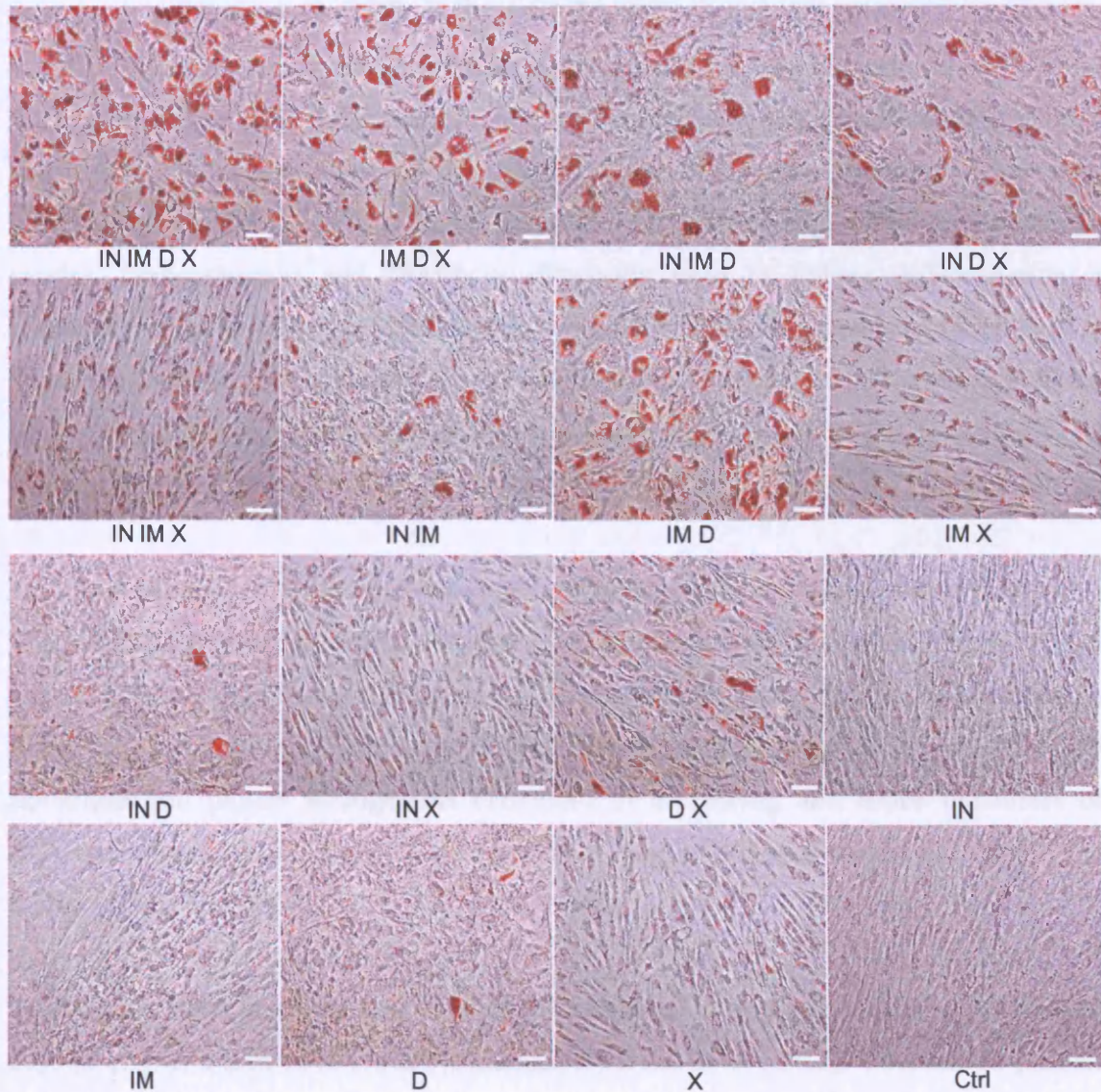
Figure 3.8. Contribution of individual components to the differentiation cocktail. MSCs were grown to confluence then induced to undergo adipogenesis using combinations of components of the differentiation cocktail as shown. Cells were fixed and stained with Oil Red O at day 14 of differentiation (2.2.3.1). Each condition was assessed in triplicate, and a representative image is shown; D, $1\mu M$ dexamethasone; X, $500\mu M$ isobutylmethylxanthine; IN, $10\mu g/ml$ insulin; IM, $200\mu M$ indomethacin; ctrl, no differentiation induced. Scale bars represent $200\mu m$.

to extensive differentiation. The addition of IBMX to this combination did not seem to significantly increase adipogenesis; rather, there appeared to be a higher level of apoptosis in undifferentiated cells. When insulin, dexamethasone and indomethacin were employed, less differentiation was observed than in hMSCs in equivalent conditions but without insulin; however the size of the lipid droplets in each cell was larger. The combination of insulin, dexamethasone and IBMX is classically used to induce differentiation in preadipocyte cell lines; in hMSCs however this combination did not induce a high level of adipocyte formation. The addition of indomethacin to this combination led to the highest proportion of differentiated cells, and thus was employed as the differentiation protocol in subsequent experiments.

3.3.2 Manipulation of hMSC differentiation to adipocytes via overexpression of PPARγ

The possibility of using genetic manipulation to study hMSC differentiation was next assessed, as this may be an efficient means to characterise interesting genes identified in future work. It has previously been shown that overexpression of PPARγ in NIH-3T3 fibroblasts can induce adipogenesis (Tontonoz *et al* 1994b), and hence the effect of PPARγ overexpression on hMSC differentiation was investigated to test the efficacy of this technique in hMSCs.

To this end, RNA was extracted from hMSCs induced to undergo adipogenesis for four days (2.6.1), and then cDNA was synthesised from the RNA (2.6.2). Two sets of primers were designed in order to amplify each isoform of PPARγ separately using the cDNA as a template for PCR (2.6.3.1). PPARγ1 and –γ2 cDNAs were then cloned separately into the retroviral vector pMSCV-Hygro (Fig. 2.1 and Table 2.1, 2.4.3). Virus was produced from these constructs via transfection of the retroviral packaging cell line 293gp (2.5.1). Next, hMSCs were transduced with this virus (2.5.3), and hMSC populations stably expressing each isoform of PPARγ, or the empty vector, were established via hygromycin selection (50μg/ml for 4-6 days) of transduced cells (2.5.5). A human liposarcoma cell line LiSa2 (2.2.4) was also transduced and drug-selected (200μg/ml hygromycin for 4-6 days) in the same manner (Wabitsch *et al* 2000). The LiSa2 cell line has been shown to have a high capacity for chemically-induced adipogenesis, and could thus be considered analogous to 3T3L1. The 3T3L1 cell line was not employed as it is murine in origin and thus the human

PPARγ cDNAs used in this study may not function fully in these cells. RNA was extracted from sub-confluent hMSCs and LiSa2s (2.6.1) and cDNA was synthesised (2.6.2) for use in RT-PCR analysis using primers that identify both PPARγ isoforms (2.6.3.1). LiSa2s and MSCs transduced with either PPARγ1 or –2 expressed high levels of PPARγ (Fig. 3.9A). PPARγ was expressed at a level below the detection threshold in empty vector-transduced or untransduced control hMSCs, and was expressed to a low level in the equivalent LiSa2 cells. It was noted that cells expressing either PPARγ isoforms exhibited slower growth than those containing the empty vector, possibly due to the pro-growth arrest properties of PPARγ (1.6.2.1). Transduced cells were grown to confluence, then incubated in hMSC growth media lacking both bFGF and induction reagents. After incubation in this media for 14 days, Oil Red O staining was used to determine whether lipid accumulation (and thus adipogenic differentiation) had occurred. No differentiation occurred in empty vector-transduced cells, but in both LiSa2 and hMSC over-expressing PPARγ1, cells containing several lipid droplets, and therefore resembling adipocytes, were visible (Fig. 3.9B). Over-expression of PPARγ2 induced differentiation more strongly than PPARγ1; the number of cells containing cytoplasmic lipid droplets increased, as did the number of lipid droplets per cell. Over-expression of either PPARγ isoform induced more LiSa2 cells to differentiate than hMSCs. Thus, PPARγ overexpression in hMSCs and LiSa2 cells led to adipogenic differentiation without the need for chemical inducers, although the differentiation rate was lower than in untransduced cells differentiated using chemical reagents.

3.4 Determination of the adipogenic differentiation time-course

The major goal of this thesis is to investigate the earliest transcriptional events of hMSC adipogenesis, and identify novel genes that act at this stage. The aim of this experimental section was therefore to use both morphological and molecular analyses of hMSC adipogenesis to establish the exact stage during differentiation of hMSCs in this system at which the earliest regulatory events are occurring.
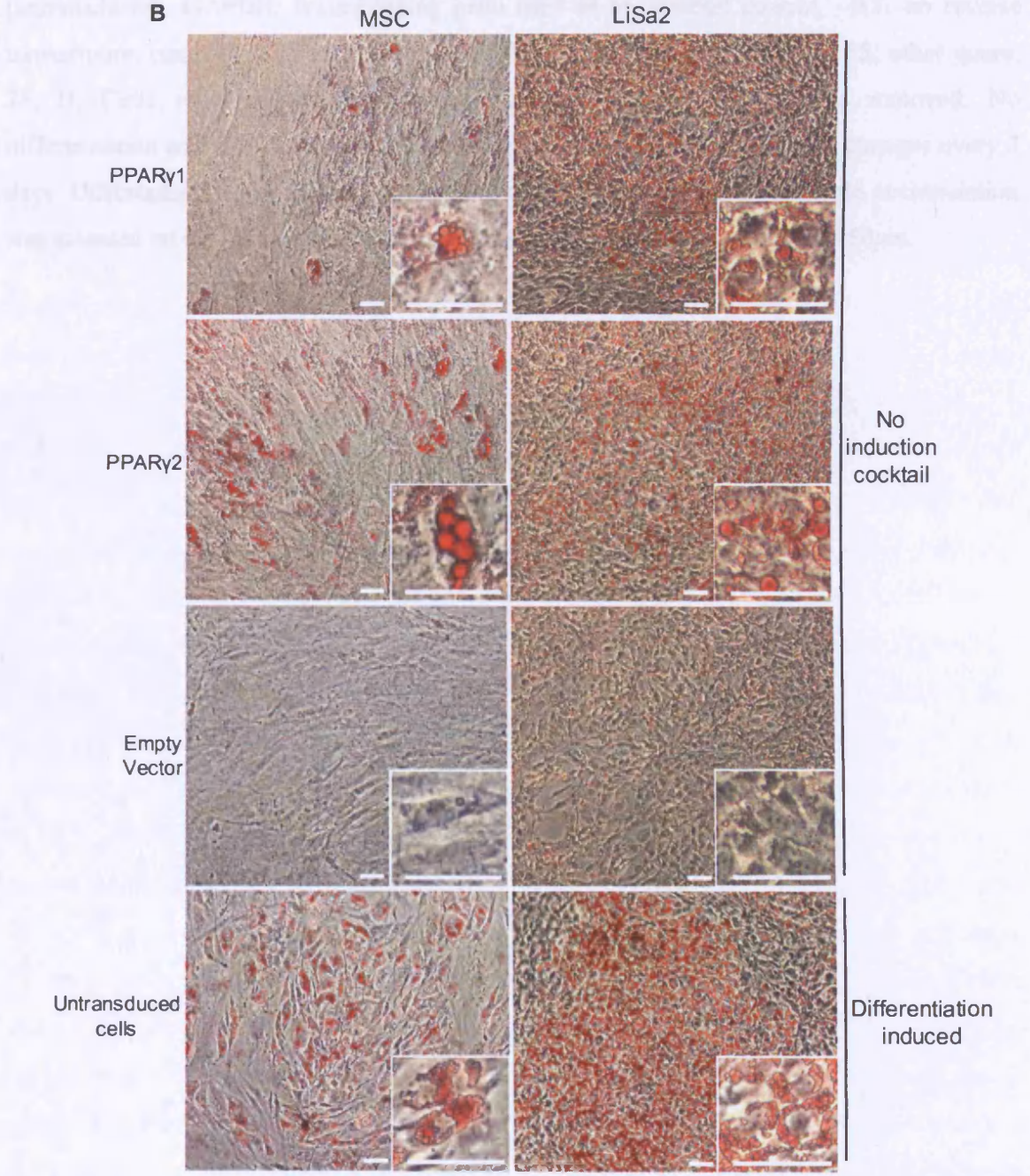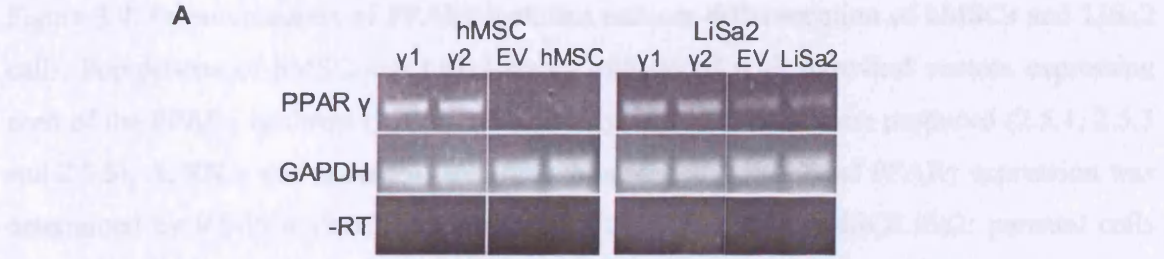
Figure 3.9. Continued overleaf.

Figure 3.9. Overexpression of PPARγ isoforms induces differentiation of hMSCs and LiSa2 cells. Populations of hMSC and LiSa2 stably transduced with retroviral vectors expressing each of the PPARγ isoforms (2.4.3), or the empty vector control, were produced (2.5.1, 2.5.3 and 2.5.5); A, RNA was extracted from subconfluent cells (2.6.1) and PPARγ expression was determined by RT-PCR (2.6.2 and 2.6.3.1). EV: empty vector, hMSC/LiSa2: parental cells (untransduced). GAPDH: housekeeping gene used as an internal control. −RT: no reverse transcription control (using adipsin primers). PCR cycle numbers: GAPDH, 25; other genes, 28; B, Cells were allowed to reach confluence, and then bFGF was removed. No differentiation cocktail was added, and cells were left for 14 days with media changes every 3 days. Differentiation was induced in parental cells as a positive control. Lipid accumulation was assessed on day 14 via Oil Red O staining (2.2.3.1). Scale bars represent 150μm.

3.4.1 Morphological changes during hMSC differentiation to adipocytes

The time-course of adipogenesis was initially studied by assessing lipid accumulation and cell morphology at different times after adipogenic induction. The adipogenic induction cocktail was added to post-confluent MSCs isolated from the bone marrow of a 31 yr old male donor (31♂), and morphological changes were monitored via Oil Red O staining and phase contrast microscopy. By day 1 of differentiation, the majority of cells lost their elongated, fibroblastic morphology and appeared rounded, similar to adipocytes (Fig. 3.10). Small lipid droplets were observed in some cells at the 2nd day of differentiation, and by day 7, cells containing multiple lipid droplets, and thus resembling mature adipocytes, were visible. As the time-course progressed, the number of lipid droplets per cell increased (14d, Fig. 3.10), but the number of cells containing lipid did not increase dramatically. After prolonged incubation in differentiation media lipid droplet size increased, perhaps due to coagulation of smaller droplets (28d, Fig. 3.10). These changes were not observed in hMSCs cultured for the same period in medium lacking inducers.

It was found in earlier experiments that hMSCs isolated from different bone marrows can have different growth rates (Fig. 3.4). The question of whether the rate of differentiation is conserved between hMSCs isolated from different bone marrows was therefore addressed by performing time-course experiments on two additional hMSC populations (from 24♂ and 30♂). As seen in Figure 3.11A and 3.11B (and additionally Figure 3.10), the number of lipid-filled adipocytes visible at day 7 remained fairly constant, and represented 58% (on average) of total cells at this stage, so the extent of differentiation (i.e. the proportion of total cells that differentiated) was similar between hMSCs isolated from different bone marrows. On average, 80% of hMSCs adopted a rounded morphology by day 1, and at day two 30% of cells on average contained lipid droplets, although this varied somewhat between hMSCs from bone marrows (15-40%). Notably, cells containing several small intracellular lipid droplets at day 1 of differentiation were visible in the 30♂ cells (Fig. 3.11B, 1d, inset), whereas in the 31♂ or 24♂ populations lipid was not visible until day 2. Taken together, these observations show that morphological markers of terminal differentiation i.e. rounded cell morphology and intracellular lipid droplets, are visible at day 1-2 of the differentiation time-course, so the process of differentiation is rapid.
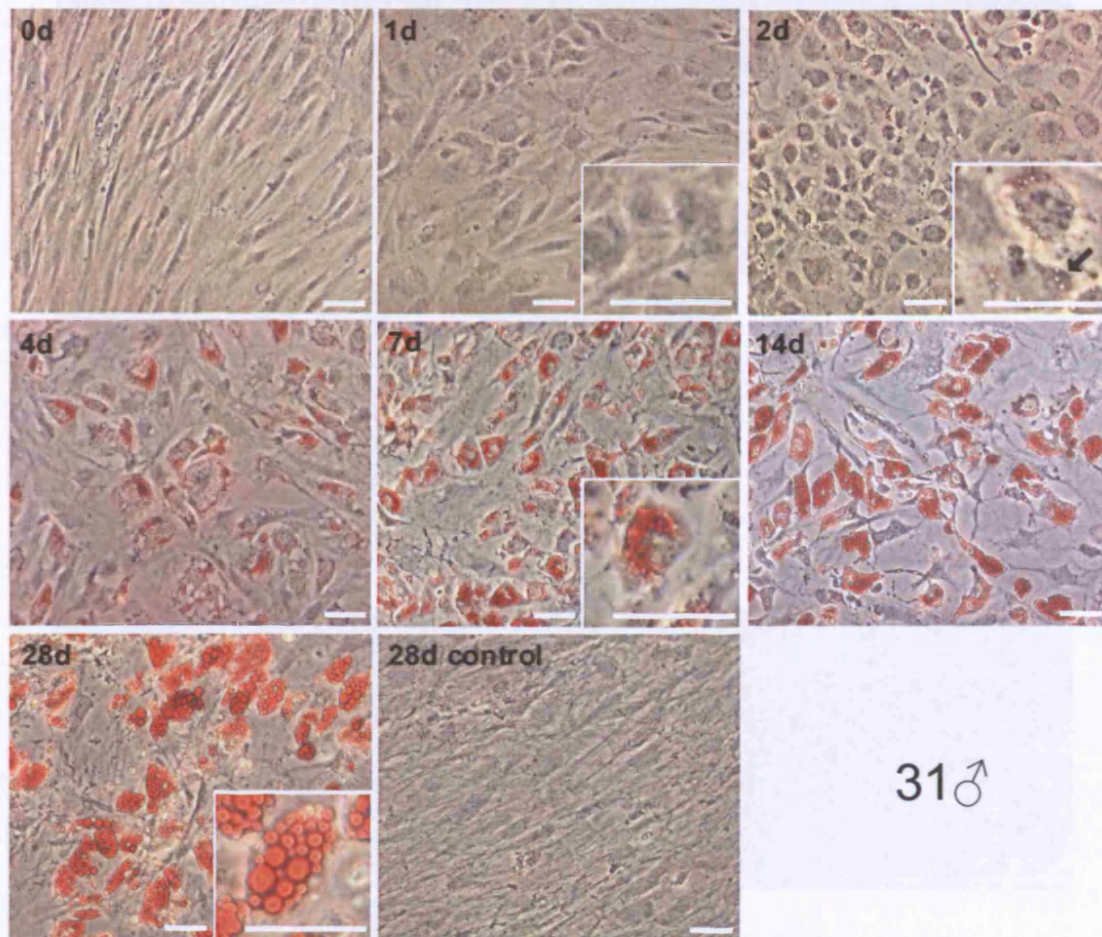
Figure 3.10. Time course of hMSC differentiation to adipocytes. Human MSCs from a single bone marrow (31yr old male donor, passage 6) were grown to confluence and differentiated to adipocytes (2.2.2.1). At the timepoints indicated, cells were fixed, stained with Oil Red O (2.2.3.1) and visualised via phase contrast microscopy. Arrows indicate lipid droplets. Scale bars represent 150μm.
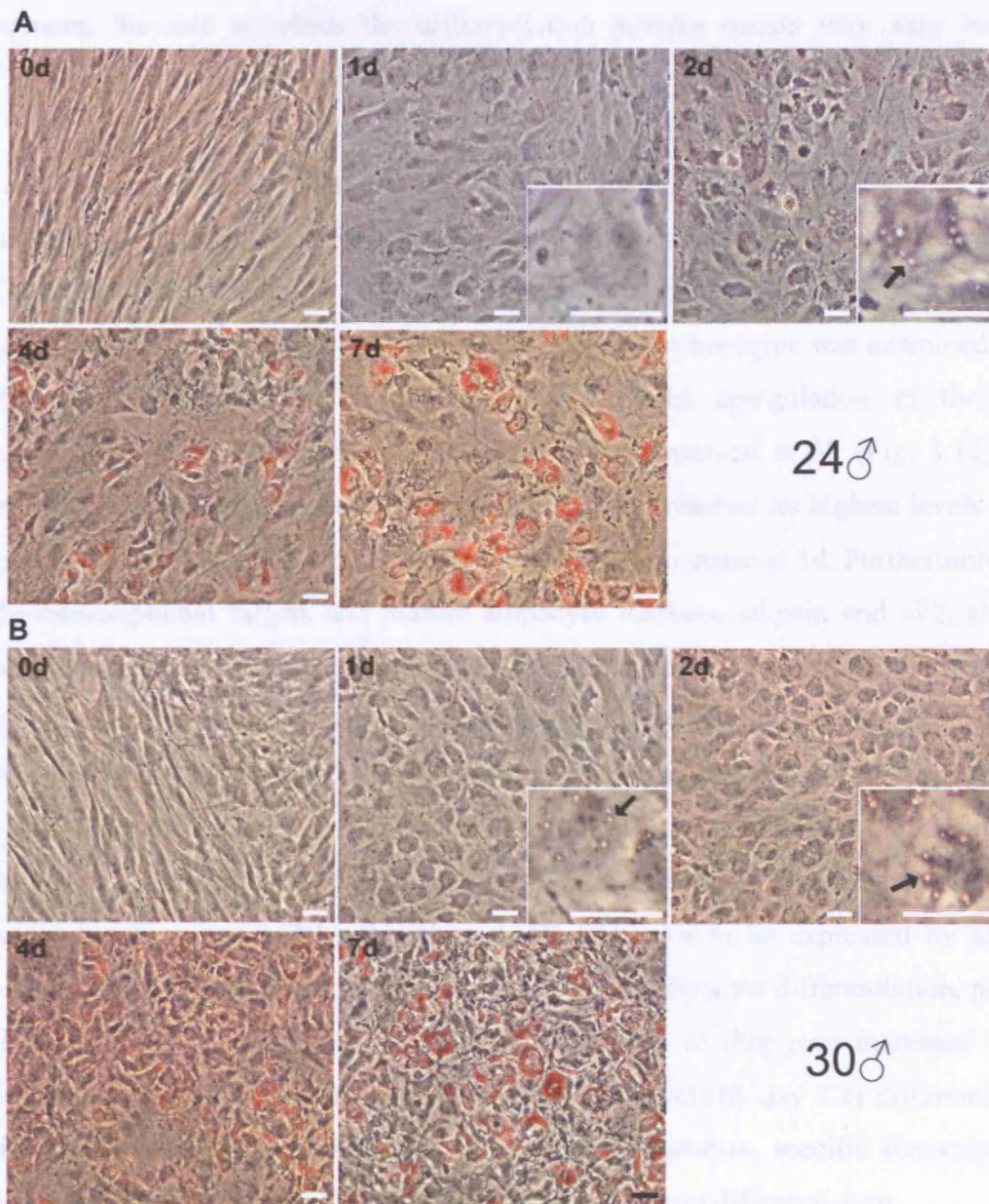
Figure 3.11. Adipogenic differentiation rates of hMSCs isolated from different donors. Human MSCs from different bone marrows (age and gender indicated, passage 6) were grown to confluence and differentiated to adipocytes (2.2.2.1). At the timepoints indicated, cells were fixed, stained with Oil Red O (2.2.3.1) and visualised via phase contrast microscopy. Arrows indicate lipid droplets. Scale bars represent 100μm.

Furthermore, the rate at which the differentiation process occurs may vary between hMSCs isolated from different bone marrows.

3.4.2 Molecular time-course of hMSC differentiation to adipocytes

To study the differentiation time-course at the mRNA level, RNA was collected from hMSCs at various times after adipogenic induction, and the expression of known adipogenic regulators and markers of the mature adipocyte phenotype was examined using RT-PCR (see Table 2.3 for primer sequences). Slight upregulation of the main transcriptional regulator of adipogenesis, PPARγ, was observed at 5h (Fig. 3.12). The expression of this gene increased substantially at 1d, and reached its highest levels at 7d. In concert with this, expression of C/EBPα also began to increase at 1d. Furthermore, two PPARγ transcriptional targets and mature adipocyte markers, adipsin and aP2, showed enhanced expression 1d and 2d, respectively, which increased further at later timepoints. As these factors are involved in the differentiation of preadipocytes or are expressed in mature adipocytes, they can be considered to represent the later phases of adipogenesis. Their upregulation during the first day of differentiation, combined with the observation of cellular lipid accumulation at day 2, indicate that a phenotype reminiscent of mature adipocytes has been attained by this time. LIFR is known to be expressed by hMSCs (Table 1.2), but is also involved in the early stages of preadipocyte differentiation, prior to C/EBPα or PPARγ (Aubert et al 1999). The expression of this gene increased at 5h, reaching a peak at 12h-1d, and then decreased to basal levels by day 7 of differentiation. This indicated that within the first 24h of hMSC adipogenesis, specific transcriptional events occur that are known to characterise early preadipocyte differentiation.

To further define the phase of differentiation in which uncharacterised, early events were occurring, a study focusing on the first day of differentiation was performed. This was accomplished using a small-scale microarray analysis of gene expression, which allowed more rapid determination of the expression of multiple genes than RT-PCR. RNA was extracted from confluent hMSCs prior to differentiation, from 7 day-differentiated cells, and from cells induced to differentiate for 9h. This timepoint was chosen because substantial upregulation of late adipogenic markers (such as PPARγ or aP2, for example) had not yet occurred, but a marker of early preadipocyte differentiation, LIFR, was
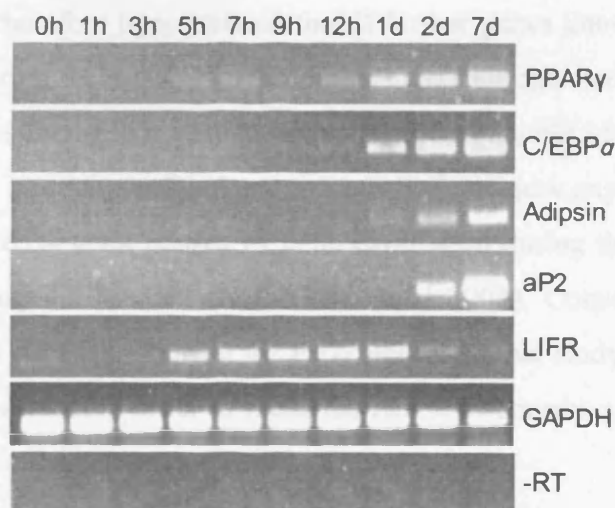
Figure 3.12. Analysis of mRNA levels of adipogenic markers during the differentiation time-course. RNA was extracted from differentiating hMSCs (2.6.1) at the timepoints indicated and RT-PCR (2.6.2 and 2.6.3.1) was used to follow gene expression across the full differentiation time-course. The expression of: major adipogenic transcription factors (PPARγ and C/EBPα), and late (aP2 and adipsin) and early (LIFR) differentiation markers was assessed. PCR cycle numbers: GAPDH, 25; all other genes 28. -RT: no reverse transcription control (adipsin primers).

upregulated. It was therefore hypothesised that if further genes known to be expressed in preadipocytes (or during their early differentiation) were upregulated at 9h compared with 0h or 7d, then studies focusing earlier than this timepoint should identify uncharacterised events. Furthermore, a previous cDNA microarray study of gene expression during hMSC adipogenesis identified several phases of gene expression during the process, and day 1 was the earliest timepoint studied (Nakamura et al 2003). Consequently, genes were selected that showed regulation during the early phases in that study and their expression was analysed in this system in order to relate the two time-courses and thus further define the unexplored areas of hMSC adipogenesis.

RNA from each timepoint was used to synthesise biotin-labelled cRNA probes (2.7.1), which were hybridised to Affymetrix U133A GeneChip Arrays (2.7.2). One array per timepoint was used as this was an exploratory experiment, so statistical significance of expression change between timepoints could not be calculated. Because of this, results were interpreted with caution, but the approach was deemed acceptable as this analysis was performed only in order to design the full experiment described in Chapter 4. The RMA method of Irizzary et al (2003a) was used to background correct and normalise the data, and produce expression values for each gene at 0h, 9h and 7d (2.7.3 and 4.2.2). "Significant" expression changes between successive timepoints (i.e. between 0h-9h or 9h-7d) for all selected genes were then determined by comparing expression values (in RMA units) at each timepoint. A difference of one expression unit (which represents an expression fold change of 2, as RMA units are on a $\log_2$ scale, 2.7.3) between two timepoints was set as the arbitrary significance cut-off, and genes with at least one "significant" change were represented in a heatmap (Fig. 3.13). A blue colour represents expression below the mean average for a particular gene over the timepoints studied, a red colour represents expression above the average, and change is measured in standard deviations from the mean.

On the whole, every gene analysed that was known to be expressed in preadipocytes or during their early differentiation, such as C/EBPδ, and/or had been identified as upregulated at day 1 (the earliest phase) in the Nakamura study (red boxes) were expressed at the highest level at 9h (Fig. 3.12). The only exception was SOD2, which was

136

0h 9h 7d

SREBP1
STAT5A
C/EBPδ
aP2
C/EBPα
Adipsin
PPARγ
Insulin-induced gene 1 (INSIG1)
Lipoprotein lipase
Retinoid X receptor A (RXRα)
Adipose differentiation-related protein (ADRP)
IGF1 Receptor
Growth arrest specific 1 (GAS1)
LIF Receptor (LIFR)
GADD45B
Dickkopf homolog 1 (DKK1)
Superoxide dismutase 2 (SOD2)
Vascular cell adhesion molecule1 (VCAM1)
Stromal-cell derived factor 1 (SDF1)
Disabled homolog 2

Nakamura *et al* (2003) observations

Upregulation at:
1d
1-3d, 7d
7-14d
14d

Downregulation at:
1-3d
1-3d, 7d

-3.0 3.0

Figure 3.13. Exploratory microarray analysis of early hMSC adipogenesis. RNA from undifferentiated, 9h- and 7 day-differentiated cells was used to synthesise biotin-labelled cRNA probes (2.7.1). These were hybridised to U133A Affymetrix microarrays (2.7.2), with one array per timepoint. Data was processed as described (2.7.3). The expression of known adipogenic markers, or genes expressed at specific timepoints in a recent hMSC adipogenesis microarray study (Nakamura *et al* 2003) was investigated in order to determine the time-course of differentiation in this system. On the heatmap, a blue colour represents expression below the mean average for a particular gene over the timepoints studied, a red colour represents expression above the average, and change is measured in standard deviations from the mean (scale bar below heatmap). Microarray data processing and production of the heatmap was performed in collaboration with Dr. M. Trotter.

137

upregulated at 9h compared to 0h, but expression was highest at 7d. A further pattern that was identified in the Nakamura study was genes upregulated at 1-3d and also 7d (orange boxes); these genes showed above average expression at 9h in this system. Additionally, genes that had shown transient downregulation at 1-3d in the work by Nakamura *et al* here were downregulated at 9h, and Disabled homolog-2, which showed a further phase of downregulation at 7d in that work, was also identified as further downregulated at 7d. Genes known to characterise the later stages of adipogenesis, and that were upregulated at the last two timepoints in the Nakamura study, were upregulated only at 7d in this work.

To summarise, this analysis shows that the series of transcriptional events known to characterise adipogenesis, and also identified previously during hMSC differentiation, occurred in the same sequence in this study (Rosen and Spiegelman 2000, Nakamura *et al* 2003). Furthermore, the 9h timepoint was found to closely resemble the earliest timepoint in the Nakamura study, and also represented a possible "preadipocyte stage", as several preadipocyte genes were expressed at this point. Therefore, future studies focusing prior to the 9h timepoint would be most likely to uncover the transcriptional events of early hMSC adipogenesis.

3.5 Discussion

This work has documented the isolation of a fibroblastic, adherent cell population from human bone marrow that displays the cell surface marker profile, differentiation potential and expansion capability characteristic of hMSCs. Further, adipogenic differentiation of these hMSCs was established, and was found to be stimulated to different extents by combinations of chemical inducers, or by PPARγ overexpression. Finally, hMSC differentiation was shown to occur rapidly under optimal conditions, and investigation of the expression of known adipogenic regulators revealed that the early events of interest were likely to occur within the first day, and more specifically within the initial 9 hours, of differentiation.

Some variability in the phenotype of hMSCs has been reported that is likely to be due to different isolation or growth protocols, so the hMSC population that was used in this study was characterised in order for comparisons to other work to be made. Both the morphology and surface marker profile (Figs. 3.1 and 3.2) of this hMSC population strongly resembled those documented elsewhere (Pittenger *et al* 1999, Conget and Minguell 1999, Majumdar *et al* 2003, Wagner *et al* 2005). When the surface marker profile of hMSCs was assessed, many markers defining hMSCs, such as SH2, SH4 and CD13 were uniformly expressed, suggesting that a uniform hMSC population had been isolated, and no contamination from other cell populations was detectable. A point of variability in this work compared with some published analyses was the expression of CD49d by hMSCs. However, this variation is not unique to this study, as hMSCs have been shown to be both positive and negative for CD49d (Pittenger *et al* 1999, Conget and Minguell 1999). Thus, as no unexpected variation in surface marker profile was detected, and most markers were expressed uniformly within the population, the cells used in this thesis are analogous to other reports of hMSCs.

In contrast to most markers, the expression of the STRO-1 antigen was variable within the population, with a subset of cell expressing it to a high level; this is similar to previous findings with freshly isolated (uncultured) hMSCs (Gronthos *et al* 2003a), meaning that it could be an inherent characteristic of the hMSC population. A further explanation for this

variability could be that the cells were analysed at passage 4 (representing around 19 days in culture), and it has been shown that hMSCs begin to lose STRO-1 expression after two weeks in culture (Simmons and Torok-Storb 1991). As the hMSC population resides in the STRO-1^{BRIGHT} bright fraction, variable STRO-1 expression could reflect a heterogeneous population, possibly analogous to the cellular hierarchies proposed for stem cell populations (Owens 1988, Weismann *et al* 2001). It is currently difficult to overcome this issue without employing technically advanced methods of cell sorting that were beyond the remit of this study, and even these techniques do not completely remove the heterogeneity. However, the cell population could differentiate with high efficiency into three mesenchymal lineages, and close to 60% of cells were capable of undergoing adipogenesis, so I consider it unlikely that committed progenitors constitute the majority of the population.

The robust multilineage differentiation ability of heterogeneous hMSC populations was also a factor in deciding not to use clonal populations to investigate hMSC adipogenesis. Further to this, isolating clonal populations is problematic, as the low culturing densities required to generate hMSC clones is deleterious to their growth. Clonal populations could have been derived via plating the mononuclear cell fraction directly after isolation from bone marrow into 96-well plates at a density that would be likely to result in a single hMSC colony per well. However, it would have been difficult to obtain large enough cell numbers at a low passage for use in subsequent microarray studies, so heterogeneous populations were used in future experiments in this thesis.

Investigation of the hMSC growth rate revealed variations between hMSCs from different donors (Fig. 3.4). This has previously been reported (Bruder *et al* 1997, Phinney *et al* 1999), so is not unusual, but is an important factor that must be taken into account when designing future experiments. Furthermore, work described here showed that hMSCs from most donors have an extensive proliferation capacity *in vitro*, but that adipogenic differentiation was lost as the passage number increased (Fig. 3.5), an effect that has previously been noted (Conget and Minguell 1999) and reveals that low passage hMSCs should be used for differentiation studies. Taken together, these observations again

highlight the similarity between the hMSCs used in this thesis and the published properties of this cell type.

Supplementation of bFGF to the growth media had a mitogenic effect on hMSCs similar to previous reports (Fig 3.6, Tsutsumi *et al* 2001), thus allowing sufficient numbers of hMSCs to be obtained after few passages for use in subsequent experiments. This supplement has not previously been used to expand hMSCs for microarray gene expression studies of adipogenesis, so it was important to determine the effects on differentiation. It was found here that bFGF supplementation during expansion did not seem to bias differentiation towards a certain lineage (either adipogenesis or osteogenesis, Fig. 3.7), which could potentially cause misleading results in gene expression studies of differentiation, and was thus adopted as the hMSC expansion method in subsequent work.

The classic method of inducing adipogenesis in hMSCs involves a modification to the protocol for preadipocyte cell lines (using insulin, IBMX and dexamethasone), specifically the addition of a putative PPARγ ligand (indomethacin). It was found that indomethacin does indeed contribute an important role to hMSC adipogenesis, especially in combination with dexamethasone as these two substances induced differentiation to a similar, but not equal, extent (by qualitative assessment) to the combination of all four reagents (Fig. 3.8). It was found that the four reagents together induced the most cells to differentiate, as well as producing many cells containing multiple lipid droplets. The morphology of differentiated cells varied widely between conditions, and although the combination of insulin, indomethacin and dexamethasone gave rise to adipocytes containing the most lipid, relatively few cells actually differentiated. The results suggest that each of the four individual reagents plays an important role in inducing hMSC adipogenesis, but their combination produces a synergistic effect. In other work, alternative PPARγ ligands, such as troglitazone (Neubauer *et al* 2004), or glucocorticoids like hydrocortisone (Gimble *et al* 1995) have been employed in the hMSC adipogenic cocktail. However, early studies of hMSC adipogenesis do not seem to have addressed the specific abilities of each of these factors to induce adipogenesis, and may have adopted the use of such substances based on empirical studies of preadipocyte induction (1.6.1). Thus,

the results described here present novel information regarding the effect of the different inducers on hMSC adipogenesis.

When PPARγ was over-expressed in hMSCs and LiSa-2 cells, it was found that differentiation occurred without the need for chemical inducers, albeit to a lower extent than in chemically induced cells (Fig. 3.9). PPARγ2 induced the most extensive differentiation, which is in line with it's adipose-specific expression and higher ligand-independent activation capacity than PPARγ1 (Mueller *et al* 2002). PPARγ-induced adipogenesis was more efficient in LiSa2s than in hMSCs as a higher proportion of cells contained lipid droplets, which could indicate that LiSa2s are more committed to undergo adipogenesis than hMSCs. However, less lipid accumulated per cell in the LiSa2 line than in hMSCs, but this may be expected of a liposarcoma cell line as liposarcomas only partially resemble adipose tissue (Helman and Meltzer 2003). A comparison of PPARγ2 over-expression in hMSCs in this study and NIH-3T3 fibroblasts in previous work (Tontonoz *et al* 1994b) shows that fibroblasts require more stimuli (i.e. PPARγ over-expression plus dexamethasone and a PPAR ligand) than hMSCs to differentiate. This reflects the fact that fibroblasts are not normally capable of adipogenic differentiation, whereas hMSCs are able to differentiate into this lineage, and so represent a stage (albeit an early, uncommitted stage) along the pathway towards adipocytes. A clear observation in this experiment was that although all cells were expressing PPARγ, only a proportion differentiated. Explanations for this could be cell-to-cell differences in the level of PPARγ expression. Additionally, differentiation in PPARγ2-expressing cells occasionally occurred in clusters. This could indicate that differentiating cells produce an endogenous PPARγ ligand that enhances differentiation in neighbouring cells, which could possibly be 15-deoxy-$\Delta^{12,14}$-prostaglandin J$_2$ (Forman *et al* 1995). These effects of PPARγ over-expression on hMSC differentiation have not previously been described, and therefore provide novel information regarding the propensity of hMSCs to differentiate to adipocytes. Furthermore, the results demonstrate that retroviral introduction of exogenous genes into hMSCs can be used to study the effect on differentiation.

The time-course of adipogenesis was investigated in terms of cell morphology and lipid accumulation, as well as on the mRNA level, in order to define the timeframe of

adipogenic differentiation in hMSCs so that subsequent gene expression studies could focus on the early regulatory events. In the majority of bone marrows, rounded cells were observed one day after initiation of differentiation and small lipid droplets accumulated by day 2 (Fig. 3.10 and 3.11). An important observation was that hMSCs from some bone marrows produced lipid droplets by day 1, indicating that the rate of differentiation was faster. Taken together with the observation of variable expansion capacity between donors, this has important implications for gene expression studies (see Chapter 4.2). These observations suggest that by day 1-2 of differentiation the transcriptional program that leads to expression of genes characterising mature adipocytes, such as those involved in lipid synthesis or uptake, had been completed in some cells. Indeed, it was found by RT-PCR that PPARγ and C/EBPα (two transcriptional regulators of late adipogenesis) were upregulated as early as the 1 day timepoint, and their targets aP2 and adipsin, which are markers of mature adipocytes, were upregulated by day 2 (Fig. 3.12).

Where the initial observation of lipid accumulation during hMSC adipogenesis has been reported in other work, it has varied from day 3 to day 5 of differentiation (Nakamura *et al* 2003, Hung *et al* 2004a). Furthermore, upregulation of PPARγ was not observed at day 1 in hMSCs (Nakamura *et al* 2003, Sekiya *et al* 2004), but was found to be upregulated at day 3 compared to day 0 in another study (Hung *et al* 2004a). This indicates that hMSC adipogenesis described in this thesis is more rapid than in other reports. A definitive explanation for this is not forthcoming; as bFGF increases hMSC proliferation rate it could be hypothesised that it would also increase their ability to differentiate, but this was not observed to be the case in the work described here (Fig. 3.7). Possible reasons for rapidity of adipogenesis in this work compared with other reports are: variations in the source and passage number of the hMSCs used - commercially obtained hMSCs such as those used by Nakamura *et al* (2003) are not as well defined in these terms. Additionally, the fact that adipogenesis was performed on subconfluent hMSCs may have hindered the rate of differentiation in the study by Sekiya and co-workers (Sekiya *et al* 2004). Other explanations for the differences are that the specific growth media and supplement used in this work favour rapid differentiation, or that the cells were maintained at optimal conditions (for example, passaged before reaching confluence), which may not have been achieved in other work. Nevertheless, the differences in differentiation rate will not affect

the validity of the work in this thesis: a careful analysis was performed here, which ensured that subsequent microarray experiments would detect previously uncharacterised events and also demonstrated the fact that comparisons can still be made between the transcriptional events identified in different studies.

From day 2 until day 7 of differentiation, both the number of cells containing lipid and the amount of lipid per cell increased. This might suggest that the process of adipogenesis is not entirely synchronous. However, these observations are consistent with other microarray studies (Nakamura et al 2003, Sekiya et al 2004) and are an inevitable consequence of the system used; for example, although cells are left for at least one day post-confluence before differentiation is induced, they may still have undergone growth arrest to different degrees. In spite of this, gene expression changes were observed using RT-PCR (Fig 3.12), which would not be possible if differentiation was asynchronous to a large degree. Furthermore, numerous other studies of hMSC differentiation have identified clear gene expression patterns during adipogenic (Nakamura et al 2003, Sekiya et al 2004, Yloslato et al 2006) as well as osteogenic (MAPCs (1.4.2.3), Qi et al 2003) and chondrogenic differentiation (Sekiya et al 2002). Thus, it can be concluded that in vitro hMSC differentiation is a robust system for studying transcriptional regulation.

The observation that little increase in the number of cells that contain lipid occurred between 7-14 days suggests that no further cells were differentiating during this time, but rather that the effect of terminal differentiation, i.e. lipid accumulation, was propagated. Thus, the 7 day timepoint is likely to represent an accurate endpoint, in terms of the transcriptional differentiation program. At this point, near to 60% of cells had accumulated lipid; this is comparable to other studies of hMSC adipogenesis (Nakamura et al 2003, Sekiya et al 2004), but is less than 3T3L1s where as many as 90% of cells will differentiate (Soukas et al 2001). This is however likely to reflect the more committed nature of the 3T3L1 line.

A further observation from the RT-PCR analysis was that LIFR, a marker of preadipocytes and their early differentiation, was upregulated at 5h and then expression decreased again at 2d. Taken together with upregulation at day 1 of later regulators of

144

adipogenesis, this shows that the early events, which are the focus of this work, occur within the first day of differentiation, and even at this stage preadipocyte markers are upregulated. As a result of these observations, the initial microarray study of gene expression was focused on the 9h timepoint of differentiation, in order to discover whether further preadipocyte genes, and also markers of the earliest phase of gene expression from a microarray study of hMSC adipogenesis (Nakamura *et al* 2003), were expressed at this stage (Fig. 3.13). It was found that all genes analysed in both this study and the Nakamura study showed comparable patterns of regulation, in terms of up or downregulation, as much as could be concluded from this study as only 3 timepoints were analysed with only one array per timepoint. Also, the phases of gene expression were similar; genes that were upregulated late (7-14d) in the Nakamura study were also upregulated late (7d) here, and genes upregulated at the earliest timepoint (day 1) were upregulated at 9h here. An explanation for these findings, as already discussed, could be that differentiation occurs more rapidly in the system used here so the "day 1" genes are upregulated earlier. However, it is also possible that these genes were also upregulated during the first day of differentiation in the Nakamura study, but it was not detected due to the timepoints employed. Nevertheless, the combined outcome from these experiments was the description of the hMSC adipogenesis time-course in this system, from which it was revealed that gene expression studies focusing on the first 9h of differentiation would be most likely to detect previously undescribed gene expression, which may represent important regulatory steps in early adipogenesis.

CHAPTER 4. TRANSCRIPTIONAL PROFILING AND GLOBAL ANALYSIS OF EARLY hMSC ADIPOGENESIS.

4.1 Introduction and aims

The molecular pathways controlling adipogenesis have been studied in great detail using murine preadipocyte cell lines such as 3T3-L1 (1.6.1). This work has led to the identification of factors such as PPARγ (Tontonoz *et al* 1994a, 1994b), which is essential in co-ordinating the expression of mature adipocyte genes (1.6.2). As these cells are committed to the adipocyte lineage, genes such as PPARγ could be considered markers of the later stages of adipogenesis. Adipogenic differentiation of hMSCs would involve additional regulatory stages early during their differentiation, which may control processes such as lineage commitment (Rosen and Spiegelman 2000). Some research has aimed to investigate the molecular control of hMSC adipogenesis, occasionally using microarrays (1.7.2 and discussed below). However, previous studies have not investigated earlier than one day after induction of hMSC adipogenesis (Nakamura *et al* 2003, Sekiya *et al* 2004), and thus little is known of the molecular events occurring at this early stage. Identifying genes and pathways that are required at this point of differentiation could extend our knowledge of the early mechanisms of stem cell differentiation.

Oligonucleotide microarrays are used to investigate gene expression via the hybridisation of fluorescently labelled target molecules, synthesised from RNA extracted from a biological sample, to a microarray containing multiple oligonucleotide probesets, each representing different genes (Knudsen 2004). The fluorescent signal from a target hybridised to a probe is a reflection of the abundance of the specific RNA in the original sample, so can be used as a measure of gene expression (1.7.1.1). The development of microarray technology combined with the sequencing of the human genome has provided an opportunity to investigate gene expression on a genome-wide scale. Affymetrix have recently developed the U133 Plus 2.0 GeneChip, which contains around 54,000 probesets that represent over 47,000 transcripts, including all known human genes plus hypothetical proteins and ESTs (Affymetrix 2003). This array therefore allows comprehensive genome-wide analysis of expression changes, and it's use in this study would provide the first insight into the global regulation of the early stages of hMSC adipogenesis.

Sophisticated data processing and analytical techniques have been developed to permit the identification of biological trends from large datasets produced by microarray gene expression analysis (1.7.1.2). The outcome is an expression value for each gene in each sample assessed, following which significant gene expression changes between samples can be calculated. A point to note regarding this technique is that it yields only a comparative analysis of gene expression levels between two or more different conditions. Thus, it is possible to tell whether a gene is expressed to a higher or lower level in sample A than in sample B. However, it does not demonstrate the absolute expression level of the gene in either sample, and it may be the case that some genes are expressed to a high level in sample A, but because their expression did not change significantly in sample B they are excluded from further analysis. Nevertheless, in the context of this differentiation time-course study, analysing significant gene expression changes should identify biological trends in the expression data, as has previously been accomplished (Hackl *et al* 2005, Yloslato *et al* 2006).

A major advantage of microarray experimentation is that as the expression of large gene sets is determined simultaneously, the role of entire pathways or cellular processes during adipogenesis can be assessed. Gene Ontology (GO) is a systematic nomenclature that can be used to identify the molecular processes that are represented by a list of genes (Ashburner *et al* 2000). Three categories are used which aim to describe how gene products behave in a cellular context; these are molecular function, biological process and cellular component. Each GO category contains numerous terms arranged in a hierarchical structure that describe different aspects of the category, and gene products can be described by one or more term from each of these categories. Molecular function describes activities at the molecular level, such as "catalytic activity" or "DNA binding". A biological process term describes a series of events which are accomplished by more than one molecular function, an example being "signal transduction". Finally, the cellular component category describes intracellular location, for example "nucleus".

Through the categorisation of genes according to their expression pattern across the differentiation time-course, combined with tools that annotate these genes using methods such as GO, the specific stage during differentiation at which different processes act can

be identified. Many automated data analysis tools exist, and not all of them utilise GO annotation; signalling pathways, gene families and metabolic processes (amongst other things) can all be analysed via this method to give an integrated view of a specific process (Zeeberg *et al* 2003, Zhang *et al* 2005, Mi *et al* 2005, Ingenuity® Systems, www.ingenuity.com). It is also possible to attach a statistical significance to these findings, thus allowing robust identification of potential areas for future research.

Although microarrays have previously been used to analyse the transcriptional events of 3T3L1 and hMSC adipogenesis, the results of these studies have left gaps in the knowledge of this process for a number of reasons. Firstly, differentiation has only been assessed from the first day of hMSC adipogenesis onwards, so much of the earliest stages of regulation are likely to have been missed by these studies. Secondly, the largest arrays that were used to study hMSC differentiation analysed 12000 genes (Sekiya *et al* 2004, Yloslato *et al* 2006), so a great deal more information could be obtained by using the latest Affymetrix arrays covering all known human genes. Finally, early microarray studies of adipogenesis did not include global annotation analysis, probably due to the small size of the microarrays used combined with the fact that automated tools to annotate large datasets have only become available in recent years. The number of studies which now contain such analyses is increasing (Hackl *et al* 2005, Tan *et al* 2006, Yloslato *et al* 2006), but this has not yet been used to it's full potential to characterise hMSC adipogenesis.

Aims

The major aims of this chapter are to:

- Obtain gene expression profiles of the early stages of hMSC adipogenesis using Affymetrix oligonucleotide microarrays
- Confirm the accuracy and reproducibility of this data, and identify and validate patterns of gene expression that occur during differentiation
- Characterise, on both global and gene-be-gene levels, the molecular processes and functional categories of genes that may control the early stages of adipogenesis
- Finally, develop selection criteria based on the known function of genes and use this to identify potential novel candidates for the molecular control of adipogenesis.

4.2 Performance of microarrays during early hMSC adipogenesis

The aims of this experimental section were to perform a differentiation time-course and generate transcriptional profiles of adipogenesis at the early stages. Secondly, through data processing and statistical analysis methods, an initial biological insight into adipogenic regulation would be gained.

4.2.1 Microarray strategy

In Chapter 3, it was found that lipid accumulation was visible at day 2 of differentiation (Fig 3.10), and markers of late adipogenesis were upregulated at day 1 (Fig 3.12). Additionally, an exploratory microarray study identified some preadipocyte markers as upregulated at 9h during differentiation (Fig 3.13). It was therefore decided that microarray experiments should focus on the initial 9h of differentiation, in order to capture the earliest transcriptional events. Furthermore, some variation in differentiation rates between bone marrows had been observed (Fig 3.11). This could mean that, if a specific timepoint was assayed in hMSCs from several bone marrows, slightly different stages of differentiation may be captured. Inaccurate microarray data may result from this, as biological variation (the genuine changes in gene expression that regulate adipogenesis) could be masked by the patient-to-patient variation. It was therefore decided to perform the microarray studies using a single bone marrow, and subsequently confirm the observations using hMSCs from other bone marrows to ensure that gene expression changes were not patient-specific. Therefore, frozen hMSCs (at passage 2) isolated from a single bone marrow (a 31 year-old male donor) were thawed, expanded (2.2.1), allowed to reach confluence, and then differentiation to adipocytes was initiated (2.2.2.1). Figure 4.1A shows the experimental design, which was set up in three independent experiments so that replicate arrays could be performed at each chosen timepoint.

Morphological changes during differentiation were followed by Oil Red O staining and phase contrast microscopy (2.2.3.1). During the first 3h of differentiation no morphological changes were observed. At 9h however, some cells were already changing from a fibroblastic morphology into a more rounded shape, characteristic of adipocytes (Fig. 4.1B). As in previous experiments, lipid accumulation was observed at day 2 and
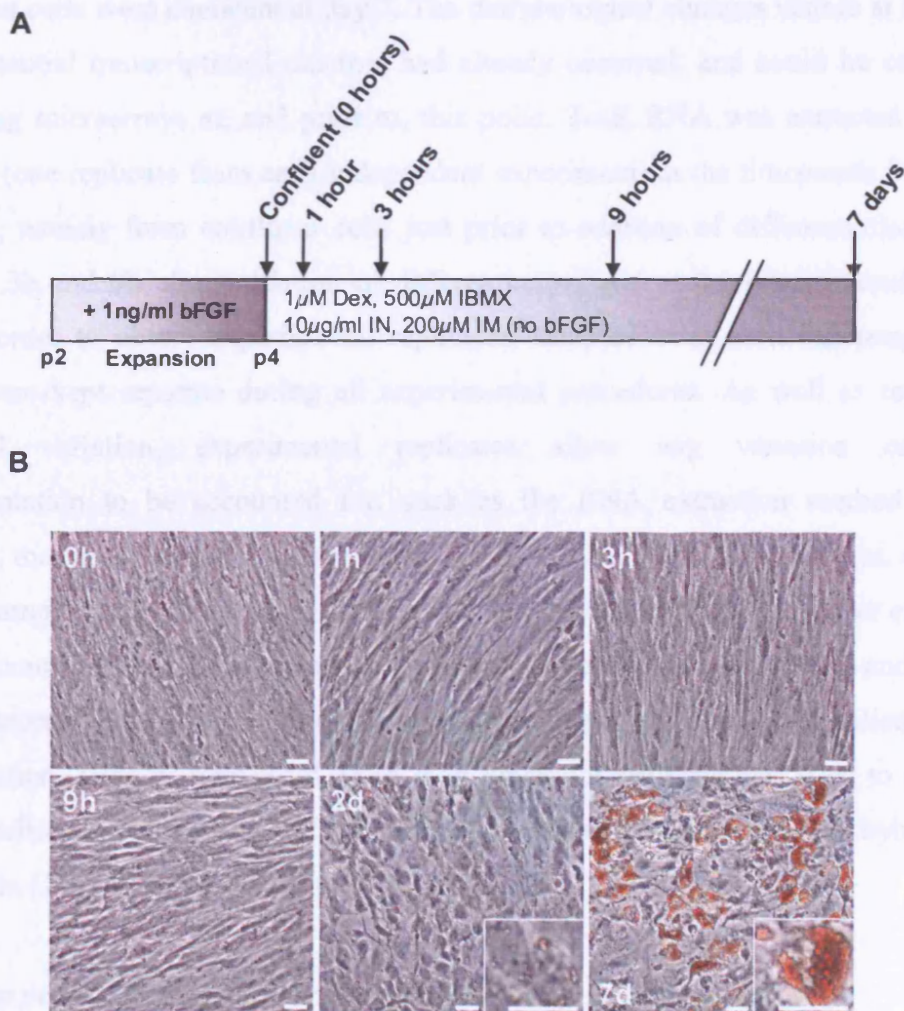
Figure 4.1. Time-course of hMSC adipogenesis for microarray analysis. A, Schematic of differentiation program and microarray timepoints. Cells were expanded to passage 4 (2.2.1), then seeded at approximately 4×10^3 cells/cm^2 and grown to confluence in media containing 1ng/ml bFGF. Differentiation was induced via removal of bFGF and addition of the differentiation cocktail (2.2.2.1). RNA was extracted (2.6.1) at the timepoints indicated from three separate differentiations using hMSCs from a single bone marrow (31 yr old male). Dex, dexamethasone; IBMX, isobutylmethylxanthine; IN, insulin; IM, indomethacin; B, Morphological characteristics of differentiating hMSCs. At the timepoints indicated hMSCs were fixed, stained with Oil Red O (2.2.3.1) and visualised via phase contrast microscopy. Scale bars represent 100μm.

lipid-laden cells were abundant at day 7. The morphological changes visible at 9h suggest that substantial transcriptional changes had already occurred, and could be captured by performing microarrays at, and prior to, this point. Total RNA was extracted (2.6.1) in triplicate (one replicate from each independent experiment) at the timepoints indicated in Fig 4.1A, namely from confluent cells just prior to addition of differentiation reagents (0h), 1h, 3h and 9h after initiation of differentiation, and at the experimental endpoint (7d). In order to obtain experimental replicates, samples from each independent time-course were kept separate during all experimental procedures. As well as representing biological variation, experimental replicates allow any variation caused by experimentation to be accounted for, such as the RNA extraction method or probe synthesis, meaning that the resulting data is more robust. Technical replicates, created by pooling samples at a certain stage of experimentation (e.g. pooling RNA after extraction), do not account for sources of technical variation that occurred after pooling, and thus give a less accurate estimation of biological variation than experimental replicates. After quantification, $10\mu g$ of total RNA from each timepoint replicate was used to synthesise biotin-labelled cRNA probes (2.7.1). Probes were then fragmented and hybridised to Affymetrix U133 Plus 2.0 GeneChip Arrays (2.7.2).

4.2.2 Data processing and statistical analysis reveal biological trends

Expression signals produced from array scanning were pre-processed using the robust multiarray analysis (RMA) method of Irizarry et al (2003). This technique was chosen as a comparison of this method and two other Affymetrix array processing methods, the MBEI method of Li and Wong (2001) and MAS5.0 from Affymetrix (2001), found it to perform to a higher degree of accuracy (Irizarry et al 2003b). Furthermore, GC-RMA, a modified version of RMA (Wu and Irizarry 2004), was not a standard data pre-processing technique at the time of this analysis. RMA produces $\log_2$-scale expression values for each probeset, in the range 4 (low expression) to 16 (high expression) (2.7.3). These expression values were used to identify significant gene expression changes between successive timepoints (0h-1h, 1h-3h, 3h-9h, 9h-7d, named "phases" of differentiation). This approach was used rather than comparing each timepoint to the 0h arrays as it was deemed a more biologically relevant strategy for analysing time-course data. The Linear Models for Microarray Analysis (LIMMA) method was employed for these calculations, which

152

combines robust estimation of variance with a standard t-test, and so gives a more accurate estimate of statistical significance associated with gene expression changes (Smyth 2004, 2.7.3). The False Discovery Rate method of Benjamini and Hochberg (FDR, Benjamini and Hochberg 1995) was used to address false positives within the resulting list of significantly changed genes. When multiple statistical tests are performed, some genes may be classified as significant by chance; these are termed false positives (2.7.3). For example, if 1000 genes are tested and a significance level of 0.05 is chosen, 50 false positives would be expected. FDR reduces the proportion of false positives in a list of genes classified as significant (according to a particular cut-off) by correcting probability p values to higher q values (2.7.3). The significance criteria of q<0.05 was adopted as it was decided that a more stringent cut-off, such as q<0.01, may miss small but important expression changes (which can often be the case for transcription factors). Furthermore, using probability values is more accurate than using a fold-change cut-offs, as variation between timepoint replicates is taken into account when statistical significance is calculated, but is not represented by fold-change values. Genes with significant changes in at least one of these phases were selected for further analysis (Appendix Tables 1 and 2).

The global expression profiles for each array were next used in average linkage clustering (2.7.3) to investigate the relationship between them. This relationship is represented by a dendrogram in which the distance between two arrays (measured by the dendrogram branch length) is an indication of the similarity between their expression profiles. The 0h and 1h replicates clustered together on the dendrogram, indicating that gene expression at these timepoints was similar (Fig. 4.2A). One 3h replicate was a slight outlier, but was most similar to the other 3h replicates, which in turn were most similar to the 0h and 1h arrays. Both the 9h replicates and 7d replicates formed two separate clusters, with the 7d arrays occupying an entirely separate branch of the dendrogram to the other arrays, indicating that between 3h-9h and 9h-7d, many expression changes had occurred that led to distinct profiles at these timepoints. The number of significant expression changes that occurred in each phase of differentiation supported these findings. Eighty-four probesets showed a significant expression change in the first hour of differentiation (Fig. 4.2B). This number increased progressively in each phase of differentiation, and culminated in more than 14000 significant expression changes occurring in the final phase, an unexpectedly
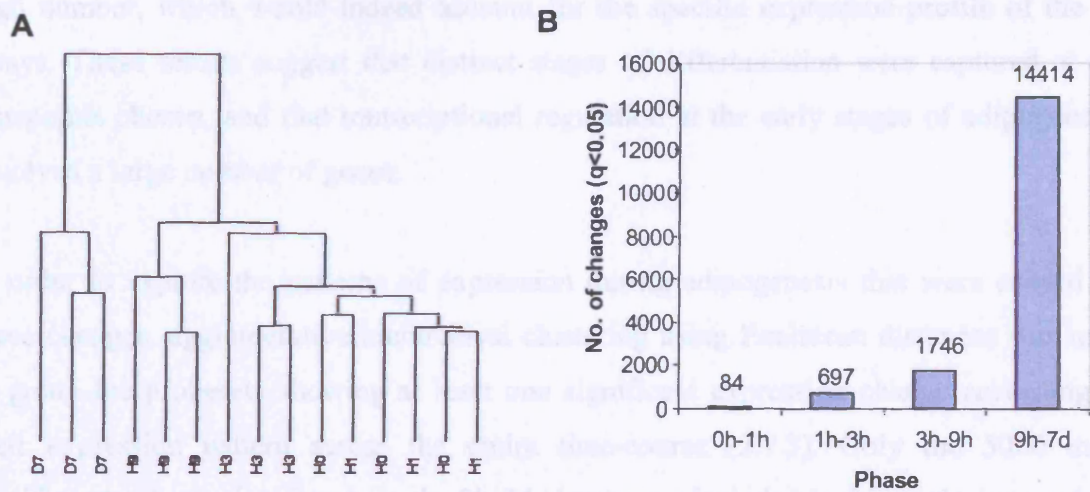
Figure 4.2. Expression changes that occur during adipogenesis result in distinct transcriptional profiles. A, Similarity of expression profiles between replicate timepoints. Global expression profiles from each microarray were used in average linkage clustering using Euclidean distances (2.7.3) to determine the relationship between them; B, Number of significant expression changes detected in microarray analysis. Normalised, $\log_2$-transformed expression values were used to calculate significant ($q<0.05$) expression changes for each probeset between successive timepoints, designated as phases of differentiation (2.7.3). Microarray data processing and production of the cluster dendrogram was performed in collaboration with Dr. S. Henderson.

high number, which would indeed account for the specific expression profile of the 7d arrays. These results suggest that distinct stages of differentiation were captured at the timepoints chosen, and that transcriptional regulation at the early stages of adipogenesis involved a large number of genes.

In order to explore the patterns of expression during adipogenesis that were caused by these changes, agglomerative hierarchical clustering using Euclidean distances was used to group the probesets showing at least one significant expression change according to their expression pattern across the entire time-course (2.7.3). Only the 5000 most significantly changed probesets in the 9h-7d phase were included in this analysis to reduce complexity; this phase was not the main focus for identification of novel early regulators, and this number was deemed sufficient to illustrate the nature of later transcriptional events. Based on the significance cut-off chosen for this study (q<0.05), 6319 probesets that were transcriptionally regulated during adipogenesis were analysed in this manner. This represents 12% of the total probesets on the microarray, again highlighting the complexity of the early transcriptional control of adipogenesis, as well as the ability of microarray profiling to identify such detailed regulation. This number corresponds to 4276 known genes, which reflects the fact that the probesets may also represent ESTs and hypothetical proteins, and also that several probesets may represent a single gene.

The expression patterns of each gene[1] are illustrated graphically on a heatmap (Eisen *et al* 1998), and a cluster dendrogram is produced to illustrate similarity in expression patterns between genes (Fig. 4.3A, Appendix Figure 1, Folder 2 and Table 3). Within the dendrogram, certain nodes (the point at which branches meet) define a gene group in which all members have a characteristic expression profile. The heatmap was broken down into 17 separate nodes (Fig. 4.3A, coloured bars; B, graphs) that represent distinct expression patterns, and cover 99% (6244) of the clustered genes (Fig. 4.3A and B).

[1]From this point onwards, the word gene will be used in place of probeset to describe an entity with a significant expression change during adipogenesis. This is for ease of explanation, but the entities could still be hypothetical proteins and ESTs, and multiple probesets could represent a single gene.
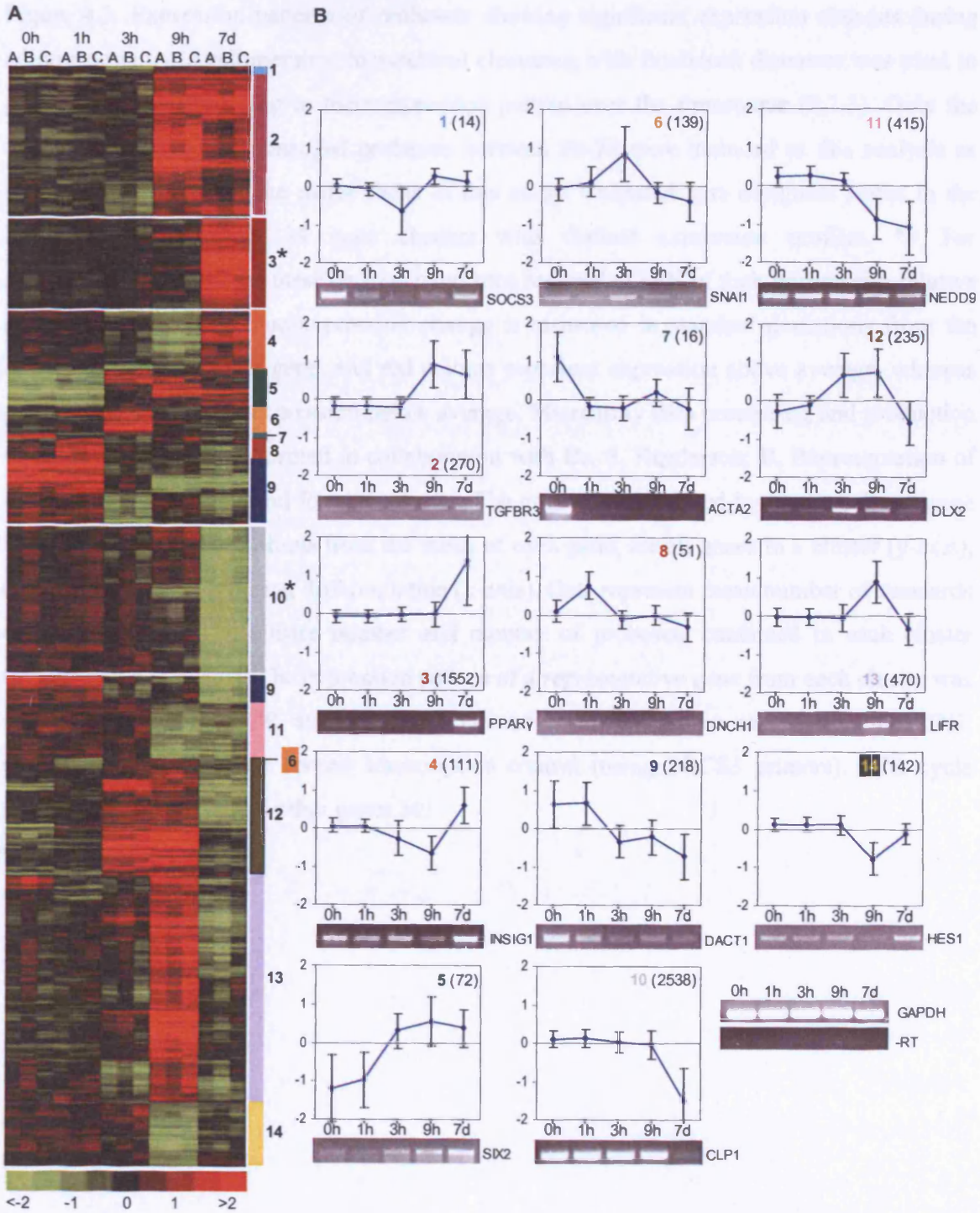
Figure 4.3. Continued overleaf.

Figure 4.3. Expression patterns of probesets showing significant expression changes during adipogenesis. A, Agglomerative hierarchical clustering with Euclidean distances was used to group probesets according to their expression pattern over the timecourse (2.7.3). Only the 5000 most significantly changed probesets between 9h-7d were included in this analysis as such changes were not the major focus of this study. Coloured bars designate nodes in the heatmap that represent 14 gene clusters with distinct expression profiles. *= For representational purposes these clusters have been reduced to 33% of their normal size relative to the other clusters. Gene expression change is measured in standard deviations from the mean expression of each gene, and red colours represent expression above average, whereas green colours represent expression below average. Microarray data processing and production of the heatmap were performed in collaboration with Dr. S. Henderson; B, Representation of the overall expression trend for each cluster. The graphs were created by plotting the average number of standard deviations from the mean of each gene, for all genes in a cluster (y-axis), and at each timepoint during differentiation (x-axis). Data represent mean number of standards deviations, ±S.D. The cluster number and number of probesets contained in each cluster (brackets) are indicated. The expression pattern of a representative gene from each cluster was confirmed using RT-PCR analysis (2.6.3.1), and is shown below each graph. GAPDH; internal control; –RT: no reverse transcription control (using SOCS3 primers). PCR cycle numbers: GAPDH: 25; all other genes 30.

In some cases, the expression patterns of two nodes were so similar that they were grouped together into "expression clusters" for further analysis. Thus, 11 out of the 14 expression clusters were represented by an individual node of the heatmap, and 3 of the clusters (6, 8 and 9) each contained 2 nodes that were grouped together. The expression patterns revealed that adipogenic regulation occurs in either transient waves of up- or downregulation (for example clusters 1 and 6), or as more sustained changes in gene expression (as in clusters 5 and 9). The number of genes in each cluster varied over two orders of magnitude and was likely to be caused by the expression pattern defining the cluster, as clusters representing early changes consistently contained fewer genes than those representing later gene expression changes. However, it seems that sustained gene expression changes, rather than transient changes, are the predominant form of regulation during adipogenesis, as clusters representing this type of pattern were on the whole the largest. The detection of temporal gene expression patterns during adipogenesis illustrates the stage at which a gene acts, as well as giving an indication as to how it might interact in a genetic network, for example it may function together with other genes that are regulated in a similar manner. For this reason, most future data analysis was based on the clusters described above.

4.3 Validation of microarray data analysis and differentiation system

4.3.1 Microarray data accurately reflect changes in mRNA levels

Microarray experiments inherently contain numerous data processing steps following any experimental work. It is therefore vital to ensure that resulting data is reproducible and accurately represents biological changes that occurred in the initial experiment. Therefore, the aim of these experiments and analyses was to verify the efficacy and reliability of both the experimental system and analysis methods used.

To assess the reproducibility of the microarray results, box-plots were produced (2.7.3) to summarise the distribution of expression values (in RMA units) of each timepoint replicate before normalisation (Fig. 4.4). Box-plots permit simple visualisation and efficient comparison of multiple datasets. The box plots revealed that the spread of

158
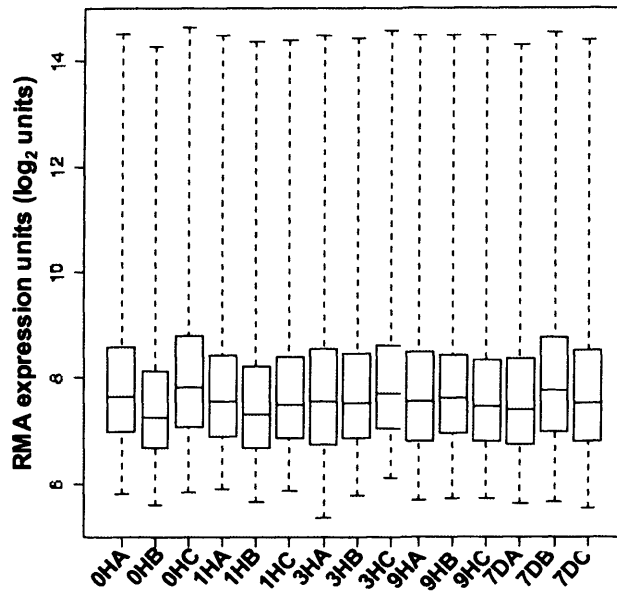
Figure 4.4. Reproducibility of the microarray data. Box-plots were created (2.7.3) to show the distribution of expression intensities for each array, after background correction and before normalisation. Median average expression values (the data point representing the 50th percentile) for each array are represented by the line within the box, and the upper and lower edges of the box represent the 75th and 25th percentile, respectively, of the expression value data for each array. Dashed bars indicate the maximum and minimum expression value on each array. Box-plots were produced in collaboration with Dr. S. Henderson.

expression values was similar between every array, as indicated by the similar size of each box (the inter-quartile range, representing the middle 50% of the data), and the fact the there was no large skew in any of the datasets (represented by the median line being close to the centre of the box). Furthermore, the range of expression values on each array was comparable. Thus, no array was an outlier compared to the other arrays, which could affect later calculations such as normalisation; by way of an example, outliers can be caused by factors such as poor washing after staining causing inaccurately high expression values over an entire array. Pearson correlation coefficients between all replicate arrays at each timepoint were also calculated (2.7.3). The average coefficient at each timepoint ranged from 0.984-0.993, showing that the data are reproducible.

In order to determine whether the microarray data accurately reflected changes in mRNA level, a representative gene from each cluster was next chosen and it's expression during differentiation analysed using RT-PCR (2.6.3.1). In every case, the expression profile of each gene closely matched the expression pattern from the microarray data (Fig 4.3B). For example, DLX2 was found in cluster 12, and hence showed upregulation at 3h followed by downregulation at 7d; RT-PCR analysis also detected similar changes in mRNA levels. This demonstrates that the data collection and processing methods employed produce results that accurately portray the changes in mRNA levels during adipogenesis.

As the microarray study used hMSCs isolated from a single bone marrow, the conservation of gene expression changes in hMSCs from other donors was investigated. Time-course differentiations were performed using hMSCs from an additional three bone marrows and the expression of a selection of genes was analysed at identical timepoints (Fig. 4.5). In the majority of cases the expression of each gene followed a similar pattern in all four bone marrows, although some variation was observed. For example, the SOCS3 trend of downregulation at 3h was conserved, but in some bone marrows further changes including downregulation at 7d were seen for this gene. These results suggest that the same transcriptional events control adipogenesis in different hMSC samples, although the timing and magnitude of change may vary to some extent. Additionally, they further confirm that the system used and microarray data obtained in this investigation can be used to gain an accurate insight into hMSC adipogenesis.
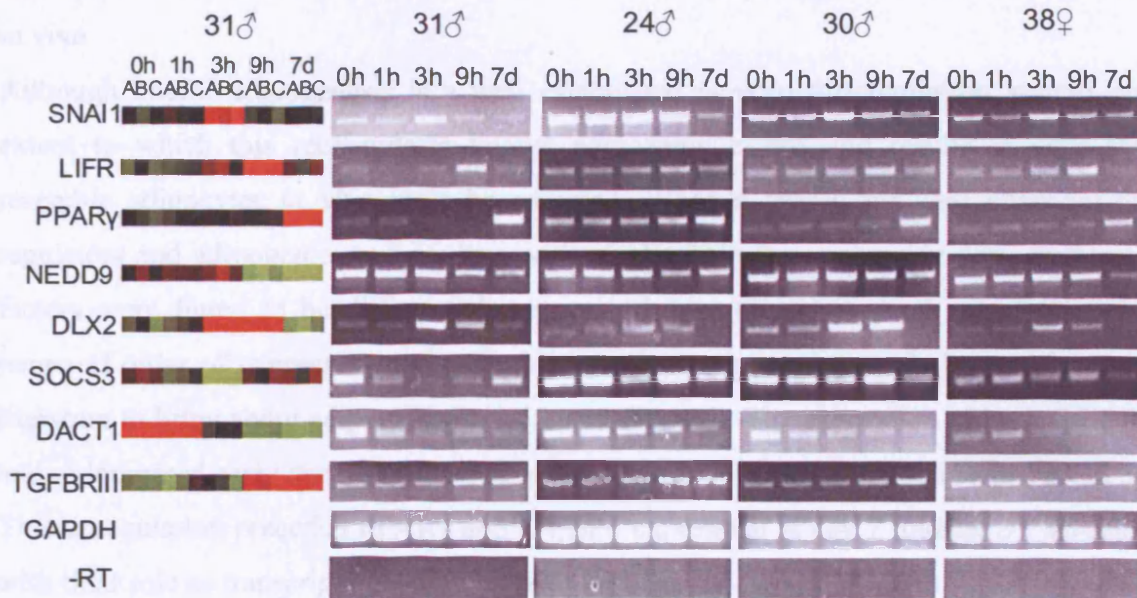
Figure 4.5. Gene expression patterns identified accurately represent regulation of hMSC adipogenesis. MSCs from three bone marrows (24♂, 30♂, 38♀) were differentiated to adipocytes (2.2.2.1) and RNA was extracted (2.6.1) at identical timepoints to those used in microarray studies (31♂). Expression of a selection of genes used to validate the microarray data were again analysed using RT-PCR (2.6.3.1). Heatmap bars represent expression of the gene in the microarray data; where multiple probesets were present for a single gene, a representative probeset was chosen. –RT: no reverse transcription control (using SOCS3 primers). PCR cycle numbers: GAPDH: 25; all other genes 30.

4.3.2 hMSC differentiation recapitulates known adipogenic regulation and adipogenesis in vivo

Although hMSC adipogenesis is a well established *in vitro* differentiation system, the extent to which this recapitulates known adipogenic events and results in cells that resemble adipocytes *in vivo* must be assessed. When a search for well characterised regulators and adipogenic markers was performed within the microarray data, numerous factors were found to be differentially expressed over the time-course and followed a temporal order of expression that resembled the transcriptional cascade known from the literature to bring about adipogenesis (1.6.2 and Fig. 4.6). The expression of both C/EBPβ and -δ increased early during differentiation, at 1h and 3h respectively (clusters 12 and 5). This upregulation preceded PPARγ and C/EBPα expression at day 7 (cluster 3), agreeing with their role as transcriptional activators of these genes (Wu *et al* 1996). LIFR is known to be expressed in preadipocytes and promotes adipogenesis partly through activation of C/EBPβ and -δ (Aubert *et al* 1999); LIFR was upregulated transiently at 9h (cluster 4). Transcriptional activators of PPARγ such as SREBP1 and STAT5A and -5B were upregulated alongside PPARγ at day 7 (cluster 3), whereas the PPARγ repressor FOXC2 was downregulated at this timepoint (cluster 6). A final observation was the upregulation also at 7d of adipocyte markers such as aP2 and fatty acid synthase (FASN), which correlates with their regulation by PPARγ and C/EBPα and also the observation of a mature adipocyte phenotype at this stage. The identification of many known adipogenic regulators as well as the temporal order of their expression in this study validates this system as a reliable model of adipogenic differentiation.

The transcriptional resemblance of 7 day-differentiated adipocytes to adipose tissue *in vivo* was investigated by determining the expression at this timepoint of genes and pathways known to be important in adipose tissue (Ailhaud 2001, Urs *et al* 2004, Viguerie *et al* 2005). Cluster 3 (genes upregulated at day 7) contained many genes known to be expressed in mature adipocytes (Table 4.1). These included secreted proteins such as leptin, which is not expressed at high levels by 3T3-L1 cells (Rosen and Spiegelman 2000). This could indicate that hMSC adipogenesis is a more robust differentiation model. Other genes identified were transport and storage molecules such as aP2 and perilipin, along with many genes involved in the metabolism of lipid and carbohydrates. An
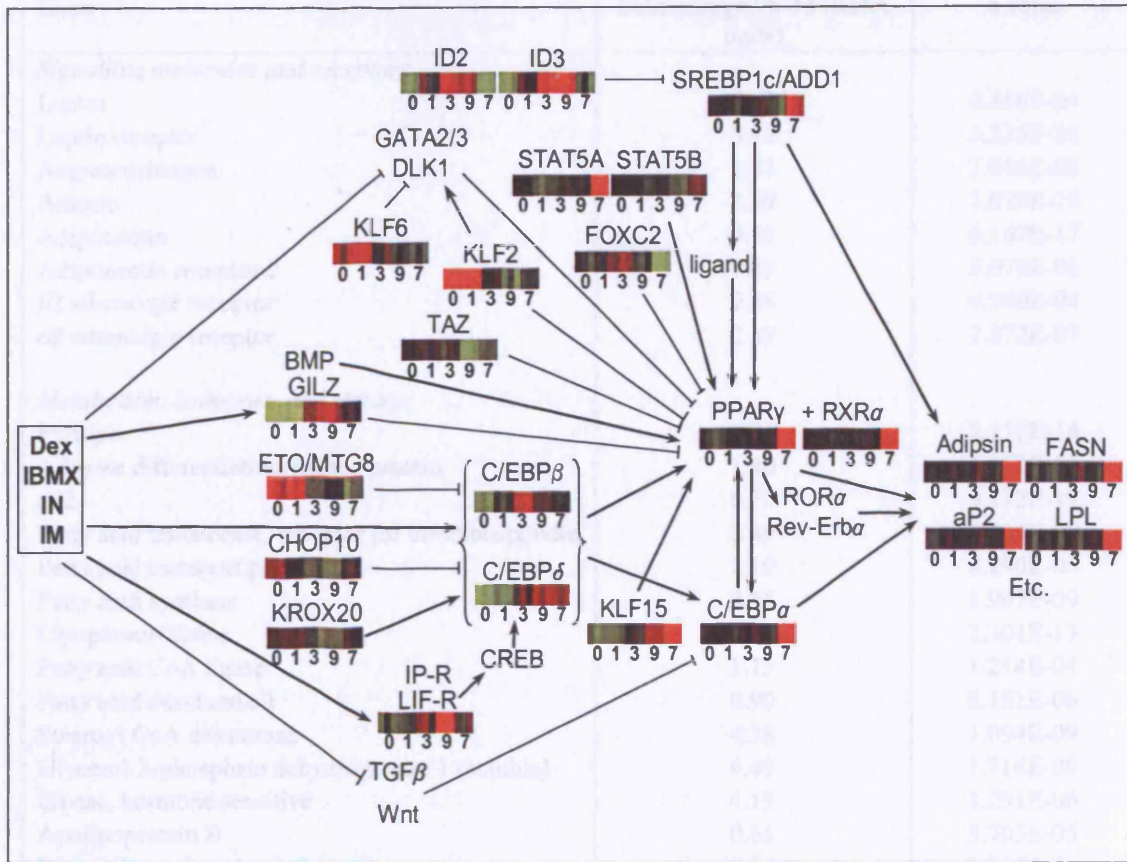
Figure 4.6. Expression of known adipogenic regulators and adipocyte markers during hMSC adipogenesis follows an ordered temporal pattern. Each heatmap bar (created by resizing sections of the heatmap in Figure 4.3, thus each replicate array is represented by a strip of colour) indicates the expression of the specific gene across the timecourse. 0, 0h; 1, 1h; 3, 3h; 9, 9h; 7, 7d. Where more than one probeset was present for the same gene, a representative probeset was included. Other probesets can be viewed in Appendix Figure 1. Genes with no heatmap bars did not exhibit a significant expression change during adipogenesis.

Table 4.1 (overleaf). Genes/pathways characterising adipocytes *in vivo* that were expressed/enriched in 7d-differentiated hMSCs. All genes were identified in cluster 3 (genes showing permanent upregulation at day 7). The entire gene list from this cluster was used to identify KEGG pathways showing significant enrichment (p<0.01, see section 2.7.4.1 for enrichment calculations).

Gene	Fold change 9h-7d (RMA units)	q value
Signalling molecules and receptors		
Leptin	0.72	6.868E-04
Leptin receptor	3.42	3.325E-10
Angiotensinogen	1.43	7.056E-06
Adipsin	7.09	2.028E-14
Adiponectin	9.23	6.167E-17
Adiponectin receptor 2	0.83	5.078E-06
$\beta2$ adrenergic receptor	0.68	6.589E-04
$\alpha2$ adrenergic receptor	2.37	2.872E-07
Metabolism, transport and storage		
Perilipin	6.01	2.112E-14
Adipose differentiation-related protein	1.40	1.172E-07
aP2	6.71	9.172E-17
Fatty acid translocase, receptor for thrombospondin	2.49	2.405E-08
Fatty acid transport protein 1	1.19	8.290E-05
Fatty acid synthase	3.45	1.997E-09
Lipoprotein lipase	7.11	2.301E-13
Fatty acid CoA ligase	1.15	1.214E-04
Fatty acid desaturase 1	0.90	8.161E-06
Stearoyl CoA desaturase	4.28	1.094E-09
Glycerol-3-phosphate dehydrogenase 1 (soluble)	4.49	1.714E-09
Lipase, hormone sensitive	4.19	1.291E-06
Apolipoprotein B	0.65	5.705E-05
Diacylglcyerol acyltransferase 1	2.07	2.147E-05
Glycerophosphate acyltransferase	2.92	2.214E-09
Phospoenolpyruvate carboxykinase	7.46	2.741E-15
Phoshodiesterase 3B	3.27	7.894E-08
Acyl-Coenzyme A dehydrogenase, C-4 to C-12 straight chain	1.52	1.346E-04

KEGG Pathway	Number of genes in pathway upregulated 9h-7d	Enrichment p value
Oxidative phosphorylation	47	2.0E-29
Fatty acid metabolism	23	6.4E-17
Citrate cycle (TCA cycle)	18	9.3E-18
Pyruvate metabolism	18	2.7E-12
Insulin signaling pathway	15	7.9E-03
Glycolysis / Gluconeogenesis	12	1.3E-04
Glycerolipid metabolism	11	1.7E-04
Adipocytokine signaling pathway	11	1.4E-03
ATP synthesis	10	1.5E-05
Biosynthesis of steroids	9	1.7E-07
Fatty acid elongation in mitochondria	7	1.8E-07

automated annotation tool (see section 4.4.1) was next used to search for members of KEGG metabolic pathways in cluster 3, in order to identify metabolic pathways that were significantly enriched in this cluster compared to a reference list (all genes on the U133 Plus 2.0 array, 2.7.4.1). Many metabolic pathways defining adipocytes were enriched in the gene cluster upregulated at day 7, including lipid metabolism, energy generation and adipocyte signalling pathways (Table 4.1). These results demonstrate that differentiation of hMSCs with adipogenesis-inducing substances produces cells which, at day 7 of differentiation, have a transcriptional profile that strongly resembles mature adipocytes.

4.4 Identification of molecular processes and factors involved in early adipogenesis

4.4.1 Molecular processes characterising early adipogenesis revealed by Gene Ontology annotation

Global annotation of the microarray data using Gene Ontology (GO) terminology was performed in order to identify the molecular processes that act at specific stages of adipogenesis. The efficacy of a range of web-based annotation tools was assessed, including Panther (Mi *et al* 2005) and GOMiner (Zeeberg *et al* 2003). GO Tree Machine and WebGestalt (Zhang *et al* 2005) were found to be the most suitable for this study, as the possible data input formats included Affymetrix probeset IDs, and a reference list containing annotations of all genes on the U133 Plus 2.0 array was available, which was not the case for some other annotation tools. Furthermore, WebGestalt could annotate the microarray data in a number of ways, and uses a statistical test to apply a significance value to each annotation, allowing the most interesting biological trends to be identified. GO annotations were obtained for genes in each cluster using WebGestalt. On average 52% of input IDs had associated GO annotations; this is because the clusters contain hypothetical proteins and ESTs as well as known genes; also not all known genes have functional annotations. Significantly enriched GO terms in each cluster, compared with what would be expected from all the genes on the U133 Plus 2.0 array, were identified using a hypergeometric test with the significance cut-off of $p<0.05$ (2.7.4.1), apart from the largest clusters (3 and 10) where a higher cut-off of $p<0.01$ was used. This was

because the large size of these two clusters, combined with the fact that they were not the focus for studying the earliest stages of adipogenesis, and hence only the most significant trends were focused on. An additional criterion that the GO term must also contain at least 10% of genes in the cluster was added to decrease the probability of a significant enrichment occurring by chance. A list of all enriched GO terms can be found in Appendix Table 4, and Figure 4.7 summarises the biological process terms that passed both significance criteria. The GO "Biological Process" category was the focus for further analysis as it was decided that this would describe the transcriptional events of adipogenesis in the most informative manner.

Developmental genes are significantly enriched in several clusters; a transient upregulation of a subset of these genes is observed at 3h (cluster 6), whereas many developmental genes, more specifically those involved in muscle development in cluster 7, are permanently downregulated at various points during differentiation (clusters 7, 9 and 11). This may indicate a requirement for specific genes to control an early stage of differentiation, but an important step during adipogenesis is the downregulation of genes that may be involved in the development of other tissues.

Transcription plays a major role throughout adipogenesis; GO terms related to transcription constitute the majority of enriched terms in the clusters representing early regulation. Such genes are more frequently found to be upregulated (clusters 5, 6, 8, 12 and 13) than downregulated (clusters 9 and 14). This suggests that adipogenesis requires the upregulation of many transcription factors at specific timepoints to regulate progression of differentiation that were not expressed, or expressed at low levels, in undifferentiated hMSCs.

The response to stress process was enriched in cluster 2 (genes permanently upregulated at 9h). This suggests that some stress-inducing conditions are encountered during early adipogenesis, likely as a result of the application of the induction cocktail, but no clear pro-apoptotic response was detected using microarrays at the early stages of adipogenesis. Genes involved in cell death and apoptotic processes are permanently upregulated at 3h (cluster 5). However, half of these genes are also termed "negative regulation of

Figure 4.7 (continued). Gene Ontology enrichment across clusters 1–14.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Cluster	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Total Input IDs	14	270	1652	111	72	189	16	51	218	2538	415	235	470	112
Total known genes	6	120	525*	62	41	79	7	24	112	525*	229	129	229	85

(A row of per-cluster expression-profile sparklines appears here.)

Biological Process

1	2	3	4	5	6	7	8	9	10	11	12	13	14	GO Term
				19			24		43					GO:0007275 Development
					2									GO:0048513 Organ development
					2									GO:0007517 Muscle development
		9					5		59					GO:0007049 Cell cycle
							3							GO:0000279 M phase
							3							GO:0000087 M phase of mitotic cell cycle
							3							GO:0007067 Mitosis
							3							GO:0000278 Mitotic cell cycle
		8												GO:0000074 Regulation of progression through cell cycle
							7		82	38			21	GO:0016043 Cell organization and biogenesis
							6			23			12	GO:0006996 Organelle organization and biogenesis
							3							GO:0051276 Chromosome organization and biogenesis
							3							GO:0007001 Chromosome organization and biogenesis (sensu Eukaryota)
							3						9	GO:0007010 Cytoskeleton organization and biogenesis
				16	33			49				70		GO:0006139 Nucleobase, nucleoside, nucleotide and nucleic acid metabolism
				12	29		9	42		28	46		20	GO:0019219 Regulation of nucleobase, nucleoside, nucleotide and nucleic acid
				12	29		9	41		28	46		20	GO:0045449 Regulation of transcription
				11	29		8	40			45		20	GO:0006355 Regulation of transcription\, DNA-dependent
				13	29		9	42		29	47		20	GO:0006350 Transcription
				12	29		8	41			46		20	GO:0006351 Transcription, DNA-dependent
				7	8			14						GO:0006366 Transcription from RNA polymerase II promoter
39				20	41		14	59		31			31	GO:0050789 Regulation of biological process
14		10		7				19						GO:0048519 Negative regulation of biological process
				6										GO:0048518 Positive regulation of biological process
37				20	39		14	57		43	65	191	30	GO:0009887 Cellular process
13		8		7				18						GO:0050794 Regulation of cellular process
				5										GO:0048523 Negative regulation of cellular process
														GO:0048522 Positive regulation of cellular process
		391		35	59			91		101	171			GO:0050875 Cellular physiological process
				19	38		13	55		42	61		30	GO:0051244 Regulation of cellular physiological process
		7		7				17						GO:0051243 Negative regulation of cellular physiological process
		416												GO:0007582 Physiological process
				19	40		13	56			47		31	GO:0050791 Regulation of physiological process
		7		7				17						GO:0043118 Negative regulation of physiological process
				5										GO:0043119 Positive regulation of physiological process
70	322	41		47			76							GO:0008152 Metabolism
26				13	29		9	44		43	48			GO:0019222 Regulation of metabolism
67	300	38			46			75						GO:0044237 Cellular metabolism
				12	29		9	43		31	48		20	GO:0031323 Regulation of cellular metabolism
		36			45			75				121		GO:0044238 Primary metabolism
									189					GO:0043170 Macromolecule metabolism
									155					GO:0044260 Cellular macromolecule metabolism
									152					GO:0044267 Cellular protein metabolism
							20		94					GO:0006464 Protein modification
									154					GO:0019538 Protein metabolism
							30		131		58			GO:0043283 Biopolymer metabolism
							20		94					GO:0043412 Biopolymer modification
				6										GO:0006793 Phosphorus metabolism
				6										GO:0006796 Phosphate metabolism
				5										GO:0006468 Protein amino acid phosphorylation
		73												GO:0006629 Lipid metabolism
		59												GO:0044255 Cellular lipid metabolism
		82												GO:0009058 Biosynthesis
		65												GO:0044249 Cellular biosynthesis
		63												GO:0006082 Organic acid metabolism
		62												GO:0019752 Carboxylic acid metabolism
		80												GO:0006091 Generation of precursor metabolites and energy
									164	76				GO:0007154 Cell communication
									54					GO:0007155 Cell adhesion
										66				GO:0007165 Signal transduction
										27				GO:0007242 Intracellular signaling cascade
		132												GO:0051179 Localization
		132												GO:0051234 Establishment of localization
		121												GO:0006810 Transport
													14	GO:0051641 Cellular localization
													14	GO:0051649 Establishment of cellular localization
													14	GO:0046907 Intracellular transport
				5									9	GO:0009653 Morphogenesis
				6										GO:0008219 Cell death
				6										GO:0012501 Programmed cell death
				6										GO:0006915 Apoptosis
				5										GO:0042981 Regulation of apoptosis
				5										GO:0043067 Regulation of programmed cell death
16														GO:0006950 Response to stress

Total Enriched GO Terms (10% of genes in cluster)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
BP	0	8	14	9	30	18	2	21	25	11	6	10	14	17
MF	0	3	3	1	3	7	2	9	10	5	4	4	4	10
CC	0	0	8	0	1	7	4	6	7	1	3	3	3	7

Figure 4.7. Continued overleaf.

Figure 4.7. Biological trends revealed by significant enrichment of Gene Ontology biological process terms in each expression cluster. GO annotations for genes in each cluster were obtained using WebGestalt (Zhang *et al* 2005), and significantly enriched GO terms were identified using the hypergeometirc test ($p<0.05$, apart from clusters 3 and 10: $p<0.01$, 2.7.4.1). An additional criterion that enriched categories must also contain $\geq$10% of genes in the cluster was used. Numbers within grey boxes indicate the number of genes from a cluster with a specific GO term. The number of dots adjacent to each GO term indicates the level in the hierarchical GO tree at which it was found. * Due to software limitations only the 525 most significantly changed genes in these clusters could be included for analysis. GO categories: BP, biological process; MF, molecular function; CC, cellular component.

apoptosis" (Appendix Table 4); although this GO term was significantly enriched (p<0.05) it contained less than 10% of genes in the cluster so was not included in the main figure. Furthermore, the morphological data (Fig. 4.2.B) indicates that there is no visible decrease in cell number at 9h that could reflect apoptotic cell death.

In the initial hour of differentiation, a small number of mitosis and chromosome organisation-related genes are transiently upregulated (cluster 8). Other cell cycle-related genes are also upregulated at 7d following downregulated expression (cluster 4); half of these are involved in negative regulation of progression through the cell cycle (Appendix Table 4). A significant number of cell cycle genes are downregulated at 7d (cluster 10), and few of these genes are involved in the negative regulation of cell cycle progression (GO term not significantly enriched). This fits with the observation that differentiation requires growth arrest, and indicates that this takes place after the first 9h of differentiation.

Cell communication is defined as a process by which a cell communicates with it's surroundings, and includes genes in the signal transduction and cell adhesion categories. These terms represent a significant proportion of genes downregulated at 9h (cluster 11), and 7d (cluster 10). This suggests certain signalling cascades may need to be repressed in order for the correct channelling of adipogenic signals during differentiation. Additionally, it indicates that as the transition from stem cell to adipocyte is made, specific subsets of genes that may have maintained the interaction of hMSCs with the environment are not required. Instead, adipocytes may require a different selection of genes that mediate this role, perhaps reflecting the altered environment in which they reside (adipose tissue), and their new endocrine role (1.5.2). Cell and organelle organisation and biogenesis are also repressed at 9h or 7d (clusters 10, 11 and 14). Such genes are involved in the assembly and maintenance of cell organelles or structures. This indicates that early adipogenesis does not involve large changes in internal cell structure.

Cluster 3, representing genes upregulated at the endpoint of differentiation, was significantly enriched for genes involved in lipid metabolism, a major process expected to operate within an adipocyte. Most of these genes are also located within the biosynthesis

category, indicating that lipid is being synthesised from smaller molecules, an energy-requiring process. Carboxylic acid metabolism and generation of precursor metabolites and energy genes are enriched, suggesting that the cells are indeed producing high amounts of energy needed for use in these metabolic processes. A large subset of genes that are upregulated at 7d are involved in localisation and transport, which encompass uptake and intracellular transport of macromolecules and larger cellular components. This fits with the dramatic change in morphology observed during adipogenesis, as well as the increased requirement for nutrients to supply the metabolic processes occurring.

4.4.2 Identification of functional trends using manual database annotation

Although GO annotation yielded detailed information regarding the processes involved in early adipogenesis, limitations to it's functionality mean that important data may be missed. For example, not all genes are described by GO terms, and the method by which gene function is annotated can sometimes be vague, or is not based entirely on experimental observation but may also involve some speculative comments. Furthermore, the hierarchical nature of the GO classification system means that large, "parent" categories that give a vague view of gene function may be deemed enriched in a specific cluster (for example "metabolism", Fig. 4.7) whereas smaller daughter categories (i.e. categories within the parent category where genes are grouped more specifically according to their role) that give a better view of adipogenic processes (for example "regulation of transcription", a daughter of "metabolism") are not. Therefore, as well as automated analyses of global processes, detailed annotation of the function of individual genes was undertaken using manual searches of online databases including EntrezGene, OMIM, PubMed and GeneCards (2.7.4.2). The analysis was focused mainly on clusters showing the earliest regulation during differentiation, namely clusters 1, 5-9, and 12, in order to gain further insight into the events characterising early adipogenesis. Several biological trends were identified that occurred either in a single cluster, or were noted throughout early differentiation.

Co-expression of genes in cluster 5 regulated by members of the induction cocktail.
Around 33% of genes in cluster 5 (containing 72 genes, permanently upregulated from 3h) for which a function could be annotated were found to be either directly regulated by

constituents of the induction cocktail used to induce adipogenesis, or to be involved in pathways stimulated by these reagents. Glucocorticoid-regulated genes were most abundant (Table 4.2). These genes included C/EBPδ and GILZ, which are known to regulate adipogenesis positively and negatively respectively (Wu *et al* 1996, Shi *et al* 2003). RASD1 is also regulated by glucocorticoids but is not known to play a role in adipogenesis; it's exact function is unknown but it may be involved in dexamethasone-induced alterations in cell morphology, growth and extracellular matrix interactions (Kemppainen *et al* 2003, Vaidyanathan *et al* 2004). It's co-expression with adipogenic regulators combined with their conserved method of regulation could indicate a role for RASD1 in cooperating with such factors to control adipogenesis.

The upregulation of genes involved in insulin/IGF signalling and cAMP-regulated processes was also observed in cluster 5. These include FOXO1A, which plays a major role in mediating insulin-induced adipogenesis (Nakae *et al* 2003). FOXO1A is also known to be important for glucocorticoid-stimulated PDK4 expression (Kwon *et al* 2004); PDK4 is also found in cluster 5. Additionally, one probeset for PDE4D (cAMP-specific) is located in cluster 5. PDE4D degrades the second messenger cAMP (the intracellular level of which is increased by IBMX), thus controlling signalling by this molecule, and some isoforms of this gene are upregulated by cAMP levels (LeJeune *et al* 2002). Although 4 other probesets for this gene are located in cluster 12, they also show upregulation at 3h. These observations indicate a high level of functional clustering at this timepoint; many genes with similar functions are co-expressed. Additionally, it highlights the ability of differentiation inducers to stimulate transcriptional regulation at a specific timepoint. These results therefore suggest that an important step in the control of adipogenesis at 3h may be the coordinated upregulation of several signalling pathways mediated by the differentiation inducers, which act in concert to transmit signals throughout adipogenesis, as their expression is permanently upregulated.

Regulation of developmental genes at a specific stage of adipogenesis. Cluster 5 (permanent upregulation from 3h) and cluster 9 (218 genes, permanent downregulation from 3h) both contained a number of genes found to play a role in the development or function of mesodermal and non-mesodermal tissues (Table 4.3). Indeed, the GO term

Cluster 5: Genes related to signalling by induction cocktail components		
Glucocorticoid-regulated/ processes	Insulin/IGF1- regulated/processes	cAMP- regulated/processes
RAS, dexamethasone-induced 1 (RASD1) Delta sleep inducing peptide, immunoreactor (GILZ) Glucocorticoid induced transcript 1 (GLCCI1) Pyruvate dehydrogenase kinase, isoenzyme 4 (PDK4) FK506 binding protein 5 (FKBP5) Metallothionein 1K (MT1K) Metallothionein 1F (functional) (MT1F) CCAAT/enhancer binding protein δ (C/EBPδ) Inhibitor of DNA binding 4 (ID4) Forkhead box O1A (FOXO1A)	Forkhead box O1A (FOXO1A) Inositol(myo)-1(or 4)- monophosphatase 2 (IMPA2) Suppressor of cytokine signalling 2 (SOCS2) Insulin receptor substrate 2 (IRS2) Phosphoinositide-3-kinase, regulatory subunit, polypeptide 1 (PIK3R1)	Phosphodiesterase 4D, cAMP-specific (PDE4D)

Table 4.2. Genes in cluster 5 (permanently upregulated at 3h) that are regulated by, or act downstream of, induction cocktail components. These functions were identified using manual database searches (2.7.4.2).

A	Genes implicated in the development/function of other lineages - Cluster 5		
Neural	Mesenchymal tissue formation		
Snail homolog 2 (Drosophila) (SNAI2) Suppressor of cytokine signaling 2 (SOCS2) Basic helix-loop-helix transcription factor 6 (HATH6) Putative small membrane protein (NID67)	Snail homolog 2 (Drosophila) (SNAI2) Inhibitor of DNA binding 1 (ID1)	Sine oculis homeobox homolog 2 (Drosophila) (SIX2) Inhibitor of DNA binding 4 (ID4)	T-box 2 (TBX2)

B	Genes implicated in the development/function of other lineages - Cluster 9			
Neural	Bone	Muscle	Cartilage	Adipose
Chromosome 5 open reading frame 13 (neuronal protein 3.1)	Osteoblast specific factor 2 (fasciclin I-like) (OSF2)	Potassium inwardly- rectifying channel, subfamily J, member 2 (KCNJ2)	A disintegrin-like and metalloprotease with thrombospondin type 1 motif, 1 (ADAMTS1)	Eight twenty one protein/myeloid translocation gene on 8q22 (ETO/MTG8)
Potassium inwardly-rectifying channel, subfamily J, member 2 (KCNJ2)	Vitamin D (1,25-dihydroxyvita min D3) receptor (VDR)	Mesenchyme homeo box 2 (MEOX2)	A disintegrin-like and metalloprotease with thrombospondin type 1 motif, 5 (aggrecanase-2) (ADAMTS5)	Ectodermal-neural cortex (with BTB-like domain) (ENC1)
Prickle-like 1 (Drosophila) (PRICKLE1)	Mesenchyme homeo box 2 (MEOX2)	Ankyrin repeat domain 1 (cardiac muscle) (ANKRD1)		High mobility group AT-hook 2 (HMGA2)
Ectodermal-neural cortex (with BTB-like domain) (ENC1)	Hepatocyte growth factor (HGF)	Tripartite motif-containing 32 (TRIM32)		Kruppel-like factor 6 (KLF6)
Ras and Rab interactor 2 (RIN2)		Muscleblind-like (Drosophila) (MBNL1)		Kruppel-like factor 2 (KLF2)
Activity-dependent neuroprotector (ADNP) Neuron navigator 1 (NAV1) Kruppel-like factor 6 (KLF6) Homeobox B2 (HOXB2)		GATA-binding protein 6 (GATA6)		Early growth response 1 (EGR1/KROX20)

Table 4.3. Genes implicated in the development/function of other lineages identified by manual annotation. A, Genes in cluster 5 (permanently upregulated at 3h) and B, genes in cluster 9 (permanently downregulated at 9h) with functions related to the development or function of mesodermal and non-mesodermal lineages. These functions were identified using manual database searches (2.7.4.2).

173

"development" was found to be enriched in cluster 9. Cluster 5 was found to contain 2 genes involved in the regulation of development during GO annotation (Appendix Table 4) but this did not pass the stringent significance criteria used. However, manual characterisation identified gene functions that were not necessarily covered by GO annotation, and in this way 20% of genes with a known function in cluster 5 were defined as having a function related to development. These included the transcription factor SIX2, the exact function of which is unknown, but SIX gene family members are known to be involved in vertebrate development and maintenance of the differentiated state of tissues (Laclef *et al* 2003, Lagutin *et al* 2003). Although involved in the development of other lineages, such genes may have an additional role in controlling adipogenesis, but functional investigation would be required to confirm this.

Genes controlling the development of a number of tissues constituted 17% of cluster 9, and thus were permanently downregulated from 3h. Many of these genes are associated with the bone, muscle and cartilage lineages, which MSCs are capable of forming. For example, the transcription factor MEOX2 was downregulated, and has been shown to be a regulator of vertebrate limb myogenesis (and skeletal development in concert with MEOX1) (Mankoo *et al* 1999, 2003). The expression of a subset of neuronal genes including ADNP, which plays a role in neurodevelopment (Pinhasov *et al* 2003), was particularly interesting as it may imply that hMSCs are capable of differentiating into this lineage - a matter under considerable debate (Lu *et al* 2004, Neuhuber *et al* 2004). Furthermore, some of the genes in this cluster were known regulators of adipogenesis, and included ENC1 which is also involved in the function of neuronal cell types (Zhao *et al* 2000). These results suggest that hMSCs express a number of genes associated with different lineages, and an important step during early adipogenesis could be the permanent downregulation of such genes in order to steer differentiation along the adipogenic pathway. Furthermore, it is possible for such genes to regulate the development of more that one tissue type, suggesting that some may have additional roles specific to the earliest stages of adipogenesis.

Occurrence of chromatin modification throughout adipogenesis. Chromatin modification genes were found to be regulated throughout adipogenesis, and as such were

174

located in numerous expression clusters. The transcriptional regulation of SMARCF1 during adipogenesis has never before been reported; here it was identified in cluster 8, and showed transient upregulation at 1h. It is thought to be a component of the SWI/SNF ATP-dependent chromatin remodelling complex, and is known to transactivate glucocorticoid receptor-dependent transcription (Inoue *et al* 2002); this is notable when associated with the upregulation of many glucocorticoid-responsive genes at 3h. Other genes associated with chromatin remodelling were located in clusters 1, 6, 8 and 9. Their frequency of regulation, both positive and negative, throughout adipogenesis, may indicate that the control of this process plays an important role at many stages of differentiation. Indeed, transcription regulators are highly enriched in clusters 6, 8 and 9 where chromatin remodellers are found (Fig 4.7), indicating that some chromatin remodellers may be expressed in order to regulate transcription at specific differentiation stages, whereas others must be repressed into order for the correct transcriptional program to proceed.

4.4.3 Role of classical signalling pathways in adipogenesis

The Wnt, TGFβ and BMP signalling pathways are known to regulate adipogenesis: Wnt and TGFβ signalling inhibit the process whereas BMP signalling promotes differentiation (Ignotz and Massague 1985, Ahrens *et al* 1993, Asahina *et al* 1996, Locklin *et al* 1999, Ross *et al* 2000, Choy *et al* 2000, Bennett *et al* 2002, Choy and Derynck 2003, Tang *et al* 2004). However, little detail is known regarding the temporal regulation of specific pathway members throughout the process of differentiation in hMSCs. This was investigated by generating graphical pathway maps to represent significant expression changes occurring in these pathways during each phase of differentiation (0h-1h, 1h-3h, 3h-9h, 9h-7d, Figs. 4.8 and 4.9). Pathway maps were generated using Ingenuity Pathway Analysis (Ingenuity® Systems, www.ingenuity.com, 2.7.4.1), and modifications were made where pathway members not included on the maps were identified during manual annotation.

Wnt pathway regulation during adipogenesis. The first transcriptional regulation of the canonical Wnt pathway occurs between 1-3h, where pathway inhibitors DKK1 and DACT1 are downregulated, along with upregulation of the Wnt receptor FZD4 and PPP2R1B, a regulatory subunit of PP2A which stabilises the β-catenin complex (Fig. 4.8).
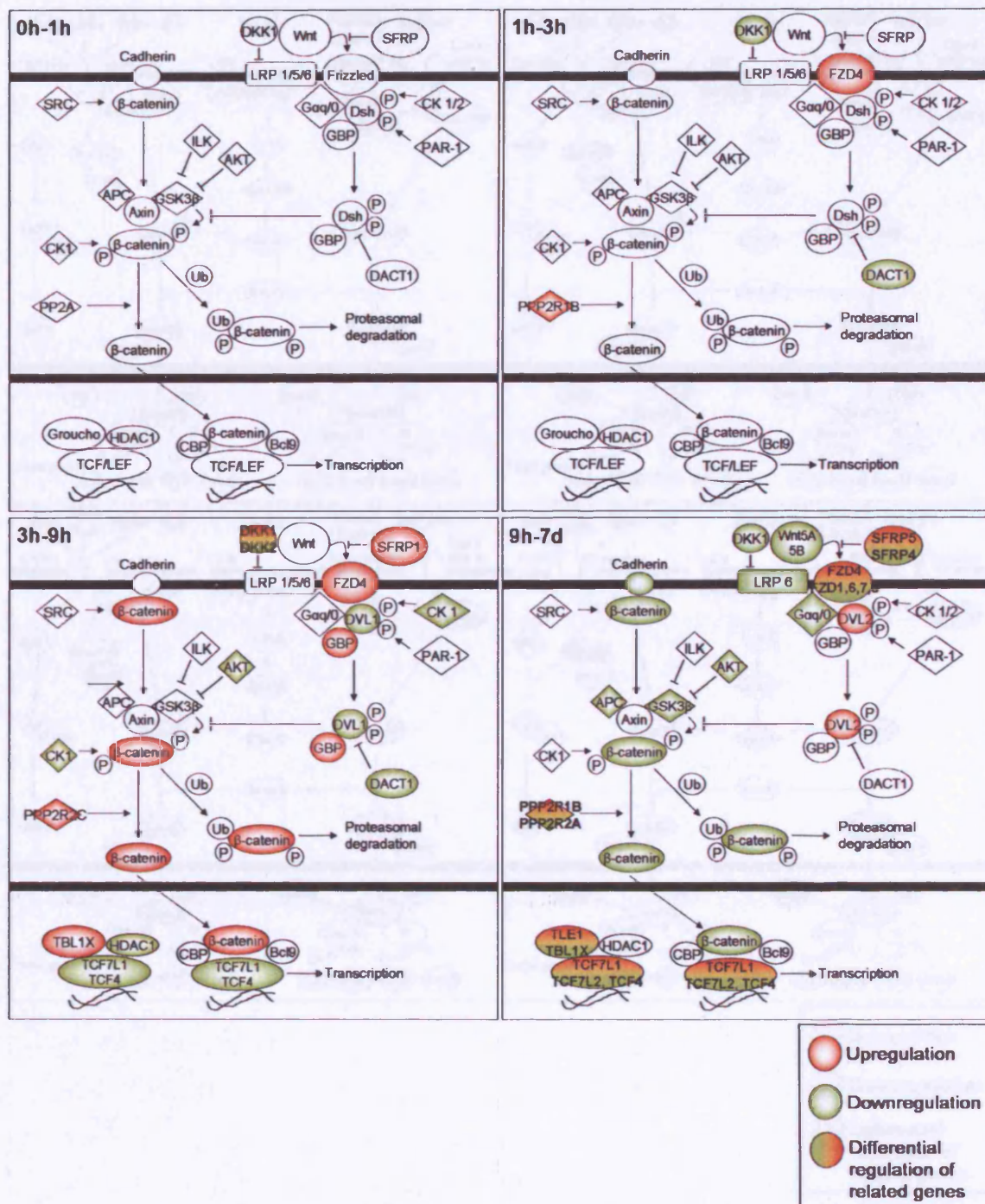
Figure 4.8. Temporal regulation of Wnt signalling pathway components during hMSC adipogenesis. Log_2 fold-change values for significantly changed genes from each phase of differentiation were mapped onto pathway diagrams using Ingenuity Pathway Analysis (2.7.4.1). Figure adapted from Ingenuity Pathway Analysis (Ingenuity® Systems, www.ingenuity.com).

Figure 4.9. Temporal regulation of TGFβ and BMP signalling pathway components during hMSC adipogenesis. Log$_2$ fold-change values for significantly changed genes from each phase of differentiation were mapped onto pathway diagrams using Ingenuity Pathway Analysis (2.7.4.1). Figure adapted from Ingenuity Pathway Analysis (Ingenuity® Systems, www.ingenuity.com).
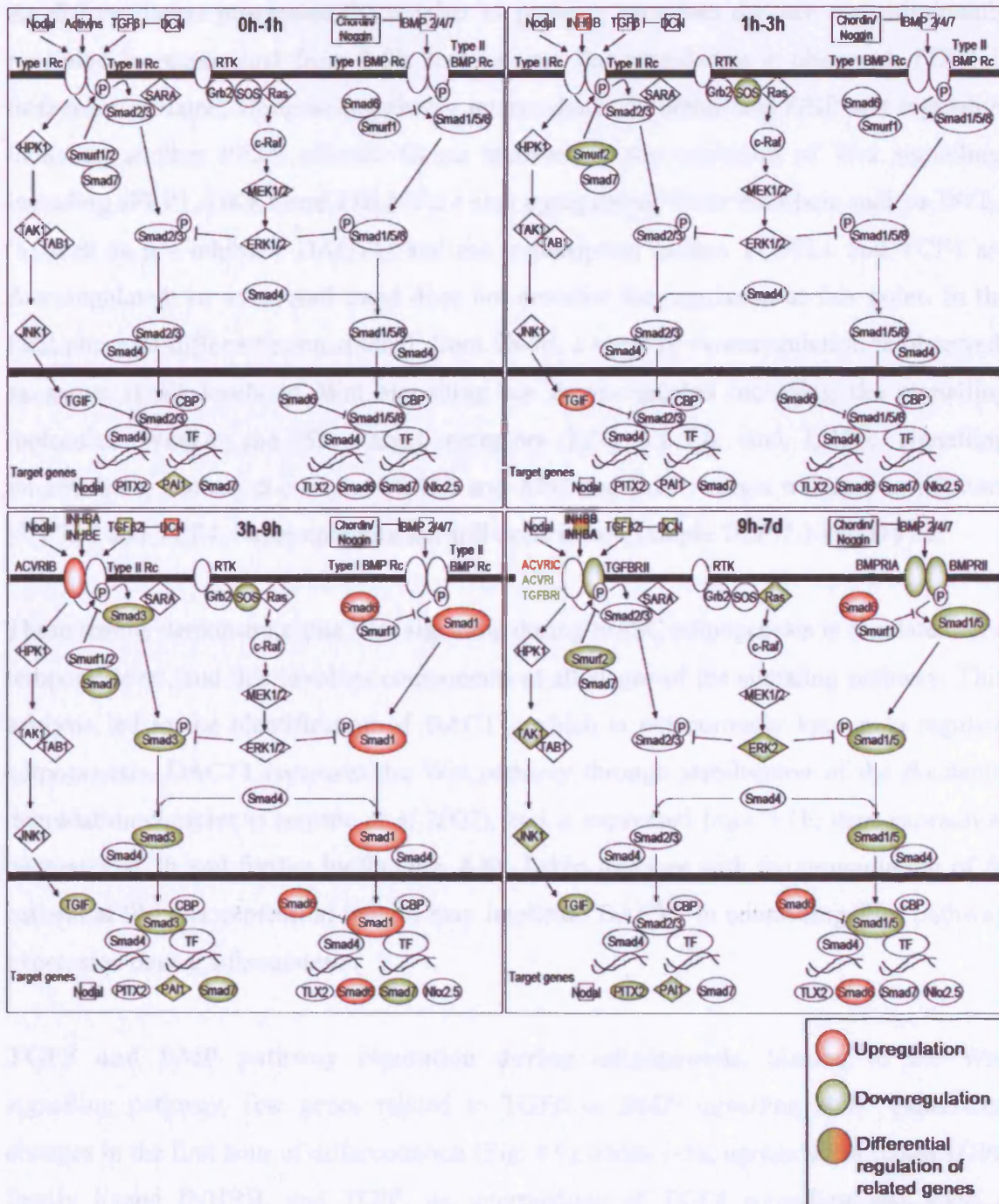
As differentiation progresses the number of pathway members that are transcriptionally regulated increases, and from 3-9h both up and downregulation is observed. FZD4 is further upregulated, along with pathway intermediates β-catenin and GBP and regulators including another PP2A subunit. Genes involved in the inhibition of Wnt signalling, including sFRP1, DKK1 and TBL1X are also upregulated. Other members such as DVL1 (as well as it's inhibitor DACT1) and the transcription factors TCF7L1 and TCF4 are downregulated, so an overall trend does not describe the regulation at this point. In the final phase of differentiation studied, from 9h-7d, a trend of downregulation is observed, as genes at all levels of Wnt signalling are downregulated including the signalling molecules Wnt5A and 5B, target receptors FZD 1,6,7,8, and LRP6, signalling intermediates such as β-catenin, GSK3β and APC and finally target transcription factors TCF7L2 and TCF4. Some upregulation still occurs, for example TCF7L1 and DVL2.

These results demonstrate that Wnt signaling during hMSC adipogenesis is regulated on a temporal level, and this involves components at all stages of the signaling pathway. This analysis led to the identification of DACT1, which is not currently known to regulate adipogenesis. DACT1 represses the Wnt pathway through stabilisation of the β-catenin degradation complex (Cheyette *et al* 2002), and is expressed from 0-1h, then expression decreases at 3h and further by 9h (Fig. 4.8). Taken together with the upregulation of β-catenin at 9h, this expression pattern may implicate DACT1 in controlling Wnt pathway expression during adipogenesis.

TGFβ and BMP pathway regulation during adipogenesis. Similar to the Wnt signalling pathway, few genes related to TGFβ or BMP signalling show expression changes in the first hour of differentiation (Fig. 4.9). From 1-3h, upregulation of the TGFβ family ligand INHBB, and TGIF, an intermediate of TGFβ signalling and itself a transcriptional target of the pathway is observed. Also at this stage, the TGFβ signalling inhibitor SMURF2 is downregulated, whereas no regulation of BMP signalling members is detected. From 3-9h, both pathways show interesting regulatory trends, as the major SMAD protein in BMP signalling, SMAD1, is upregulated, along with a BMP pathway target SMAD6. This is in contrast to the TGFβ branch, where SMAD3 is downregulated along with TGIF, various ligands and the pathway targets PAI1 and SMAD7. ACVR1, a

TGFβ-type receptor, and DCN which sequesters TGFβ ligands, are upregulated. Regulation of these pathways then shows significantly different patterns in the final phase of differentiation. Every BMP pathway member that changes expression is downregulated, apart from the target protein SMAD6. This pattern is mirrored in the TGFβ pathway, where kinases that mediate an alternative route for TGFβ signalling are downregulated, combined with multiple ligands and receptors.

These observations indicate that the TGFβ and BMP pathways are regulated in contrasting directions during adipogenesis; the general trend for TGFβ pathway members during differentiation is downregulation, whereas some BMP signalling components are transiently upregulated at 9h. This could suggest different requirements for these pathways at specific stages of adipogenesis. Analysis of the TGFβ signalling pathway revealed regulation of TGIF, which has not been shown previously during adipogenesis (Fig. 4.9). TGIF, as well as being transcriptionally induced by TGFβ1 signalling, acts as a co-repressor of SMAD2/3-mediated transcriptional activation by forming repressive complexes containing histone deacetylases (Wotton et al 1999, Chen et al 2003), so may be involved in a negative feedback loop to inhibit TGFβ signalling. The pattern of TGIF upregulation at 3h followed by downregulation alongside other TGFβ components could indicate a role for TGIF during adipogenesis.

4.5 Identification of novel candidate regulators of early adipogenesis

Characterisation of novel regulators of early adipogenesis would further knowledge of the molecular control of early stem cell differentiation. The identification in this study of 6319 probesets regulated during adipogenesis provides a unique opportunity to do this. Manual functional annotation of clusters representing the earliest regulation (described in section 4.4.2) as well as some analysis of additional clusters was therefore used to identify candidates for novel regulators of adipogenesis. A number of selection criteria were designed to identify such genes, which were deemed of interest if they passed at least one of these criteria:

- Functional association with differentiation or tissue development, especially of mesenchymal cells/tissues.

- Functional association with genes (e.g. C/EBP proteins) or processes (e.g. cell shape change) known to play a part in adipogenesis, but with no known role in the process.

- Specifically defined as regulators of signalling pathways involved in adipogenesis (e.g. Wnt pathway), but with no known role in the process.

- Regulated by the reagents used to induce adipogenesis, or acting as part of their downstream pathways.

- Transcription factors/factors involved in the regulation of transcription were deemed particularly interesting as they have the ability to control large subsets of downstream genes, so may have a significant role in differentiation.
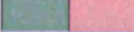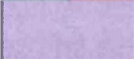
This analysis identified over 100 interesting candidate genes, summarised in Table 4.4, which were regulated during adipogenesis but had no known role in the process. This number highlights the potential of this study to identify possible novel early adipogenic regulators. Below, selected candidates of particular interest are discussed.

HOXB2 and SNAIL1 were both chosen mainly due to their known roles in the development of other tissues. SNAIL1 is a transcriptional repressor of ectodermal genes within the neural crest and mesoderm during development, in a process known as epithelial-mesenchymal transition (Leptin 1991). Transient upregulation of SNAIL1 at 3h (cluster 6) could indicate a requirement for suppression of non-mesodermal genes during adipogenesis that may be mediated by this gene. HOXB2 belongs to the homeobox family of transcription factors that are known to be involved in development, and is thought to have a specific role in neuronal patterning and differentiation in certain areas of the mouse hindbrain (Gavalas et al 2003). HOXB2 is permanently downregulated at 3h (cluster 9); if downregulation of genes involved in the differentiation of other tissues (such as HOXB2) is crucial for the progression of adipogenesis, alteration of the expression of this factor may affect differentiation of hMSCs to adipocytes.

Cluster	Gene Name	Gene Symbol	Reason for selection	Present in shRNA library Y/N
1	Helicase (DNA) B	HELB		N
1	Ret finger protein	RFP		Y
2	Transforming growth factor, beta receptor III (betaglycan, 300kDa)	TGFBR3		Y
5	Casein kinase 1, delta	CSNK1D		Y
5	Growth arrest-specific 1	GAS1		Y
5	Glucocorticoid induced transcript 1	GLCCI1		N
5	Basic helix-loop-helix transcription factor 6	HATH6		Y
5	Inositol(myo)-1(or 4)-monophosphatase 2	IMPA2		Y
5	Putative small membrane protein NID67	NID67		N
5	Phosphodiesterase 4D, cAMP-specific (phosphodiesterase E3 dunce homolog, Drosophila)	PDE4D		Y
5	Period homolog 1 (Drosophila)	PER1		Y
5	Proline-rich nuclear receptor coactivator 1	PNRC1		N
5	Prickle-like 2 (Drosophila)	PRICKLE2		N
5	RAS, dexamethasone-induced 1	RASD1		Y
5	Sine oculis homeobox homolog 2 (Drosophila)	SIX2		Y
5	Snail homolog 2 (Drosophila)	SNAI2		Y
5	Suppressor of cytokine signaling 2	SOCS2		Y
5	T-box 2	TBX2		Y
5	Transforming growth factor beta-Stimulated protein TSC-22	TSC22		Y
6	Basic transcription element binding protein 1	BTEB1		Y
6	Chromobox homolog 4 (Pc class homolog, Drosophila)	CBX4		Y
6	cAMP responsive element binding protein 3-like 2	CREB3L2		N
6	cAMP responsive element modulator	CREM		Y
6	Cofactor required for Sp1 transcriptional activation, subunit 7, 70kDa	CRSP7		Y
6	Diphtheria toxin receptor (heparin-binding epidermal growth factor-like growth factor)	DTR		Y
6	Ephrin-B2	EPHB2		Y
6	F-box and leucine-rich repeat protein 12	FBXL12		Y
6	Forkhead box Q1	FOXQ1		N
6	GTF2I repeat domain containing 1	GTF2IRD1		Y
6	Jumonji homolog (mouse)	JMJ		Y
6	Kruppel-like factor 4	KLF4		Y
6	Kruppel-like factor 13	KLF13		Y
6	Large tumor suppressor, homolog 2 (Drosophila)	LATS2		Y

Cluster	Gene Name	Gene Symbol	Reason for selection	Present in shRNA library Y/N
6	Lipoma HMGIC fusion partner-like 2	LHFPL2		Y
6	Insulin receptor tyrosine kinase substrate	LOC55971/ IRTKS		Y
6	Mesoderm development candidate 1	MESDC1		Y
6	Mlx interactor	MONDOA		Y
6	Nuclear receptor subfamily 4, group A, member 1	NR4A1		Y
6	Nuclear receptor subfamily 4, group A, member 2	NR4A2		Y
6	Nuclear receptor subfamily 4, group A, member 3	NR4A3		Y
6	PR domain containing 4	PRDM4		Y
6	LIM domain protein	RIL		Y
6	Snail homolog 1 (Drosophila)	SNAI1		Y
6	Sprouty homolog 4 (Drosophila)	SPRY4		N
6	Slingshot 1	SSH1		Y
6	Transcription factor 8 (represses interleukin 2 expression)	TCF8		Y
6	Transducin-like enhancer of split 3 (E(sp1) homolog, Drosophila)	TLE3		Y
6	Tumor necrosis factor receptor superfamily, member 1B	TNFRSF1B		Y
6	Transducer of ERBB2, 1	TOB1		Y
6	Zinc finger protein 36, C3H type, homolog (mouse)	ZFP36		N
6	Zinc finger protein, multitype 2	ZFPM2		Y
6	Zinc finger protein 331	ZNF331		N
7	S100 calcium binding protein A10 (annexin II ligand, calpactin I, light polypeptide (p11))	S100A10		Y
8	Cysteine-rich motor neuron 1	CRIM1		N
8	E1A binding protein p400	EP400		N
8	Helicase, lymphoid-specific	HELLS		N
8	Myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLL		Y
8	Nuclear factor of activated T-cells 5, tonicity-responsive	NFAT5		Y
8	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily f, member 1	SMARCF1		Y
8	Sprouty homolog 2 (Drosophila)	SPRY2		Y
8	Zinc finger protein 288	ZNF288		N
9	Rho GTPase activating protein 18	ARHGAP18		N
9	KIAA0537 gene product aka ARK5 AMP-activated protein kinase family member 5	ARK5		Y
9	Bromodomain containing 8	BRD8		Y
9	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2	CITED2		Y

Cluster	Gene Name	Gene Symbol	Reason for selection	Present in shRNA library Y/N
9	Core promoter element binding protein/ Kruppel-like factor 6	COPEB/ KLF6		Y
9	Dapper homolog 1, antagonist of beta-catenin (xenopus)	DACT1		Y
9	Deleted in liver cancer 1	DLC1		Y
9	Dual specificity phosphatase 14	DUSP14		Y
9	E2F transcription factor 7	E2F7		N
9	Eukaryotic translation initiation factor 2C, 4	EIF2C4		N
9	GRB2-associated binding protein 1	GAB1		Y
9	GATA binding protein 6	GATA6		Y
9	HMG-box transcription factor 1	HBP1		N
9	Hepatocyte growth factor (hepapoietin A; scatter factor)	HGF		Y
9	Homeo box B2	HOXB2		Y
9	Immediate early response 3	IER3		Y
9	Inhibitor of growth family, member 3	ING3		Y
9	Jub, ajuba homolog (Xenopus laevis)	JUB		N
9	V-jun sarcoma virus 17 oncogene homolog (avian)	JUN		Y
9	Kruppel-like factor 7 (ubiquitous)	KLF7		Y
9	Mesenchyme homeo box 2 (growth arrest-specific homeo box)	MEOX2		Y
9	Nuclear receptor coactivator 6	NCOA6		Y
9	Protein kinase (cAMP-dependent, catalytic) inhibitor alpha	PKIA		Y
9	Protein phosphatase 2C, magnesium-dependent, catalytic subunit	PPM2C		N
9	Protein phosphatase 1, regulatory (inhibitor) subunit 12A	PPP1R12A		Y
9	Prickle-like 1 (Drosophila)	PRICKLE1		N
9	Protein kinase C-like 2	PRKCL2		Y
9	Pumilio homolog 2 (Drosophila)	PUM2		N
9	RAD54B homolog	RAD54B		Y
9	Regulator of G-protein signalling 4	RGS4		Y
9	E3 ubiquitin ligase SMURF2	SMURF2		Y
9	Suppressor of cytokine signalling 5	SOCS5		Y
9	Son of sevenless homolog 1 (Drosophila)	SOS1		Y
9	Nuclear antigen Sp100	SP100		Y
9	Transcriptional regulator interacting with the PHS-bromodomain 2	TRIP-Br2 aka SERTAD2		Y
9	Vitamin D (1,25- dihydroxyvitamin D3) receptor	VDR		Y
9	Zinc finger protein 143 (clone pHZ-1)	ZNF143		Y
9	Zinc finger protein 238	ZNF238		Y
10	Development and differentiation enhancing factor 1	DDEF1		N
10	FAT tumor suppressor homolog 1 (Drosophila)	FAT		Y

183

Cluster	Gene Name	Gene Symbol	Reason for selection	Present in shRNA library Y/N
10	Serum/glucocorticoid regulated kinase	SGK		Y
10	TGFB-induced factor (TALE family homeobox)	TGIF		Y
12	Dual specificity phosphatase 1	DUSP1		Y
12	Dual specificity phosphatase 5	DUSP5		Y
12	Potassium voltage-gated channel, Isk-related family, member 4	KCNE4		N
13	Homeo box A7	HOXA7		Y
13	BMP and activin membrane-bound inhibitor homolog (Xenopus laevis)	BAMBI		N

KEY	Reason for selecting gene of interest (GOI)
	Function associated with differentiation/development (of mesenchymal tissues)
	Association with adipogenic factors/processes but no known role in the process
	Novel regulators of signalling pathways involved in adipogenesis
	Regulated by the reagents used to induce adipogenesis, or part of their downstream pathways
	Transcription factor/involved in regulation of transcription

Table 4.4. Novel candidates for adipogenic regulators identified by functional annotation. Functional information for genes showing early regulation was obtained using manual database searches and interesting candidate genes were selected using stringent criteria related to relevant functional information (Key). These functions were identified using manual database searches (2.7.4.2).

The identification of TCF8 and FAT as genes of interest was based primarily on their roles in processes related to adipogenesis, or a relationship to adipose tissue. TCF8 resides at a major susceptibility locus for obesity on chromosome 10 (Hager *et al* 1998, Hinney *et al* 2000). It is a transcription factor that has been shown to augment BMP2 and TGFβ1 signalling (Postigo 2003), and plays a role in skeletal development (Takagi *et al* 1998), indicating that it also has a role in pathways known to control mesodermal differentiation. TCF8 was upregulated early in another microarray study of hMSC adipogenesis (Sekiya *et al* 2004). Here, it showed early upregulation from 1-3h followed by downregulation at 7d in one probeset (cluster 6); 3 other probesets for this gene showed only downregulation at 7d (cluster 10). The probeset showing upregulation at 1-3h had the identification "_s_at", meaning that it shared some probes in common with another gene. However, the other gene (SNF1-like kinase) was represented by 2 further probesets on the U133 Plus 2.0 array (as stated on NetAffx, www.affymetrix.com/analysis/index.affx), but these probesets did not show altered expression during adipogenesis. It could not be definitively concluded therefore that cross-hybridisation by SNF1-like kinase transcripts had caused the observed expression change, so TCF8 was included for further analysis. The patterns of expression change might suggest that TCF8 could play an early role in adipogenesis, but is not required at the later stages, or must be downregulated in order for the final stage of differentiation to proceed. It would be interesting to investigate further whether this gene plays a role in adipogenesis.

The FAT gene encodes a transmembrane receptor which is part of the cadherin superfamily (and was so named according to Drosophila FAT mutants being oversized due to increased cell proliferation). The Drosophila orthologue of this gene controls cell proliferation during development, and it is thought that FAT may also be important in cell communication (Dunne *et al* 1995). This gene showed downregulation at 7d (cluster 10). This regulation combined with the putative function of FAT may suggest that this gene has an early role in controlling cell proliferation and communication during early adipogenesis, or that it must be repressed during later differentiation.

DACT1 and TGIF, described in section 4.4.3, were identified due to their role as regulators of the Wnt and TGFβ signalling pathways, respectively. TOB1 was also

185

identified in this study. TOB1 interacts with the BMP-SMADs 1, 5 and 8 and inhibits their transcription factor activity (Yoshida *et al* 2000). TOB1 was upregulated at 3h, and then at 9h became downregulated (cluster 6) – this was the time at which expression of SMAD1 was seen, again implicating this gene as being involved in the regulation of BMP signalling during adipogenesis.

SGK was included as a candidate gene for a number of reasons. It is transcriptionally regulated by serum and glucocorticoids (Itani *et al* 2002), and has cAMP-dependent kinase activity, providing evidence of a relationship between this gene and the differentiation inducers used. It has also been shown to be activated by TGFβ signalling and high glucose levels, and is implicated in diabetic nephropathy (Saad *et al* 2005). During adipogenesis, SGK was upregulated at 1-3h followed by strong downregulation at 7d (cluster 10). This may indicate an early role for SGK, possibly mediated by one/several of the differentiation inducers, followed by repression of the gene in mature adipocytes which may be required for normal cellular function.

Finally, several genes that were identified as candidate adipogenic regulators during manual annotation have recently been confirmed in other work as regulators of either 3T3L1 or hMSC adipogenesis, although the mechanisms by which this occurs were not known in every case (Table 4.5 and Fig. 4.6). DTR was identified in this study as a potential candidate as its levels were shown to correlate with fat accumulation in humans (Matsumoto *et al* 2002) and a mitogenic role in fibroblasts and smooth muscles was described (Higashiyama *et al* 1991). In July 2005 DTR was shown to stimulate MSC proliferation whilst preventing differentiation (Krampera *et al* 2005). The upregulation of DTR at 3h in this study could possibly fit with an early proliferative role, and its downregulation from 9h-7d would be expected in order for differentiation to proceed. The potential to identify novel adipogenic regulators among this candidate list is therefore exemplified by this gene.

Gene	Function	Expression group	Reference
KLF6	Promotes adipogenesis through repression of DLK1	9	Li *et al* 2005
KLF15	Promotes adipogenesis through induction of PPARγ expression in cooperation with C/EBPα	3	Mori *et al* 2005a
KROX20	Promotes adipogenesis partly through C/EBPβ	4	Chen *et al* 2005
O/E-1	Promotes adipogenesis; mechanism unknown	3	Akerblad *et al* 2002
ETO/MTG8	Inhibits early adipogenesis through C/EBPβ	9	Rochford *et al* 2004
TAZ	Molecular switch promoting osteogenesis whilst inhibiting adipogenesis in MSCs	14	Hong *et al* 2005
GILZ	Sequence-specific transcriptional repressor of PPARγ	5	Shi *et al* 2003
KLF2	Inhibits preadipocyte differentiation to adipocytes, partly through DLK1 restoration	9	Banerjee *et al* 2003, Wu *et al* 2005
DTR	Promotes MSC proliferation but inhibits differentiation	6	Krampera *et al* 2005

Table 4.5. Genes recently implicated in adipogenesis that were identified in expression clusters.

4.6 Discussion

The experiments described in this chapter have demonstrated that early adipogenesis involves complex transcriptional regulation that occurs in defined temporal patterns. Identification and analysis of gene clusters with specific expression patterns revealed novel information regarding molecular processes that are important during early adipogenesis. From the large dataset produced, it was possible to identify numerous novel candidate regulators of adipogenesis.

These results have shown that adipogenesis involves transcriptional regulation that is initiated as early as 1h after the induction of differentiation and that 6319 probesets show significant expression changes over the time-course (given the significance cut-off criteria employed in this study). This indicates the rapidity by which the addition of hormonal inducers of differentiation was transduced into transcriptional changes, and demonstrates the overall complexity of adipogenic regulation. This level of detail and the early transcriptional response have not previously been described during adipogenesis. The largest arrays previously used on hMSC or murine 3T3L1 adipogenesis covered 12,000 (Sekiya *et al* 2004, Yloslato *et al* 2006) or 16,000 genes (Hackl *et al* 2005) respectively, whereas the microarrays used in this study covered all known human genes. The number of genes with significant expression changes during adipogenesis varied in other work but was at most 1016 (Burton *et al* 2004). This is likely to be a reflection of the different array sizes and selection criteria used to identify significant genes. I do not consider that the selection criteria used here, that identified over 6300 significant genes, were not stringent enough, as stricter criteria would be likely to miss small but important expression changes.

The increasing number of significant changes detected in each phase of differentiation may be a reflection of the increasing length of time between timepoints studied (Fig 4.1 and Fig 4.2B). However, when the expression profiles of replicate arrays were compared (Fig 4.2A), the 7d and 9h arrays, and also 3h though to a lesser extent, were found to display expression profiles unique to the particular timepoint, indicating that different stages of the transcriptional regulation of adipogenesis were captured at each timepoint.

188

Hierarchical clustering of significantly changed genes revealed that adipogenesis involves temporally ordered changes in gene expression which occur in distinct patterns. Fourteen separate gene clusters, each with a different regulatory pattern, were identified in this analysis. Smaller microarray studies of adipogenesis that employed hierarchical clustering detected 5-9 gene clusters (Ross et al 2002, Burton et al 2004, Nakamura et al 2003), and it would be expected that with a larger gene set, as in this study, more detailed regulation, and therefore a higher number of clusters, would be detected. Several studies used additional clustering methods such as k-means clustering and self-organising maps to further analyse these regulatory patterns (Burton et al 2004, Hackl et al 2005) but it was decided here that hierarchical clustering had identified a reasonable number of clearly separable clusters containing genes with similar regulatory patterns, so further analysis was unnecessary. All genes tested by RT-PCR analysis showed mRNA level changes that recapitulated the expression patterns detected by microarray analysis, so it can be concluded that the microarray analysis accurately reflected biological changes (Fig 4.3). It was also shown that every regulatory pattern assessed was conserved in differentiated hMSCs from several other donors (Fig. 4.5), which is an analysis that is infrequently performed in other hMSC work, but which demonstrates the important fact that the observations are not an artefact of the particular bone marrow used.

Evidence supporting this differentiation system as an accurate model of adipogenesis came from the identification of many crucial adipogenic regulators in the heatmap (Fig. 4.6). A small number of known regulators, for example GATA2 and -3, were not significantly changed during adipogenesis. A possible explanation for this is that they may not have passed significance level applied in this investigation (q<0.05). Indeed, transcription factors such as these need only small expression changes to have significant effects. Post-transcriptional changes such as phosphorylation, rather that transcriptional control, could be the major regulator of genes such as CREB, providing a possible reason for it's absence from the genes showing significant transcriptional change during differentiation. Furthermore, as explained earlier (section 4.1), it is possible that these genes were expressed to a high level during adipogenesis, but if their expression did not change then the comparative nature of microarray studies means that they would not have been included in the analysis. Other studies of hMSC or 3T3L1 adipogenesis tend to

identify some, but not all, known adipogenic regulators through microarray analysis of gene expression changes (for example, Sekiya *et al* 2004, Burton *et al* 2004), indicating that this is a common issue amongst experiments of this nature, although an additional causative factor in other studies is that the microarrays used may not have represented all known adipogenic regulators.

The expression at 7d of numerous genes known to play a role in adipocytes *in vivo*, as well as the enrichment at this endpoint of relevant metabolic pathways, indicates that at 7d cells resemble adipocytes *in vivo* both transcriptionally and morphologically (Table 4.1 and Fig. 4.1B). The 7d adipocytes were morphologically similar to brown adipocytes (multiple fat droplets rather than a single large droplet, 1.5.1), but studies of brown adipocyte morphology have mainly been performed in rodents. Furthermore, other studies of hMSC differentiation reveal multi-vesicle adipocytes (Pittenger *et al* 1999, Sekiya *et al* 2004), so this morphology could reflect either an inter-species difference in morphology, or that the induction cocktail use does not fully recreate the *in vivo* signals that initiate adipogenesis. However, as upregulation of the BAT-specific marker UCP1 was not observed, the transcriptional program occurring is not likely to be favouring differentiation towards brown adipocytes.

Genes that are regulated in a similar manner throughout adipogenesis, and hence fall into the same expression cluster, may have related roles. It was therefore possible to use the gene clusters to identify global events and processes that act at specific stages of differentiation. As the microarray used in this study covered all known human genes, this analysis has provided a level of detail and accuracy regarding global processes of adipogenesis that is greater than previous studies, which used smaller arrays and less comprehensive analysis (Sekiya *et al* 2004, Yloslato *et al* 2006). The information gained is limited only by the annotation method used. However, both automated and manual annotation were used to infer gene function, so a high level of detail was obtained.

Gene clusters representing expression changes at the earliest timepoints investigated contained high proportions of transcription factors and developmental genes (Fig. 4.7 and Table 4.3). Another early trend was the upregulation at 3h of genes involved in response

to the inducers of differentiation (Table 4.2). Later regulatory events included the downregulation of cell communication and the cell cycle, as well as upregulation of certain metabolic processes (Fig. 4.7). These findings suggest that the earliest events of adipogenesis are centred on the transduction of environmental stimuli into cellular responses which involve the action of many transcription factors. Some developmental genes and transcription factors are repressed whereas others, which were not expressed (or expressed at lower levels) in hMSCs, are upregulated. Such events are likely to direct cells along the adipogenic lineage and control the downstream events identified. These findings highlight the importance of the early stages of adipogenesis and also the potential from this study to understand more regarding the control of stem cell differentiation.

The results described here provide novel information in a number of ways. Similar to the findings described here, other studies using GO annotation also identified a decrease in cell cycle genes and an increase in lipid metabolism and bioenergetics genes during hMSC adipogenesis (Yloslato et al 2006). However, this study augments such work as the list of genes involved in these processes is more complete, and many more trends that may describe the earlier events of differentiation have also been identified. Adipogenesis in hMSCs has previously been shown to involve downregulation of genes specific to other lineages (Hung et al 2004a). Furthermore, a recent study showed that a subset of developmental genes, including members of the HOX (HOX A5, C8 and C9) and TBX (TBX15) families, were differentially expressed in different fat depots of mice or humans (subcutaneous and visceral, 1.5.2), and the expression of two of these genes (HOXA5 and TBX15) correlated with the extent of obesity in humans (Gesta et al 2006). It was concluded that such genes may play a role in body fat distribution and obesity. The discovery here of HOXA7, HOXB2 and TBX2 as differentially expressed during adipogenesis adds further detail to the array of developmental genes that may be involved, and the previous findings (Gesta et al 2006) add weight to a possible role for these genes in controlling adipogenesis. It was shown through manual annotation that certain processes, such as chromatin remodelling, occur throughout differentiation. It has recently been shown that downregulation of histone deacetylases, a subset of chromatin remodellers, during early 3T3L1 adipogenesis is required for their differentiation, and that histone acetylation at the promoters of adipocyte-specific genes it tightly associated with

their expression during differentiation (Yoo *et al* 2006). The identification of regulation of chromatin remodellers in this study, and their specific times of regulation, may serve to enhance understanding as to how they may control differentiation at the early stages.

Analysis of the expression of Wnt, TGFβ and BMP signalling pathways components provided an overview of transcriptional regulation of pathway members during hMSC adipogenesis (Figs. 4.8 and 4.9). As the major signalling events of these pathways occur at the post-transcriptional level (e.g. phosphorylation), gene expression changes may not directly affect signalling, so the functional consequences of the transcriptional changes cannot be concluded without further investigation. However, these analyses provide novel information in a number of ways; they indicate not only the exact pathway components that are regulated, but the specific stage of differentiation at which regulation occurs.

Although recent studies have identified some members of the Wnt pathway that are expressed in preadipocytes (Bennett *et al* 2002) and hMSCs (Etheridge *et al* 2004), expression of Wnt signalling pathway components has not been followed in detail throughout hMSC differentiation. The results described in this thesis define the temporal expression patterns of specific pathway members that are regulated during differentiation, and show that many of the same components are regulated that were identified previously (e.g. FZD1, TCF4, sFRP4, DKK1), but identifies regulation of further genes (e.g. DACT1, GBP, PP2A subunits). Additionally, this work provides a basis for future functional studies aimed at elucidating the stage at which Wnt signalling plays a role during hMSC adipogenesis. Activation of Wnt signalling at any point during the first 3 days of 3T3L1 differentiation was found to inhibit adipogenesis (Bennett *et al* 2002), and although it is known that Wnt signalling can inhibit hMSC adipogenesis, the exact timepoint at which this may occur is not known. The observation here that many Wnt pathway members are downregulated after 9h in this study might indicate that, similar to preadipocytes, hMSCs at the later stages of differentiation would be refractory to Wnt activators, so studies should focus on earlier timepoints to investigate the effects of Wnt signalling. Finally, the temporal view of transcriptional regulation of Wnt signalling could also be integrated with other regulation identified at similar timepoints, allowing inferences as to crosstalk between pathways during differentiation to be made.

TGFβ and BMP signalling are known to exert different effects on adipogenesis in hMSCs. The continued downregulation of many TGFβ pathway components identified (Fig 4.9) would fit with it's role as an inhibitor of differentiation, and the observation that this begins as soon as the 3h-9h phase might indicate that downregulation of this pathway is required early in order for differentiation to progress. SMAD3 mediates the inhibitory action of TGFβ in preadipocytes through inhibiting C/EBP protein transactivation function (Choy and Derynck 2003). Indeed, downregulation of SMAD3 rather than SMAD2 was observed in this work, and this occurred at the same timepoint that C/EBPβ and -δ expression increased. Therefore, many of the observations made here fit with previous work, but provide additional information as to the exact pathway members regulated and the timepoints at which this regulation occurs.

BMP-2,-4 and -7 can activate adipogenesis in bone marrow stromal cells (Ahrens et al 1993, Asahina et al 1996, Tang et al 2004). A significant upregulation of SMAD1, a transducer of BMP signalling was observed, as well as targets of this pathway, between 3h and 9h. SMAD1 and 5, as well as BMP-type receptors, were then downregulated during the latest phase of differentiation studied. Such temporal patterns have not previously been reported. It is tempting to speculate that any activatory role of BMP signalling would be more likely to be occurring during this early 3-9h phase when signalling components are upregulated, and later on, when BMP signalling is no longer required, the pathway is downregulated on the transcriptional level. Functional confirmation of these suggestions would add further knowledge regarding the methods by which these pathways control hMSC adipogenesis.

The major aim of this thesis was to identify novel factors involved in regulating the early stages of hMSC adipogenesis. Candidate genes were selected using functional information obtained from manual annotation of genes showing early regulation. The selection criteria employed optimised the likelihood of these genes having a role in adipogenesis; indeed, since candidate genes were identified, other research groups have described the role of a small number of these genes in adipogenesis (Table 4.5 KLF6 and DTR, Li et al 2005, Krampera et al 2005). This further supports the potential of other candidate genes in

playing a role in adipogenesis. In addition, genes shown to control the early stages of 3T3-L1 adipogenesis, for example KLF2 (Banerjee *et al* 2003, Wu *et al* 2005) and ETO/MTG8 (Rochford *et al* 2004) also showed early regulation during hMSC adipogenesis (cluster 9, downregulation from 3h, Table 4.5). This correspondence between the 3T3L1 and hMSC differentiation programs is likely to be due to the fact that similar induction cocktails are used to initiate adipogenesis in both systems, so some genes may respond in a similar manner. However, many more genes are co-regulated with these factors in this study, and thus are interesting candidates for novel adipogenic regulators. Moreover, as the hMSC system has been employed here it may be that these additional genes are not expressed in 3T3L1s. Importantly, the majority of candidate adipogenic regulators identified here are unique to this study.

Some candidate genes, including SNAI2, GAS1 and TCF8, were identified in previous studies of hMSC adipogenesis. GAS1 and SNAI2 were found by Nakamura *et al* to be upregulated from day 1 to day 9 during differentiation, and GAS1 was upregulated at day 3 in the study of Hung and co-workers (Nakamura *et al* 2003, Hung *et al* 2004a). Here, these genes were also shown to be upregulated, but it was found that this upregulation occurred as early as 3h (Cluster 5, Table 4.4), highlighting the importance of studying these early timepoints. TCF8 was upregulated at day 1 followed by downregulation in the work of Sekiya *et al* (2004); it was shown here to be down regulated at 7d, and one probeset showed upregulation at 1-3h (cluster 6); this regulation has not previously been reported and would be interesting to investigate further. A recent study of genes regulated during multilineage hMSC differentiation and their subsequent "de-differentiation" back to "hMSCs" after removal of the differentiation cocktails identified several genes in common with this thesis (Song *et al* 2006). Interestingly, one of these was DACT1, which showed decreased expression during adipogenic and osteogenic differentiation and was upregulated during "dedifferentiation", and hence was defined as a gene involved in maintaining the stem cell properties of hMSCs, but this function was not tested. DACT1 was also downregulated during differentiation in the work described here, so this correlation further demonstrates the potential of this work to identify novel genes that control the ability of hMSCs to undergo adipogenesis.

CHAPTER 5. INVESTIGATION OF THE ROLE OF NOVEL
CANDIDATE GENES IN hMSC ADIPOGENESIS THROUGH
RNAi SCREENING TECHNIQUES.

5.1 Introduction and aims

RNA interference is an endogenous cellular mechanism that can be exploited to silence the expression of a target gene at the mRNA level through the introduction of double-stranded RNA molecules targeting that gene (Fire *et al* 1998, section 1.7.3.1). Initial work in *C. elegans* used long dsRNAs to induce gene silencing, but such molecules trigger an interferon response in mammalian cells (Williams and Haque 1997, section 1.7.3.2). The technique was adapted for use in mammalian cells via the use of short, 21-23nt dsRNA oligonucleotides, named siRNAs, which mimic intermediates of the endogenous RNAi pathway, and thus mediate gene silencing (Fig. 1.5, Elbashir *et al* 2001). siRNAs bring about only a transient effect on gene expression. Hence, methods for stable expression of siRNA in mammalian cells were developed involving vectors from which 19-29nt shRNAs could be stably expressed (Brummelkamp *et al* 2002, Paddison *et al* 2002, 1.7.3.2).

The use of RNAi to identify genes involved in stem cell differentiation has several advantages, as there is a clear loss-of-function phenotype, and stably-expressed shRNAs can be used to ensure a constant level of gene knock-down throughout the relatively long time-course of differentiation. Indeed, several categories of viral vector have been used to stably deliver shRNA into a variety of cell types, including stem cells. Lentiviral vectors have been used to facilitate shRNA expression in ESCs (Rubinson *et al* 2003, Clements *et al* 2006), HSCs (Scherr *et al* 2003) and hMSCs (Cho *et al* 2006, Clements *et al* 2006), and have the advantage that they can infect non-dividing as well as dividing cells. Although adenoviral vectors are also used for delivering shRNA into many cell types, they have been less successful at delivering shRNA into stem cells, perhaps because the primary receptor for adenoviruses is expressed poorly on such cells (Zou and Yoder 2005). A wide range of retroviral vectors are available for the delivery of shRNA into cells, and their relatively small size allows for ease of manipulation (Brummelkamp *et al* 2002, Paddison *et al* 2002).

In the developing era of functional genomics, RNAi is becoming a widely used tool to investigate the function of multiple genes simultaneously. Although other functional

genomics techniques such as microarrays identify transcriptional changes that may indicate a role for a gene in a specific process, the use of techniques such as RNAi is required in order to functionally confirm such a hypothesis. This has the potential to yield much detail regarding a particular process such as stem cell differentiation. A major advance in this field was the production of RNAi libraries, in which hundreds or thousands of siRNAs or shRNAs are designed to target individual genes (1.7.3.3). Large-scale RNAi libraries targeting a broad spectrum of genes (Berns *et al* 2004, Paddison *et al* 2004) as well as smaller libraries focusing on specific processes or functions (Aza-Blanc *et al* 2003) have been developed, and are thus versatile in allowing broad studies, or targeted screens focusing on a selection of genes, to be performed. Moreover, novel genes involved in numerous processes including fat storage in *C. elegans* (Ashrafi *et al* 2003), apoptosis (Aza-Blanc *et al* 2003) and the p53 pathway (Berns *et al* 2004) have been identified, highlighting the efficacy of these RNAi libraries.

RNAi libraries can be used in several formats, including selection-based screening and systematic screening. Selection-based screens involve transduction of a cell population with a shRNA pool targeting multiple genes. A selection is then applied according to the function or process under investigation, such that shRNA-mediated silencing permits preferential selection of cells with a new phenotype. This technique has most often been used when the shRNA of interest confers a growth advantage under selection, but fluorescent reporters can also be used (Echeverri and Perrimon 2006). Identification of shRNAs conferring this new phenotype can be accomplished through the use of molecular barcodes (1.7.3.3, Westbrook *et al* 2005, Cullen *et al* 2005). However, this approach is limited in the range of functions that can be studied, and it has been reported than the simultaneous silencing of multiple genes, as could be the case in a pooled screen, results in weakened silencing of each gene (Echeverri and Perrimon 2006). Systematic screens entail high-throughput transduction of cells with individual siRNA/shRNAs, followed by assay for a particular function, which may involve a fluorescent reporter system. This method is more experimentally intensive than a selection-based screen (depending on the number of genes being studied), but is a more direct approach as any shRNAs producing a phenotype are immediately known (Brummelkamp *et al* 2003, Cullen and Arndt 2005).

In 2004, Bernards and co-workers documented the production of an RNAi library (the NKi library) targeting 7914 human genes (Berns *et al* 2004). Three shRNAs were designed per gene, and when a pool of the 3 shRNAs was employed, an average inhibition of 70% was achieved for 70% of all genes (Berns *et al* 2004). The NKi RNAi library is designed such that a 19nt sequence targeting a specific gene is separated from it's antisense version by a specific hairpin loop sequence. This cassette is under the control of the H1 promoter, and resides in a retroviral vector with puromycin resistance. The shRNAs in this library were designed using strict criteria to maximise knock-down efficiency and minimise off-target effect (Berns *et al* 2004). Design criteria included that shRNAs must share minimal sequence identity with other genes in order to reduce off-target effects, and they must target every transcript variant of a specific gene. The 3 shRNAs targeting each gene are cloned separately into this vector, and are subsequently pooled to form a mixed plasmid population (Fig 5.1A). This library has proved successful in a number of screens (Berns *et al* 2004, Kolfshoten *et al* 2005).

Aims

The main aim of these experiments was to implement a high-throughput shRNA screen of candidate genes and assess the effect of shRNAs on target gene expression and hMSC adipogenesis. Specific aims required to achieve this are:

- to isolate shRNAs of interest from the NKi library and develop strategies to deliver shRNA to stem cells using retroviral vectors

- to perform loss-of-function screens, in different formats, and assess the effects of shRNA expression on target gene expression and adipogenesis

- finally, to develop a lentiviral strategy to transduce hMSCs with custom-designed shRNAs for a smaller number of the most interesting candidates and assess the effect on gene expression and differentiation.

5.2 Development and optimisation of a technique for retroviral delivery of shRNAs isolated from an RNAi library

Microarray profiling of early hMSC adipogenesis identified 108 candidate novel regulators (Table 4.4). In order to assess the role of such a large number of candidates, the use of the NKi retroviral RNAi library was adopted. The aim of this experimental section was to develop methods to use shRNAs from this library in a functional screen.

shRNAs targeting 77% (83) of the candidate adipogenic regulators were present in the NKi library (Table 4.4). It was necessary to select 21 of these candidates for investigation, due to technical difficulties with the shRNA library discussed below. Table 5.1 therefore lists the most interesting 21 candidate genes with shRNAs present in the library, chosen according to their annotated functions. PPARγ was also targeted in this library, and hence could possibly be used as a positive control (if the shRNA was functional), as knock-down of this major adipogenic regulator has been shown to prevent adipogenesis in 3T3L1s (Hosono *et al* 2005).

5.2.1 shRNA isolation method I

The library is stored as bacterial glycerol stocks with a single well of a 384-well plate containing bacteria transformed with all 3 shRNAs for a specific gene (Fig. 5.1A). The initial method for isolating the shRNAs for each candidate gene (or gene of interest, GOI) was therefore to inoculate bacterial cultures directly from the library stocks and isolate a mixed population of plasmids containing shRNAs for each of the 83 GOIs (method I, Fig 5.2A). Plasmids containing shRNA inserts can be identified by a simple diagnostic digest using *Eco*RI and *Xho*I restriction enzymes (2.4.2.1, Fig. 5.1B), yielding a 6kb fragment representing the vector backbone, plus a ~320bp fragment containing the shRNA insert and H1 promoter (plasmids with no shRNA insert would show a ~260bp fragment). Examination of restriction digestion products revealed multiple band sizes in many samples (Fig, 5.2B, lane iii). These bands were visible for 34 out of 83 (42%) genes, and ranged in size from ~500bp to over 6kb (Fig. 5.2B and C). It is highly likely that these bands represent multiple recombination products; it was discovered that a high level of recombination occurs between the viral LTRs of this vector during replication in bacterial

Gene name	Symbol	Cluster	Interesting features/reasons for selection
Dapper homolog 1, antagonist of beta-catenin (xenopus)	DACT1	9	Wnt signalling inhibitor during development
FAT tumor suppressor homolog 1 (Drosophila)	FAT	10	Cell proliferation and communication during development
Homeo box A7	HOXA7	13	Thought to regulate gene expression and differentiation during development
Homeo box B2	HOXB2	9	Role in devlopment, neuronal patterning and differentiation
Insulin receptor tyrosine kinase substrate	IRTKS	6	Possible association with insluin signalling
Jumonji homolog (mouse)	JMJ	6	Required for nerual tube formation and heart development in mice, anti-proliferative role
Lipoma HMGIC fusion partner-like 2	LHFPL2	6	Possible association with lipoma
Mesenchyme homeo box 2 (growth arrest-specific homeo box)	MEOX2	9	Mesoderm induction, KO mouse has muscular/skeletal abnormalities
Mesoderm development candidate 1	MESDC1	6	Possible association with mesoderm development
Serum/glucocorticoid regulated kinase	SGK	10	Regulated by glucocorticoids, TGFβ signalling and glucose levels
Sine oculis homeobox homolog 2 (Drosophila)	SIX2	5	Associated with ligament and tendon differentiation, maintainance of differentiated state
Snail homolog 1 (Drosophila)	SNAI1	6	Represses ectodermal genes within the embryonic mesoderm
Snail homolog 2 (Drosophila)	SNAI2	5	Epithelial-mesenchymal transition; expressed in emerging mesodermal cells
Sprouty homolog 2 (Drosophila)	SPRY2	8	Negative regulator of bFGF signalling
Slingshot 1	SSH1	6	Activated downstream of PI3K, role in insulin-induced membrane protrusion and actin dynamics
T-box 2	TBX2	5	Limb pattern formation and transcritpitonal regulation of genes resuired for mesoderm
Transcription factor 8 (represses interleukin 2 expression)	TCF8	6	Resides at a major susceptibility locus for obesity, regulates SMAD proteins
Transforming growth factor, beta receptor III (betaglycan, 300kDa)	TGFBR3	2	TGFBR3-specific antisera inhibit mesenchyme formation
TGFB-induced factor (TALE family homeobox)	TGIF	10	Forms transcriptional repressor complex with SMAD2/4
Transducer of ERBB2, 1	TOB1	6	Regulates BMP-SMAD activity
Zinc finger protein, multitype 2	ZFPM2	6	Controls expression of GATA proteins, specifically GATA4 during cardiac development
Peroxisome proliferator-activated receptor, gamma	PPARγ	3	Major positive adipogenic regulator - possible positive control

Table 5.1. Most interesting 21 out of the 83 candidate adipogenic regulators (plus PPARγ) for which shRNAs were designed in the NKi shRNA library, and reasons for their selection as defined by interesting functional attributes.
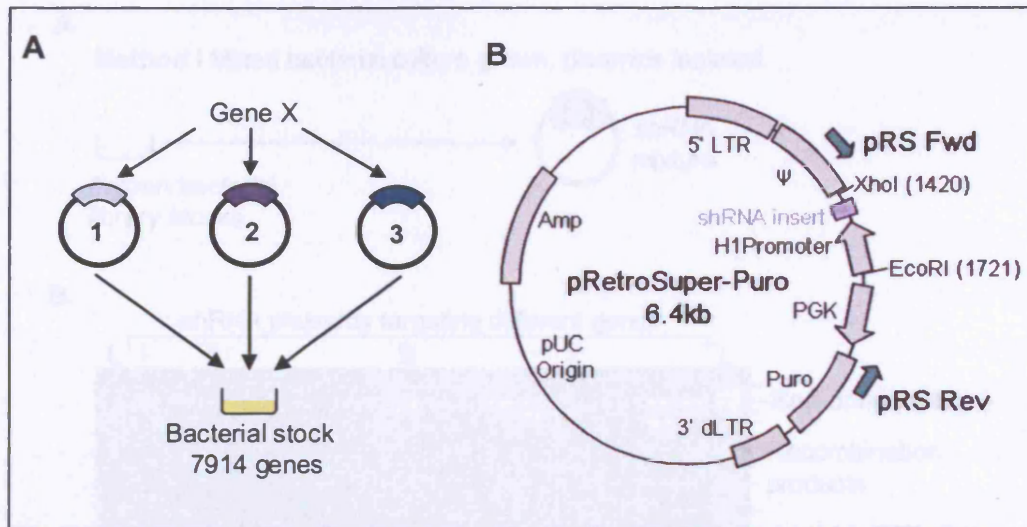
Figure 5.1. Schematic showing NKi RNAi library design and retroviral vector. A, Design of NKi library. Three shRNAs were designed per gene and cloned individually into pRetroSuper-Puro. Plasmids containing individual shRNAs were mixed and stored as bacterial glycerol stocks, with one well per gene. The library represents 7914 human genes. B. pRetroSuper-Puro. Retroviral vector, based on pMSCV retroviruses with the addition of the shRNA expression cassette. shRNA expression is driven by the H1 promoter, and ampicillin and puromycin resistance ORFs allow for drug selection in bacterial and mammalian cells, respectively. pRS Fwd and pRS Rev primers can be used to amplify the region containing the shRNA and H1 promoter.

Method I Mixed bacterial culture grown, plasmids isolated

Frozen bacterial library stocks → shRNA mixture

B

shRNA plasmids targeting different genes

Backbone (~6kb)
Recombination products
Insert (~320bp)

C

Outcome of direct isolation (method I)	Number of GOIs with this outcome (/83)
Insert present, no recombination	16 (19%)
Recombination (insert present/absent)	34 (42%)
Low plasmid yield	26 (31%)
No bacterial growth from library	7 (8%)

Figure 5.2. Outcome of shRNA library screening method I. A, Bacterial cultures were inoculated directly from bacterial glycerol stocks of the library and the three shRNA vector constructs (plasmids) for each gene were isolated as a mixed plasmid population (2.4.2.6); B, Representative agarose gel separation of *Eco*RI/*Xho*I restriction digestions of mixed plasmid populations for different GOIs (2.4.2.1). This technique was used to identify GOIs for which the shRNA insert was present (320bp band). Multiple bands were observed for several constructs as a result of recombination inherent in the library. L, 1kb Plus DNA ladder (Invitrogen); (i) Vector backbone and insert bands of correct size; (ii) No/low plasmid yield: possible contamination of library with ampicillin-resistant bacteria; (iii) Some/multiple recombination products; C, Summary of the outcome of method I.

cultures, thus resulting in the elimination of shRNA inserts (personal communication from NKi library distributors). Some recombined plasmids contained the shRNA insert (for example, Fig. 5.2B, lane iii), whereas the majority did not, suggesting that as suspected, recombination was occurring and resulting in loss of the shRNA insert. A further 26 (31%) genes gave a low plasmid yield even after two separate plasmid isolations, which was reflected by faint bands or no bands after digestion, indicating that there may have been some contamination in the library with ampicillin-resistant bacteria, or that a portion of the vector containing only the ampicillin ORF was present (Fig. 5.2B, lane ii). Bacterial cultures for 7 (8%) of genes could not be grown. Finally, 16 (19%) GOIs contained no recombination and a reasonable to high level of shRNA insert (Fig. 5.2B, lane i). It was clear from these observations that for the majority of genes, direct shRNA isolation would yield too little un-recombined shRNA vectors for use in the screening strategy.

5.2.2 shRNA isolation method II

A PCR screening method was therefore developed in an attempt to identify individual bacterial colonies that contained plasmids encoding shRNAs (i.e. plasmids in which the shRNA had not been excluded by recombination) (Fig.5.3A). As stated above, it was not possible to use this more detailed strategy to identify shRNAs for all 83 GOIs represented in the RNAi library, so further experiments were limited to the top 21 candidates (Table 5.1), plus PPARγ. These candidates were chosen due to their interesting functions, rather than because shRNA insert bands were identified in method I, so it was possible that some of these shRNAs contained recombination. To perform PCR screens, bacterial clones were isolated for each gene via streaking of library stocks on LB-agar plates containing 100μg/ml ampicillin. Ten bacterial colonies per gene were then suspended in dH₂O and were used as template DNA for PCR with the pRS primers (Fig. 5.1B) for the identification of positive bacterial clones, i.e. clones containing shRNA-encoding plasmids (Fig. 5.3B i, 2.6.3.1). Positive clones were identified by a ~640bp band after PCR amplification and separation of PCR fragments on agarose gels (2.6.3.3). The presence of shRNA inserts in plasmids purified from bacterial colonies identified as positive in the PCR screen was confirmed by EcoRI/XhoI restriction digestion of these plasmids (Fig. 5.3B ii). Plasmids containing shRNAs were then sequenced with the pRS Seq primer (Table 2.3, 2.4.2.7) to identify the 3 different shRNAs for each gene. Using the

203
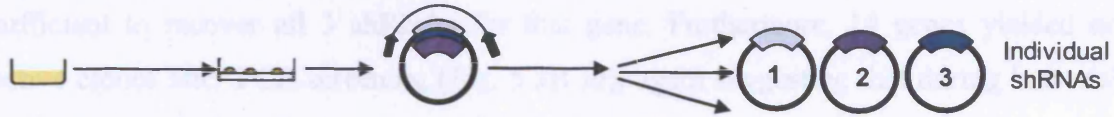
Figure 5.3. Outcome of shRNA library screening method II. A, Library bacterial stocks for each top 22 GOI (inclusive of PPARγ) were streaked onto LB-agar + ampicillin (100μg/ml) plates to isolate bacterial clones. Bacteria were used directly as a PCR template with the pRS primers to identify positive clones for each gene, indicated by a 640bp PCR fragment (2.6.3.1); B, (i), An acceptable number of positive clones were identified for some genes; (ii) these were amplified via overnight culture in LB + ampicillin (100μg/ml), then plasmids were isolated (2.4.2.6) and confirmed as positive with EcoRI/XhoI restriction digestion. Plasmids were sequenced to recover each of the 3 shRNAs for that gene; (iii), Some genes yielded no positive clones in the PCR screen. The constructs for these genes were then sub-cloned according to method III. L, 1kb Plus DNA ladder; C, Outcome of method II.

204

PCR screening method, sufficient (5-10) positive clones were identified for 6 of the 22 genes (Fig. 5.3C). For two genes, 1-4 positives were isolated, but this was deemed to be insufficient to recover all 3 shRNAs for that gene. Furthermore, 14 genes yielded no positive clones after PCR screening (Fig. 5.3B iii), again suggesting that during bacterial amplification of the library for distribution, shRNA-positive clones were lost by recombination, or alternatively that contamination with ampicillin-resistant bacteria was present. Thus, this method resulted in the isolation of shRNA inserts for 6 genes. These genes were, on the whole, those shown to have visible insert bands in method I. It was therefore likely that the shRNA inserts for other genes were present at such low levels that a further method was required to recover them.

5.2.3 shRNA isolation method III

Method III involved utilising a PCR cloning strategy. Bacterial library stocks for each gene for which no, or few, shRNAs were isolated in method II, were used as a PCR template (either directly from the frozen library stocks, or after overnight culture, 2.4.4.1) to amplify the region containing the shRNA expression cassette (containing the shRNA insert and H1 promoter, 2.4.4.1, Fig. 5.1B). *EcoRI/XhoI* restriction digestion of the PCR fragment was used to excise the shRNA expression cassette and these fragments were subsequently ligated back into an empty retroviral vector containing the same restriction sites. Following sub-cloning, method II was repeated to screen for bacterial clones containing plasmids in which the shRNA had ligated, in order to recover individual shRNAs (Fig. 5.4A). Initially, a version of pRetroSuper-Puro which contained no shRNA insert (Fig. 5.4D lanes 1 and 2) was used that had been isolated from the RNAi library. The likely absence of an endogenous shRNA in this vector precluded inadvertent re-ligation of endogenous shRNA rather than target shRNA.

As a result of isolation methods II and III, shRNAs were recovered for 19 out of the 22 GOIs; for the remaining 3 genes, it was not possible to amplify the bacterial stocks (Table 5.2). When the shRNA-positive plasmids were sequenced (using the pRS Seq primer, 2.4.2.7) and BLAST searches were performed it was revealed that for 7 of the 19 genes, shRNAs targeting other non-candidate genes had been isolated; on occasion shRNAs for 2 non-candidate genes were identified, making it unlikely that the wrong well had been
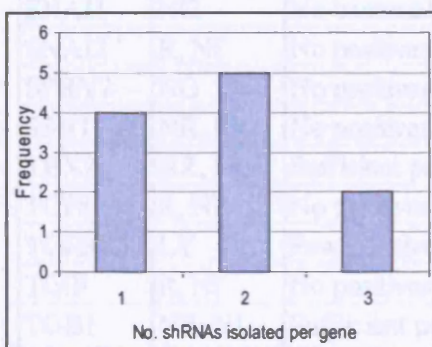
A

Method III PCR amplify shRNAs from bacterial stocks, subclone and isolate bacterial clones, repeat method II
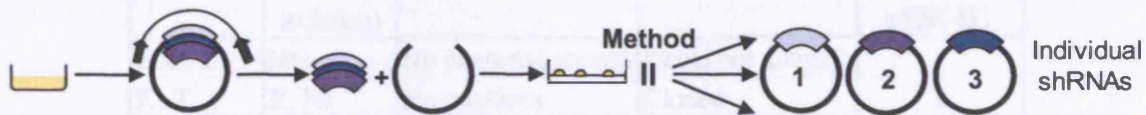
B

Outcome of sub-cloning (method III)	No. GOIs with this outcome (/16 with few/no +ves from method II)
Could not amplify insert	3
Contaminating shRNAs isolated	5
shRNA for GOI isolated but mutated	1
shRNA for GOI isolated	7

C

D
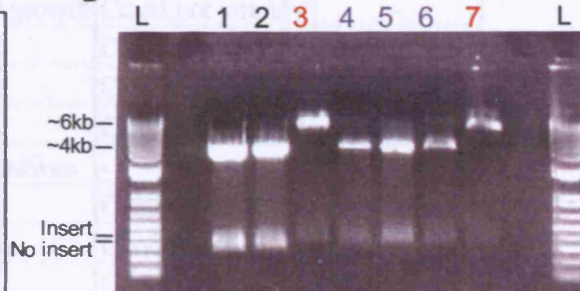
Figure 5.4. Outcome of shRNA library screening method III. A, Library bacterial stocks were used directly as a PCR template to amplify mixed shRNA inserts. These were sub-cloned back into pRetroSuper-Puro (2.4.4.1) and individual shRNAs isolated as in method II; B, Outcome of shRNA isolation using methods II and III. shRNAs for GOIs were isolated via PCR screening (method II), sub-cloning (method III), or both; C, Number of shRNAs recovered for each GOI; D, The pRetroSuper-Puro vector backbone used for sub-cloning had recombined. Two preps of empty vector (lanes 1 and 2) plus representative shRNA plasmids isolated by PCR screening (method II, lanes 3 and 7), or sub-cloning (method III, lanes 4-6) were digested with *Eco*RI and *Xho*I restriction enzymes (2.4.2.1). L, 1kb Plus DNA ladder.

| Gene name | Outcome of library method: | | | Number of shRNAs cloned into pSilR-H1 |
	I (Direct isolation)	II (PCR screening)	III (Sub-cloning)	
DACT1	NG	No bacterial growth	Could not amplify	-
FAT	R, NI	No positives	Cloned	2
HOXA7	LY	Sufficient positives	-	2
HOXB2	R, NI	No positives	Cloned	2
IRTKS	R, NI	No bacterial growth	Could not amplify	-
JMJ	R, NI	No positives	Contamination	-
LHFPL2	NR, I	Sufficient positives	-	3
MEOX2	NR, I	Sufficient positives	-	1
MESDC1	NG	No bacterial growth	Contamination	-
SGK	R, I	Sufficient positives (contamination)	-	-
SIX2	R, NI	No positives	Cloned	2
SNAI1	NG	No bacterial growth	Could not amplify	-
SNAI2	R, NI	No positives	Cloned	1
SPRY2	NG	No positives	Contamination	-
SSH1	NR, I	No positives	Cloned - mutation	-
TBX2	NR, I	Sufficient positives	-	1
TCF8	R, NI	No positives	Cloned	3
TGFBR3	LY	Few positives	Contamination	-
TGIF	R, NI	No positives	Cloned	2
TOB1	NR, NI	Suffiicient positives (contamination)	-	-
ZFPM2	R, NI	Few positives	Contamination	-
PPARg	-	-	Cloned	1

Table 5.2. Summary of the outcome of shRNA isolation strategies for the 22 GOIs from the shRNA library. NG, No growth; R, recombination; NR, no recombination; I insert band visible; NI, no insert band visible; LY low plasmid yield. Each shRNA was cloned into pSilencer H1-5.0 Retro.

picked in the library glycerol stocks (Table 5.2). These contaminating genes were found to reside in wells adjacent to the GOI in the original library stocks, and it was likely shRNA levels for GOIs were so low in these cases that they may have been present at the same levels as contaminating genes, making isolation of contaminating shRNAs likely. Furthermore, although a single shRNA for one GOI was isolated, it was found to contain a mutation rendering it non-functional (Table 5.2). Thus, as a combined result of shRNA recovery methods II and III, shRNAs for 11 out of the 22 GOIs (including PPARγ) were isolated (Fig. 5.4B, 5.3C, Table 5.2). The number of shRNAs isolated per gene varied from 1-3, with most genes represented by 2 shRNAs (Fig. 5.4C). It is likely that certain shRNAs for a specific gene were overrepresented in the library compared with other shRNAs for that gene, due to bacterial growth reducing the complexity of the shRNA pool (Downward 2004). This would occur when bacteria selectively amplify certain shRNA inserts over others, possibly due a specific shRNA having a deleterious effect on bacterial growth or being more difficult to amplify, eventually resulting in the loss of a specific shRNA from the bacterial pool. This could account for the isolation of all 3 shRNAs for only 2 genes. Indeed, although sequencing was performed on 5-10 positive clones per gene, in 4 cases these all represented a single shRNA sequence.

During this process, it was unfortunately found that the pRetroSuper-Puro backbone used for sub-cloning had recombined; restriction digestions of bacterial clones identified from PCR screens (method II) showed a normal 6kb backbone plus insert band (Fig. 5.4D, lanes 3 and 7), whereas the vector stock plus shRNA plasmids isolated via sub-cloning (method III) revealed a backbone lacking around 2kb (Fig 3.4D, lanes 1, 2, 4-6). In preliminary experiments it was indeed found that hMSCs transduced with shRNAs recovered using the PCR screening method (method II) could survive puromycin selection, whereas shRNAs that had been sub-cloned into the (recombined) pRetroSuper-Puro backbone did not confer puromycin resistance to hMSCs so were non-functional (data not shown). However, this work had resulted in the isolation of multiple constructs containing specific shRNAs that could be easily sub-cloned into further vectors.

pSuperRetro-GFP-Neo (pSR-GFP/Neo, OligoEngine) was chosen for sub-cloning as it contained a selection cassette resulting in a Neo^R-GFP fusion protein (Fig.5.5A), so
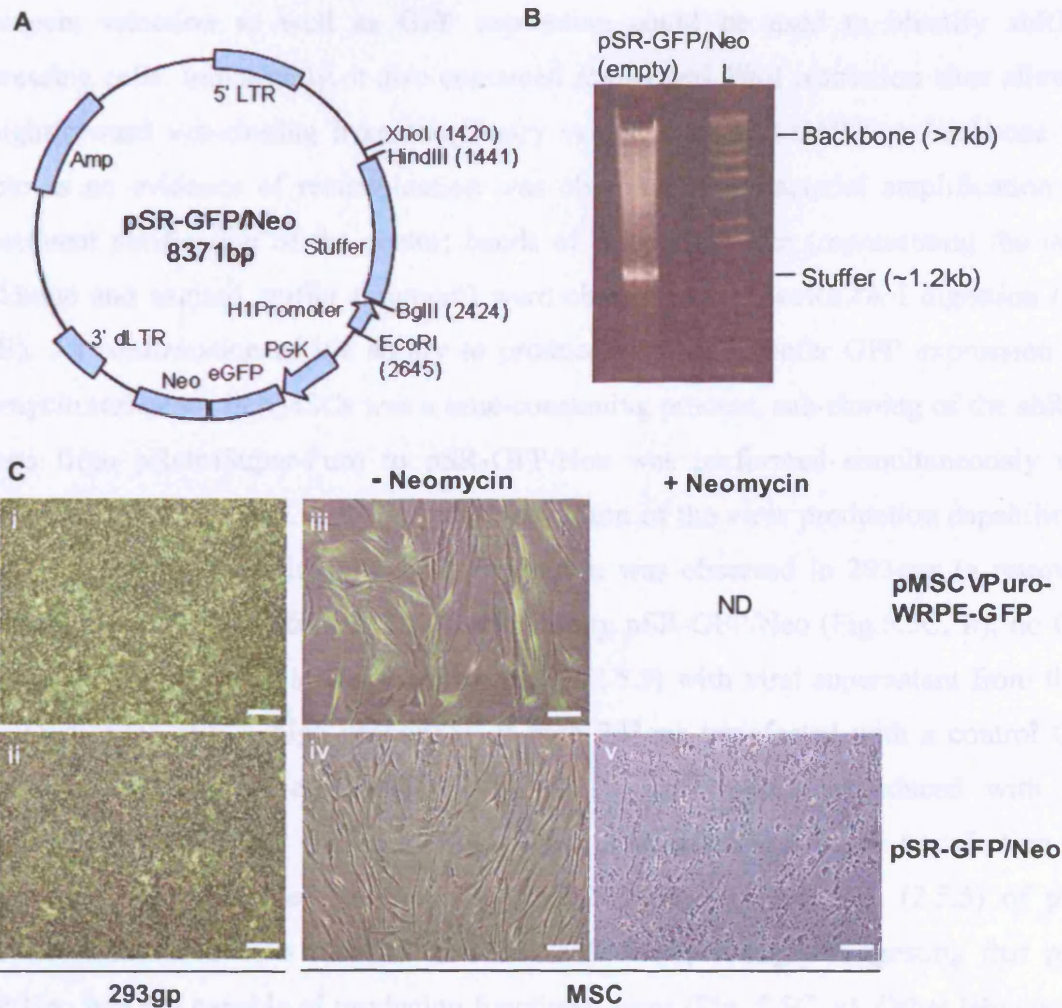
Figure 5.5. pSR-GFP/Neo does not recombine, and may not produce competent virus. A, Schematic of pSuperRetro-Neo-GFP; B, *Eco*RI/*Xho*I restriction digest (2.4.2.1) of empty pSR-GFP/Neo isolated after bacterial amplification yielded 1kb stuffer band and ~7.3kb backbone of expected size, with no recombination products. L, 1kb Plus DNA ladder; C, pSR-GFP/Neo does not produce functional virus. The 293gp viral packaging cell line was co-transfected with 2μg control vector pMSCVPuro-WRPE-GFP (i) or empty pSR-GFP/Neo (ii) plus 2μg VSVG helper plasmid (2.5.1). Viral supernatants were harvested 48h after media change, and used at a 1:2 dilution (supplemented with 4μg/ml polybrene) to transduce hMSCs (2.5.5). This was repeated to increase transduction rates. Two days after the 2nd transduction, cells expressing GFP were visible in control vector-transduced cells (iii), but pSR-GFP/Neo transduction produced no visible GFP +ve cells (iv). Neomycin (500μg/ml) selection of pSR-GFP/Neo-transduced hMSCs was carried out for 5d (2.5.5). pSR-GFP/Neo-transduced cells did not survive selection (v). Scale bars represent 200μm. ND, not determined.

neomycin selection as well as GFP expression could be used to identify shRNA-expressing cells. Importantly, it also contained *Eco*RI and *Xho*I restriction sites allowing straightforward sub-cloning from the library vector. The pSR-GFP/Neo backbone was stable as no evidence of recombination was observed after bacterial amplification and subsequent purification of the vector; bands of the correct size (representing the intact backbone and excised stuffer fragment) were obtained after *Eco*RI/*Xho*I digestion (Fig. 5.5B). As confirmation of it's ability to produce virus and confer GFP expression and neomycin resistance in hMSCs was a time-consuming process, sub-cloning of the shRNA inserts from pRetroSuper-Puro to pSR-GFP/Neo was performed simultaneously with vector validation (2.4.4.2). Unexpectedly, validation of the virus production capability of the vector showed that although GFP expression was observed in 293gps (a retroviral packaging cell line) transfected (2.5.1) with empty pSR-GFP/Neo (Fig.5.5C, ii), no GFP expression was observed in hMSCs transduced (2.5.3) with viral supernatant from these cells (Fig. 5.5C, iv). A high proportion of both 293gps transfected with a control GFP construct (pMSCV-Puro-eGFP-WPRE, Fig. 2.1), and hMSCs transduced with this construct, expressed GFP (Fig. 5.5B, i and iii), demonstrating that the transfection and transduction methodologies had been successful. Neomycin selection (2.5.5) of pSR-GFP/Neo-transduced cells resulted in 100% cell death, strongly suggesting that pSR-GFP/Neo was not capable of producing functional virus (Fig. 5.5C, v). Other laboratories had indeed reported similar observations to the company producing this vector (OligoEngine, personal communication). Therefore, pSR-GFP/Neo could not be used for the shRNA screening experiment.

Due to the above problem a third vector was selected. pSilencer H1-5.0 Retro (pSilR-H1) is analogous to the RNAi library vector, as it contains a puromycin resistance cassette, and the presence of *Eco*RI and *Xho*I sites again allowed for use of a simple sub-cloning protocol (Fig. 5.6A). No evidence of recombination was observed within this vector upon bacterial amplification (Fig. 5.6B, lane i). Before sub-cloning shRNA inserts into this vector, its functionality was confirmed. MSCs were transduced with virus produced from pSilR-H1 or the pMSCV-Puro-eGFP-WPRE control vector and puromycin selection (2.5.5) killed the majority of untransduced cells (Fig. 5.6B, i). A stable population could be expanded from the surviving pSilR-H1-transduced cells (Fig. 5.6B, ii), indicating that
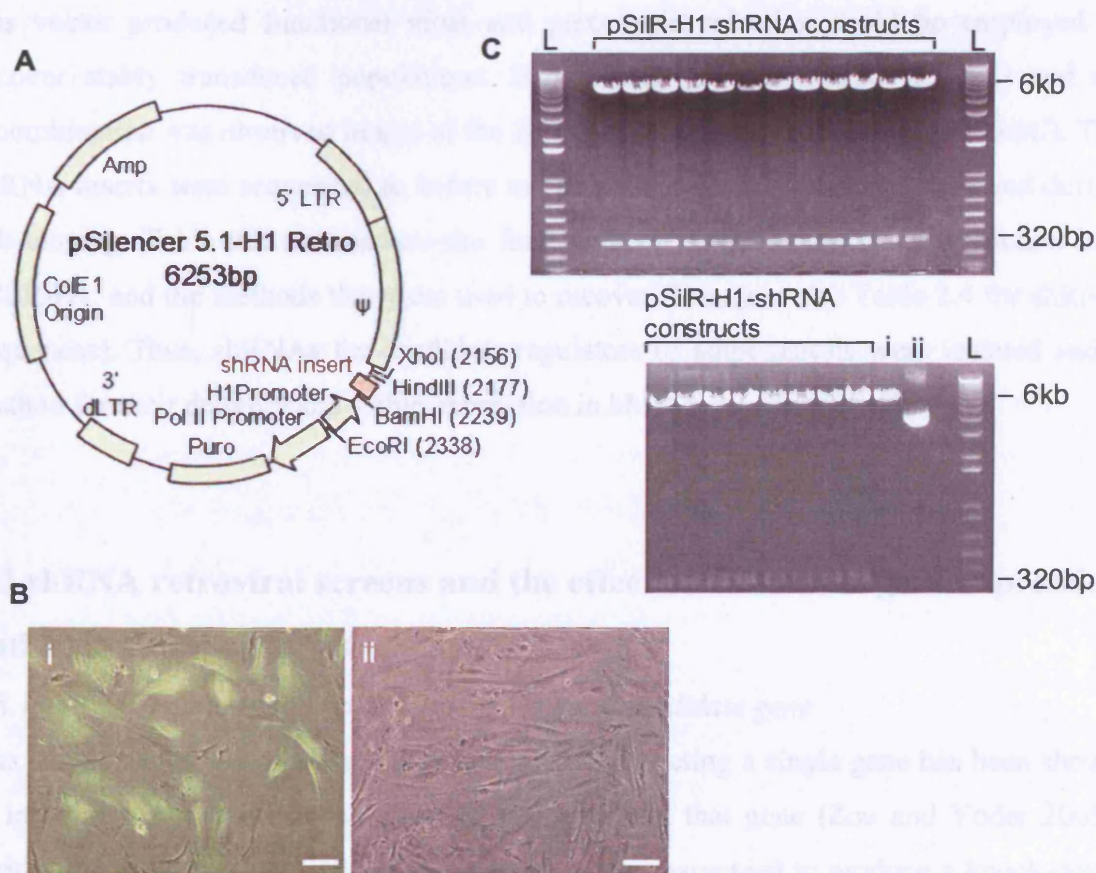
Figure 5.6. pSilR-H1 does not recombine and produces functional virus. A, Schematic of pSilR-H1; B, pSilR-H1 produces functional virus. hMSCs were transduced with viral supernatant produced from 293gps co-transfected with 2μg pMSCVPuro-WRPE-GFP control vector (i) or 2μg pSilR-H1 (ii) plus 2μg VSVG helper plasmid (2.5.1 and 2.5.3). Transduction was performed in duplicate, and 2d after the 2nd transduction, selection with puromycin (750ng/ml) was initiated (2.5.5). Selection killed the majority of non-transduced cells after 4 days, as assessed by GFP expression in control vector-transduced cells (i), and resulted in a surviving population of pSilR-H1-transduced cells (ii). Scale bars represent 200μm; C, pSilR-H1 does not recombine. *Eco*RI/*Xho*I restriction digest (2.4.2.1) of empty vector after bacterial amplification (lane i; lane ii = uncut vector). All other lanes: *Eco*RI/*Xho*I restriction digest of pSilR-H1 containing all GOI shRNAs. Correct backbone and insert bands are present. L, 1kb Plus DNA ladder.

211

this vector produced functional virus and puromycin selection could be employed to recover stably transduced populations. Sub-cloning was completed (2.4.4.2) and no recombination was observed in any of the 20 shRNA-containing plasmids (Fig. 5.6C). The shRNA inserts were sequenced as before to ensure that no mutations had occurred during sub-cloning. Table 5.2 summarises the final number of shRNAs that were cloned into pSilR-H1, and the methods that were used to recover them (see also Table 2.4 for shRNA sequences). Thus, shRNAs for candidate regulators of adipogenesis were isolated and a method for their delivery and stable expression in hMSCs was developed.

5.3 shRNA retroviral screens and the effect on candidate gene expression and hMSC adipogenesis

5.3.1 shRNA screen using pooled shRNAs for each candidate gene

The simultaneous use of several different siRNAs targeting a single gene has been shown to increase the level of knock-down of expression of that gene (Zou and Yoder 2005). This could be because not all siRNA sequences are guaranteed to produce a knock-down (1.7.3.2), and it could be hypothesised that different siRNAs may act in synergy to produce a higher knock-down than if used separately. As more than one shRNA had been isolated for 7 out of the 11 GOIs (Fig. 5.4C), this technique was initially adopted for the loss-of-function screen in order to maximise any knock-down effects.

The viral packaging cell line 293gp was transfected with either pooled shRNAs (1:1 ratios of each shRNA) for those genes with more than one shRNA, or with single shRNAs, plus VSVG helper plasmid (2.5.1). The PPARγ shRNA was included in these experiments so that, if this shRNA was functional (which was not guaranteed), it could act as a positive control to demonstrate the functional consequences of silencing an adipogenic regulator on the ability of hMSC to differentiate to adipocytes. Additionally, a scrambled hairpin (Scr) with no significant sequence homology to any human gene (Ambion 2005) was included as a control to assess non-specific effects of hairpin introduction. The viral supernatant produced was filtered, diluted 1:2 with fresh media and applied to passage 5 hMSCs at 30-50% confluence with 4μg/ml polybrene (2.5.3). This process was repeated to

increase the percentage of transduced cells; retroviruses can only transduce dividing cells as they do not contain a nuclear localisation signal that allows them to pass through the nuclear membrane, so they can only access the genome when the nuclear membrane is absent during cell division. A GFP-expressing control virus (pMSCV-Puro-eGFP-WPRE) was also transduced into hMSCs as a control to monitor the success of each experimental stage, as it was assumed that transfection, transduction and drug selection efficiencies would be comparable between this control vector and all other vectors. After selection of transduced cells with puromycin for 3-4d (2.5.5), FACS analysis of hMSCs transduced with the GFP-expressing control virus (2.3.2) revealed a shift in the mean population fluorescence after drug selection (Fig. 5.7A), which indicated that puromycin selection had killed the majority of untransduced cells. Furthermore, although a proportion of cells expressed a high level of GFP, most transduced cells were expressing GFP at a low level. Cells were grown to confluence and then induced to differentiate into adipocytes using the classic induction regimen (2.2.2.1). RNA was also extracted at confluence (2.6.1) for use in RT-PCRs to assess knock-down effect, as it had been confirmed that all GOIs were expressed to some extent in undifferentiated hMSCs (Fig. 5.8A).

Differentiation was continued for 14 days then cells were fixed and stained with Oil Red O (2.2.3.1) to assess the extent of adipocyte formation (Fig. 5.7B and C). The average number of adipocytes per field of view ranged from 27-42 for cells transduced with a shRNA-containing vector (Fig. 5.7B). Non-transduced hMSCs yielded an average of 106 adipocytes per field of view, significantly more than cells containing the scrambled hairpin (p<0.001, using a two-sample Student's t-test), suggesting a possible effect of hairpin introduction or viral transduction on differentiation. However, there were no statistically significant differences in the extent of differentiation between the scrambled hairpin and any of the shRNAs targeting candidate genes (p>0.05). It was next determined whether any of the shRNAs had reduced the expression of their target genes at the mRNA level. cDNA was synthesised (2.6.2) from the RNA collected from confluent cells, and RT-PCR analysis using gene-specific primers (2.6.3.2, and Table 2.3 for primer sequences) revealed that for each GOI-shRNA, expression was similar to that in hMSCs transduced with the scrambled hairpin, but this was less that it's expression in untransduced hMSCs (Fig. 5.8B). Expression of GAPDH, a housekeeping gene that
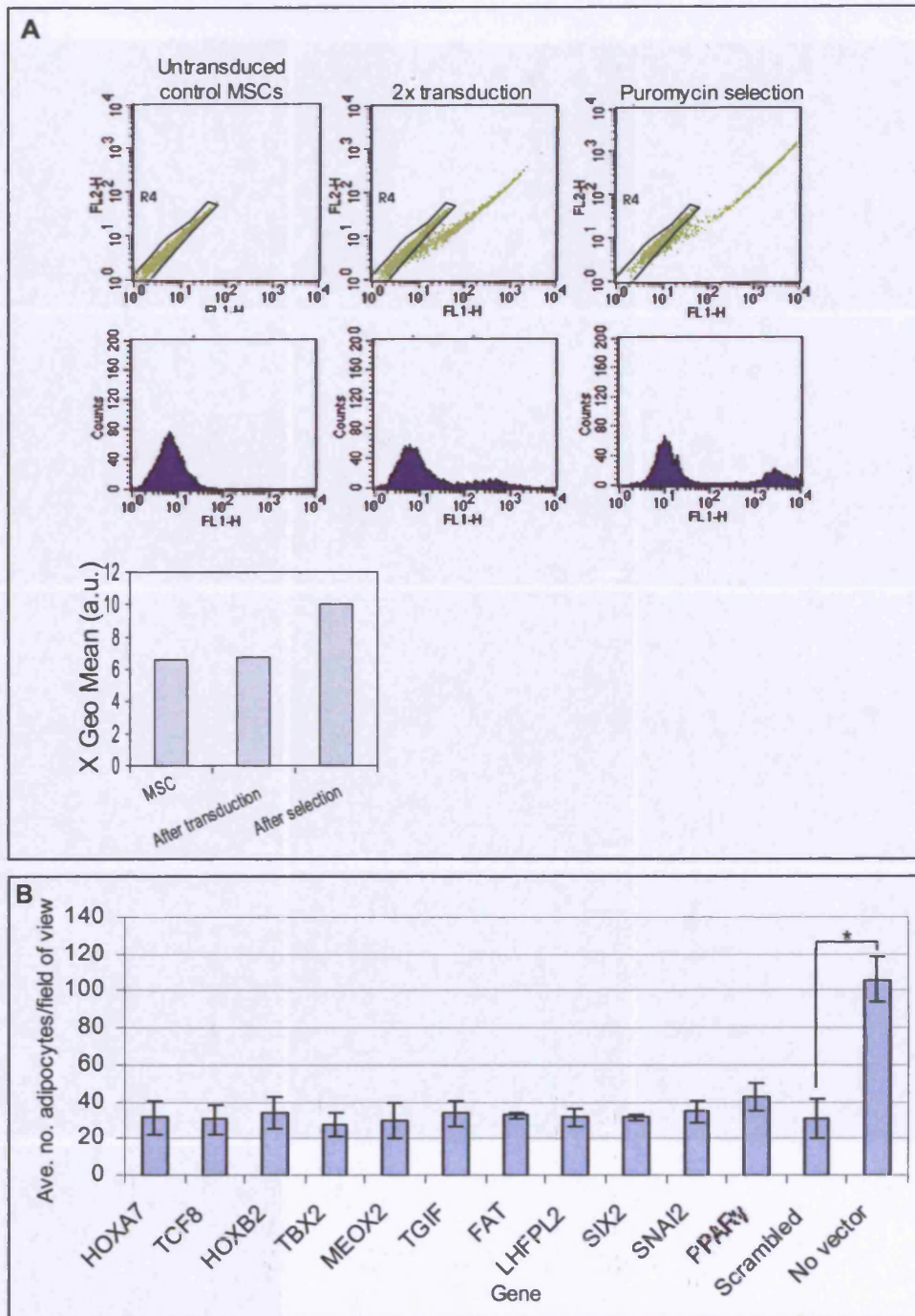
A

Untransduced control MSCs | 2x transduction | Puromycin selection

B

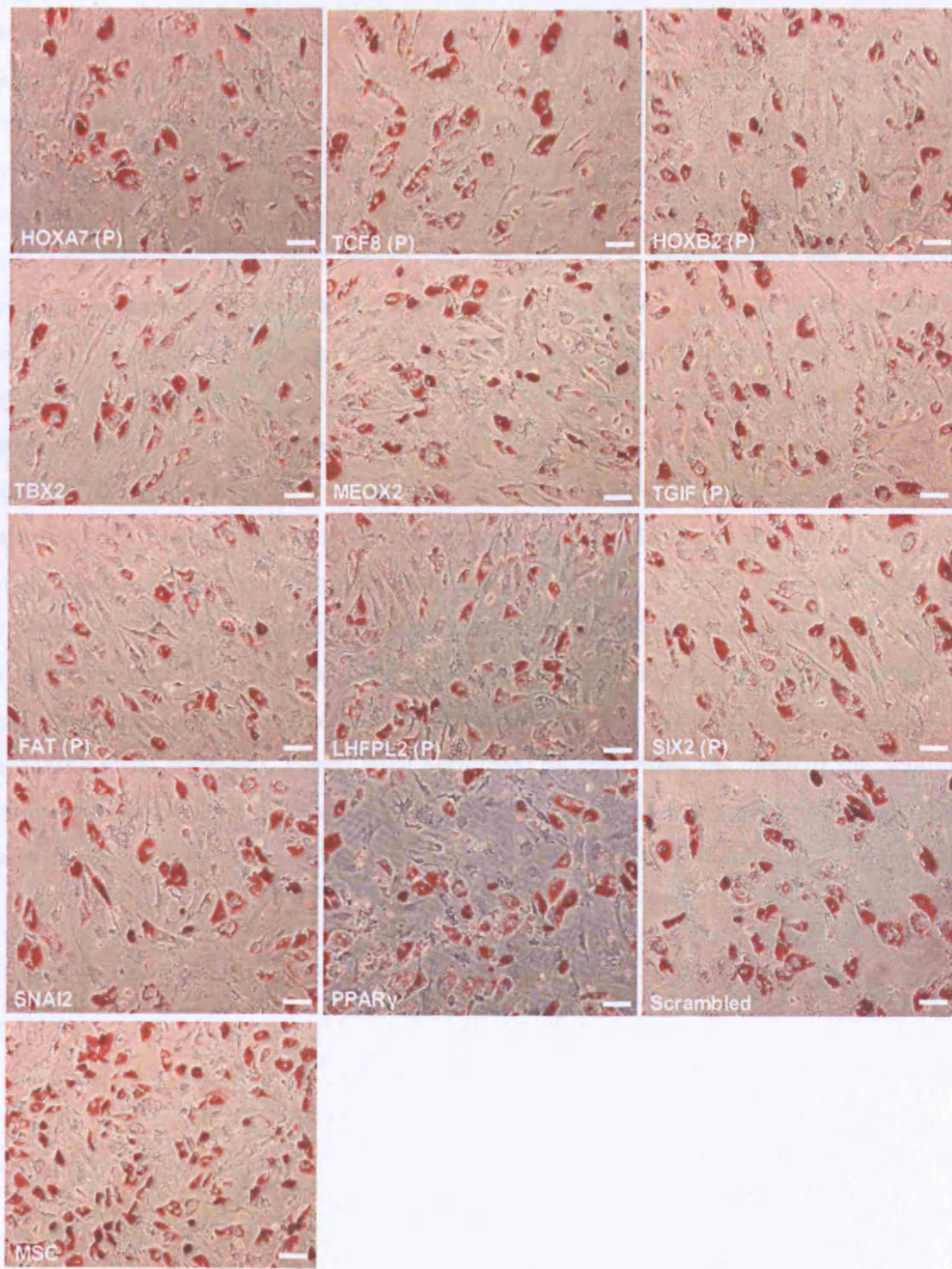Figure 5.7. Continued overleaf.

C

Figure 5.7. Continued overleaf.

215

Fig 5.7. Effect on adipogenesis of transduction with shRNA (pools) targeting candidate genes. hMSCs underwent 2 rounds of transduction (2.5.3) with viral particles produced from transfection of 293gps with pooled shRNAs for each GOI (where more than one shRNA had been isolated), plus scrambled shRNA and PPARγ control (2.5.1). Cells were treated with puromycin (750ng/ml for 4 days) to select for stably transduced cells (2.5.5), then grown to confluence and induced to differentiate (2.2.2.1); A, FACS analysis of hMSCs transduced with virus produced from pMSCVPuro-GFP-WRPE (2.3.2). Shown is the number of GFP+ events in un-transduced hMSCs, hMSCs transduced with the GFP virus prior to puromycin selection, and hMSCs transduced with the GFP virus, after puromycin selection. Upper graphs represent the distribution of events (green dots) on the FL1 (green) versus FL2 (red) filters, and the gate (R4) used to separate GFP- from GFP+ events. Lower histograms (purple-filled) represent the number of GFP+ events at different fluorescent intensities. The bar chart represents the mean population fluorescence in each condition; B and C, At day 14 of differentiation, cells were fixed and stained with Oil Red O (2.2.3.1). Experiments were performed in duplicate; B, Quantification of the effect on differentiation of shRNA expression in each condition. Data represent number of adipocytes per field of view and are mean ± SD. Four fields of view were counted per condition. *, p<0.001 using a two-sample t-test (assuming equal variance). T-tests were performed by comparing the mean of the Scrambled control sample to all other means; C, representative image showing extent of differentiation for each shRNA pool, or individual shRNA. Scale bars represent 200μm.
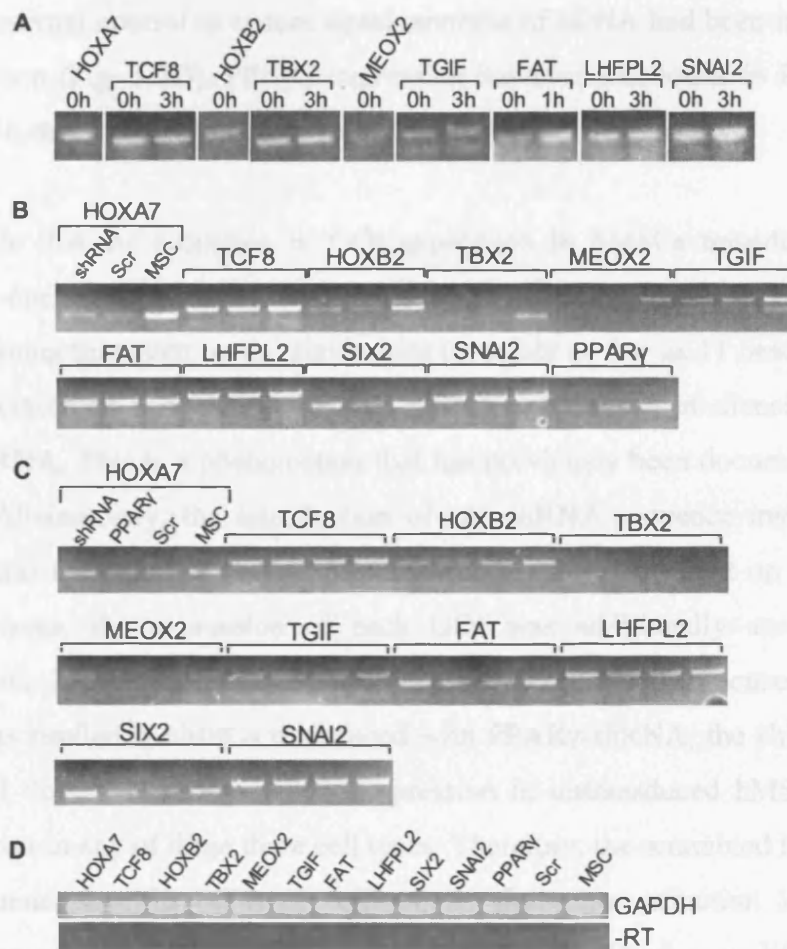
Fig 5.8. Effect of (pooled) shRNA screen on candidate gene expression in hMSCs. RNA was extracted from confluent hMSCs stably expressing shRNA(s) for each GOI (2.6.1). cDNA was synthesised (2.6.2) and used as a PCR template with gene-specific primers to investigate gene expression levels (2.6.3.2); A, the timepoint at which gene expression knock-downs should be investigated was determined by confirming that confluent hMSCs (0h) expressed each GOI to some degree. Additional timepoints during differentiation (1h or 3h, as indicated) were investigated for some genes. SIX2 and PPARγ expression in confluent hMSCs had already been confirmed (Fig. 4.3); B, the expression of each GOI was compared in 3 conditions: hMSCs expressing shRNA targeting the GOI (shRNA), scrambled hairpin (Scr) or untransduced hMSCs (MSC); C, The expression of each GOI in PPARγ-shRNA cells was assessed as an additional control to examine whether the reduced GOI expression in Scr-shRNA cells was due to a transcriptional response specifically to the Scr 19nt sequence; D, GAPDH internal control and no reverse transcription controls (using adipsin primers) for each condition. PCR cycle numbers: GAPDH: 25; all other genes: 28.

217

served as an internal control to ensure equal amounts of cDNA had been used, was equal in each condition (Fig. 5.8D). PPARγ expression however was lower in PPARγ-shRNA transduced cells that in Scr-shRNA transduced cells (Fig. 5.8B).

It was possible that the reduction in GOI expression in hMSCs transduced with Scr-shRNA was due to sequence-specific, off-target effects of the scrambled shRNA sequence, meaning that even partial similarities (possibly as few as 11 bases) in sequence between the scrambled shRNA and the GOI mRNA could result in silencing of the GOI by the Scr-shRNA. This is a phenomenon that has previously been documented (Jackson *et al* 2003). Alternatively, the introduction of any shRNA sequence into hMSCs may induce a general reduction in gene expression (that is not dependent on sequence). To address this issue, the expression of each GOI was additionally assessed in cells transduced with PPARγ-shRNA. It was observed that, in many cases, target gene expression was similar in hMSCs transduced with PPARγ-shRNA, the shRNA targeting that gene, and Scr-shRNA (Fig. 5.8C). Expression in untransduced hMSCs was again often higher than in any of these three cell types. Therefore, the scrambled hairpin was not inducing sequence-specific off-target effects, as the same reduction in target gene expression was seen in cells containing the PPARγ-shRNA (i.e. a different shRNA sequence). Instead, it is possible that a more general response to viral transduction or hairpin introduction was occurring in all transduced cells.

5.3.2 shRNA screen using individual shRNAs for each candidate gene

Although it has been reported that pooled shRNAs targeted against a specific gene may result in a more efficient expression knock-down, this effect was not observed in the previous experiments. It was therefore possible that the presence of non-functional shRNAs in the pools may have dampened the knock-down effect of other, functional shRNAs that were also present. Hence, a further screen was implemented to assess the effect of individual shRNAs for those genes where shRNAs had previously been pooled.

The experimental set-up was performed as before. An additional empty vector control (pMSCV-Puro, Fig. 2.1) was included to assess the effect of the transduction process on hMSC differentiation. Double transduction of passage 6 hMSCs followed by drug

selection yielded a high level of stably transduced cells, as assessed by FACS analysis of cell transduced simultaneously with the pMSCV-Puro-eGFP-WPRE control virus (Fig. 5.9A). MSCs were then grown to confluence and induced to differentiate. After 13 days of differentiation, adipogenesis was quantified. It was found that the average number of adipocytes per field of view ranged from 40-73 in hMSCs expressing GOI-shRNAs (Fig. 5.9B). However, TBX2-X showed a significantly lower average number of adipocytes per field of view (mean average: 30, $p<0.01$ compared to scrambled control, or $p<0.05$ compared to empty vector control); these adipocytes also contained less lipid droplets than in other conditions (Fig. 5.9C). LHFPL2-X, LHFPL2-Y and TGIF-X shRNA expression resulted in significantly less adipocyte formation than in Scr-shRNA-transduced cells ($p<0.05$); however, this result was not statistically significant if the empty vector control was used for comparison ($p>0.05$). Scr-shRNA cells showed 75, empty vector 60 and untransduced hMSCs 76 adipocytes per field of view, on average (Fig. 5.9B and C); although these differences were not statistically significant this indicates that there is some experimental variation between control samples, which means that statistical significance should be interpreted with caution. As the reduction in adipocyte formation with TBX2-X was significant when compared to both controls, this result is considered reliable. In the majority of cases, the extent of differentiation did not differ significantly between control samples and cells transduced with GOI-shRNAs.

RT-PCR analysis was again used to investigate whether any reduction in gene expression had occurred (Fig. 5.10). The expression of each GOI was assessed in cells transduced with each shRNA targeting that GOI, plus Scr-shRNA, empty vector-transduced and untransduced cells. No large differences in expression of any GOI were seen between any of these conditions, including TBX2-X or LHFPL2-X which had shown lower differentiation. In these experiments, higher gene expression levels were not detected in untransduced hMSCs. These observations suggest that, as in the pooled shRNA screen, no knock-down of gene expression was observed with any of the shRNAs targeting GOIs, which could explain the observation of no reduction in adipogenesis in the majority of cases.
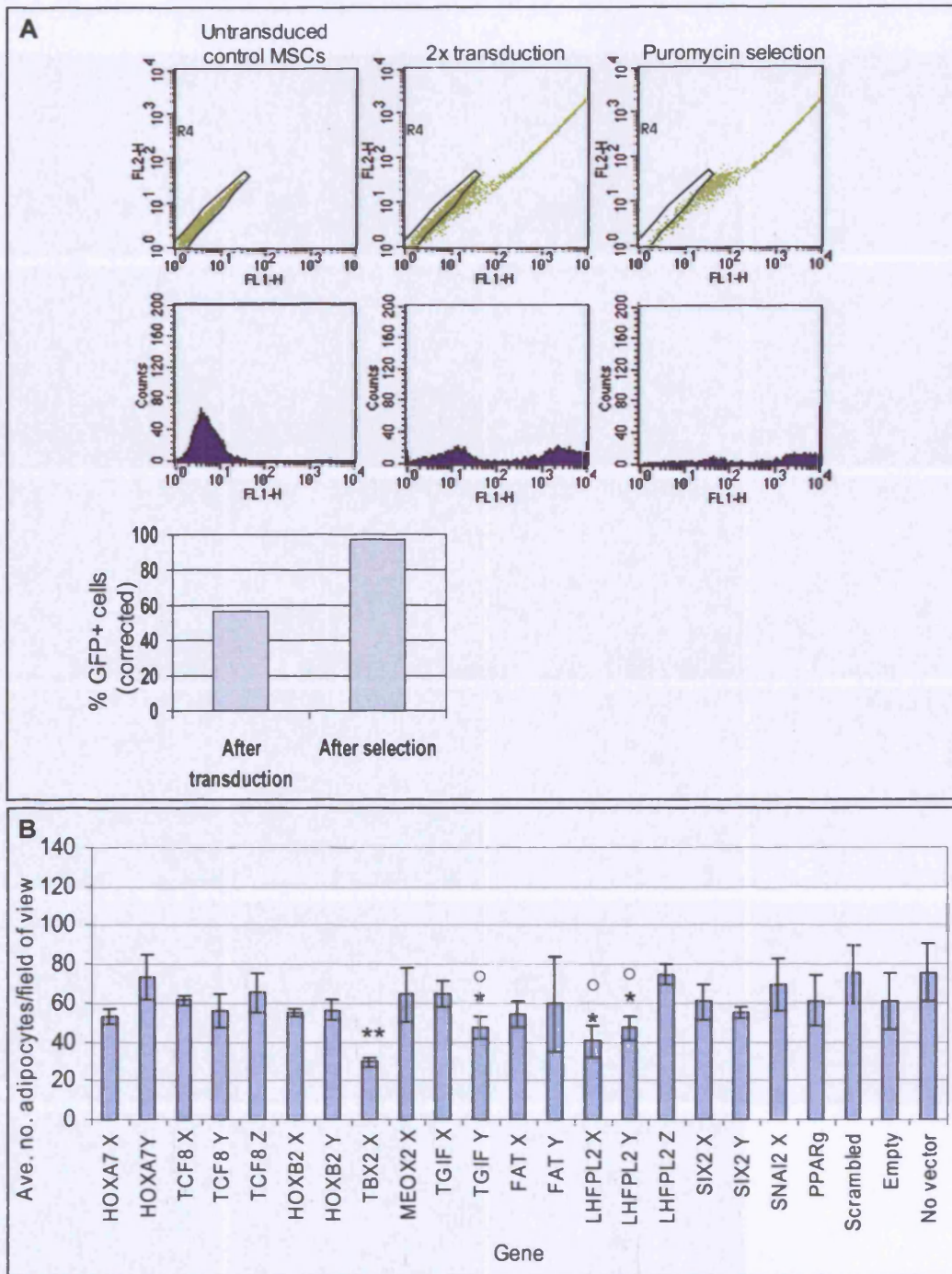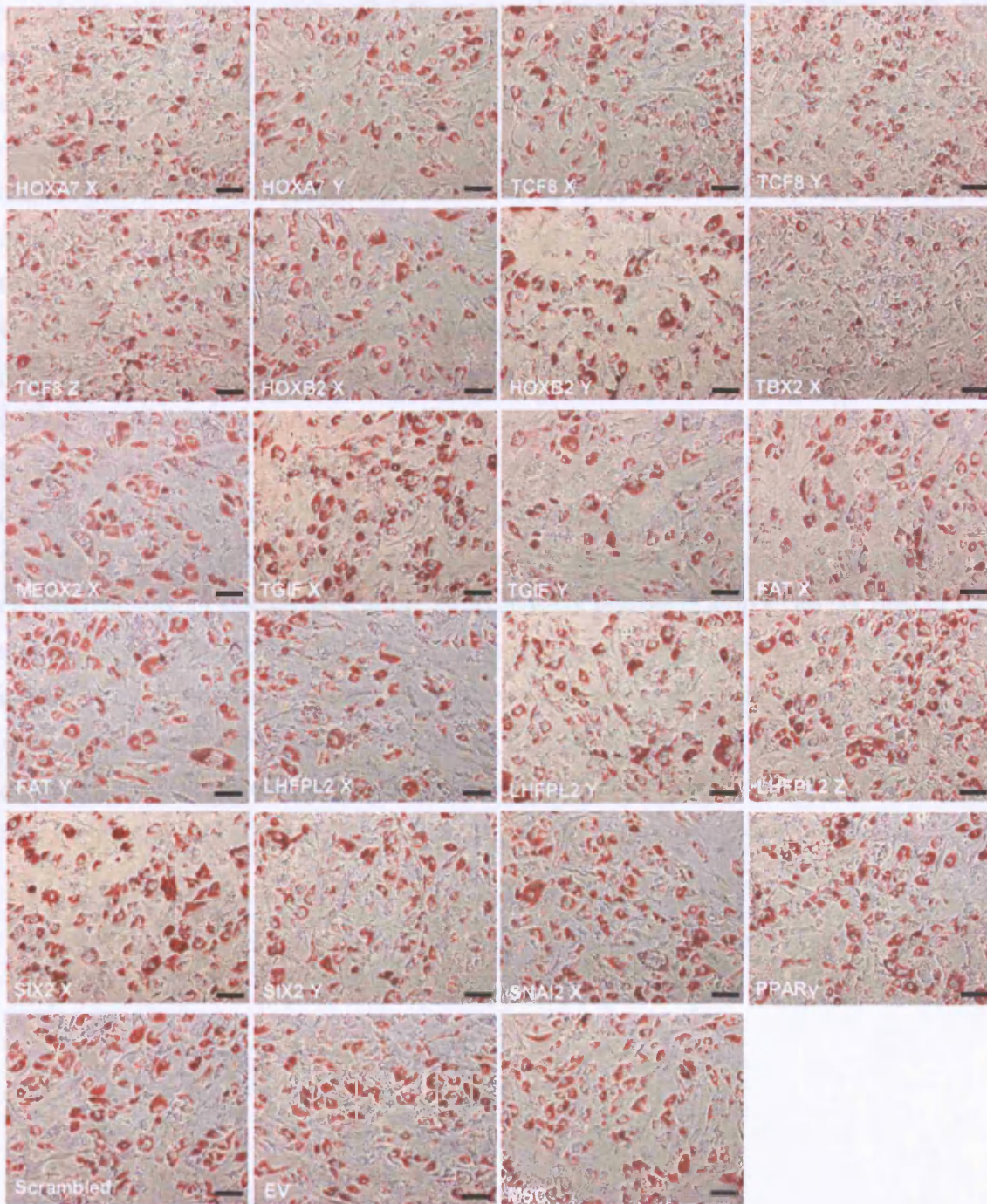
Figure 5.9. Continued overleaf.

Figure 5.9. Continued overleaf.

Fig 5.9. Effect on adipogenesis of individual shRNAs targeting candidate genes. hMSCs were transduced with viral particles (2.5.3) produced from transfection of 293gps with each shRNA, plus PPARγ, scrambled shRNA and an equivalent empty vector control (2.5.1) Cells were treated with puromycin (750ng/ml for 4 days) to select for stably transduced cells (2.5.5), then grown to confluence and induced to differentiate (2.2.2.1); A, FACS analysis of hMSCs transduced with virus produced from pMSCVPuro-WRPE-GFP (2.3.2). Shown is the number of GFP+ events in un-transduced hMSCs, hMSCs transduced with the GFP virus prior to puromycin selection, and hMSCs transduced with the GFP virus, after puromycin selection. Upper graphs represent the distribution of events (green dots) on the FL1 (green) versus FL2 (red) filters, and the gate (R4) used to separate GFP- from GFP+ events. Lower histograms (purple-filled) represent the number of GFP+ events at different fluorescent intensities. The bar chart represents the percentage of GFP+ cells in each condition, corrected for the untransduced control; B and C, at day 13 of differentiation, cells were fixed and stained with Oil Red O (2.2.3.1). Experiments were performed in duplicate; B, Quantification of the effect on differentiation of shRNA expression. Data represent number of adipocytes per field of view and are mean ± SD. Three fields of view were counted per condition. *, $p<0.05$ and **, $p<0.01$, using a two-sample t-test (assuming equal variance) when the mean value for each condition was compared with the Scrambled control; o, $p>0.05$ (not significant) when the mean value for each condition was compared with the empty vector control; C, representative image showing extent of differentiation for each shRNA. Scale bars represent 250μm.
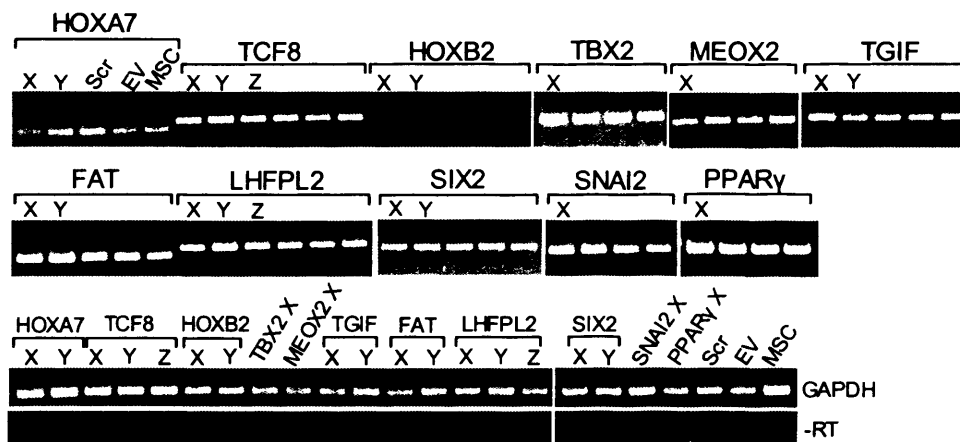
222

Fig 5.10. Effect of individual shRNA expression on candidate gene mRNA levels. RNA was extracted (2.6.1) from confluent hMSCs stably expressing shRNA for each GOI. cDNA was synthesised (2.6.2) and used as a PCR template with gene-specific primers (2.6.3.2). The expression of each GOI was compared in 4 conditions: hMSCs expressing each shRNA targeting the GOI (X, Y, or Z), scrambled hairpin (Scr), empty pMSCV-Puro vector (EV) or untransduced hMSCs (MSC). GAPDH and no reverse transcription controls (using adipsin primers) for each condition. PCR cycle numbers: GAPDH: 25; TBX2, MEOX2, SIX2: 35; all other genes: 30.

5.4 Development and optimisation of a technique for lentiviral delivery of custom-designed shRNAs

Due to the problems associated with the previous strategy, a second shRNA strategy was designed to maximise the chance of observing a knock-down of gene expression. Lentiviruses are capable of transducing non-dividing as well as dividing cells due to the presence of a nuclear localisation signal that facilitates their transport through the nuclear membrane of non-dividing cells. The use of lentiviruses would therefore increase transduction rates. Furthermore, this strategy, where possible, used shRNA sequences obtained from published research, which were therefore more likely to silence target gene expression. The aim of this experimental section was therefore to develop and optimise the use of lentiviral vectors and custom-designed shRNAs for investigating candidate gene function during adipogenesis.

5.4.1 shRNA design and lentiviral cloning strategy

This strategy was implemented for the top GOIs, namely DACT1, SNAI1, SIX2, and TCF8. It was not possible to recover shRNAs for DACT1 and SNAI1 from the RNAi library (Table 5.2), so these genes would be investigated for the first time using this strategy. SIX2 and TCF8 shRNAs had produced no expression knock-down in either the pooled or individual shRNA screens (sections 5.3.1 and 5.3.2), so this strategy provided a further opportunity to investigate their function. DACT1 (Yau *et al* 2005) and SNAI1 (Kajita *et al* 2004) shRNA sequences were obtained from published work so only one shRNA construct was required for each of these genes. Literature searches revealed no published 19nt shRNAs for SIX2 or TCF8. Although shRNA design software is becoming more advanced, it still does not guarantee that a single shRNA will silence a target gene. Therefore, to maximise the chance of obtaining a knock-down, software produced by OligoEngine (available at www.oligoengine.com) was used to design two 19nt shRNA sequences for both SIX2 and TCF8 (named SIX2/TCF8-a and -b, Table 2.4).

DNA oligonucleotides were designed (and purchased from OligoEngine), such that the sense and antisense 19nt sequence for each shRNA were separated by a hairpin loop sequence (Fig 5.11). *Bgl*II and *Hin*dIII restriction site sequences were included at the 5'
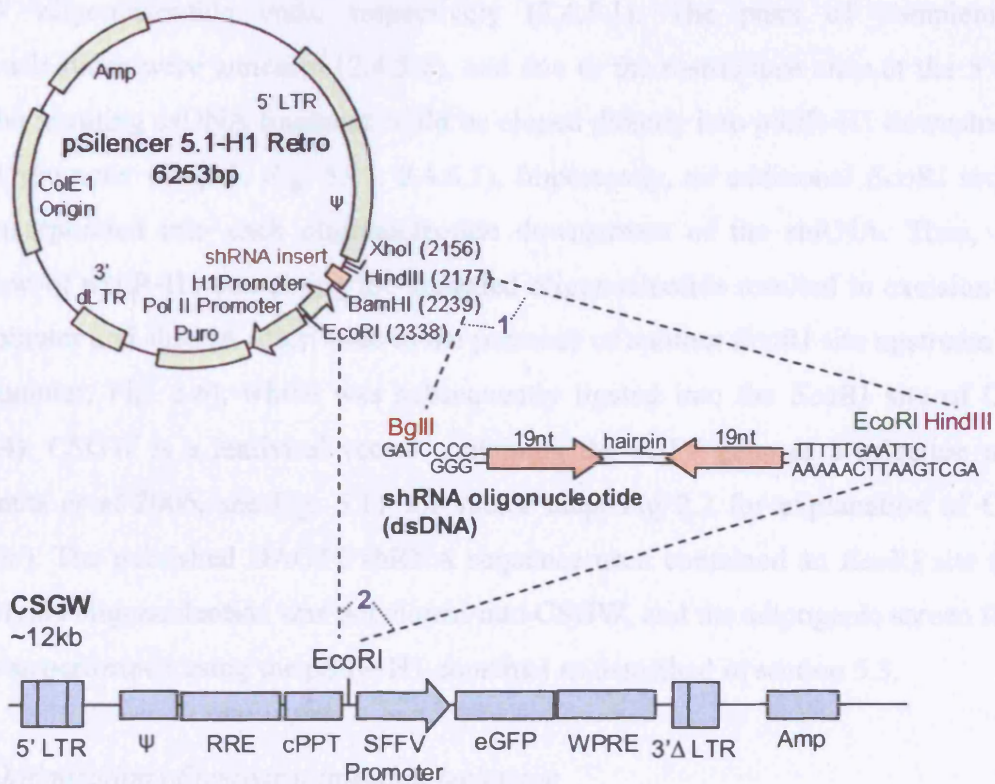
Figure 5.11. Design of lentiviral shRNA strategy. shRNA sequences for the top GOIs were either obtained from published work or designed using OligoEngine RNAi design software. The shRNA oligonucleotides were designed such that *Bgl*II (which produces compatible ends with *Bam*HI digests) and *Hin*dIII sites could be used to clone the shRNA oligonucleotide downstream of the H1 promoter in pSilR-H1 (step 1). Additionally, an extra *Eco*RI site was added to the oligonucleotide downstream of the shRNA so that *Eco*RI restriction digestion of the resulting pSilR-H1-shRNA plasmid would excise a fragment containing the H1 promoter and the shRNA (step 2). This fragment could then be ligated into the *Eco*RI site of CSGW.

225

and 3' oligonucleotide ends, respectively (2.4.5.1). The pairs of complementary oligonucleotides were annealed (2.4.5.2), and due to the restrictions sites at the 5' and 3' ends the resulting dsDNA fragment could be cloned directly into pSilR-H1 downstream of the H1 promoter (Step 1, Fig. 5.11, 2.4.5.3). Importantly, an additional *Eco*RI sequence was incorporated into each oligonucleotide downstream of the shRNA. Thus, *Eco*RI digestion of pSilR-H1 containing the annealed oligonucleotide resulted in excision of the H1 promoter and shRNA insert (due to the presence of another *Eco*RI site upstream of the H1 promoter, Fig. 5.6), which was subsequently ligated into the *Eco*RI site of CSGW (2.4.5.4). CSGW is a lentiviral vector containing the eGFP gene as a selection marker (Clements *et al* 2006, see Fig. 5.11 for vector map, Fig 2.2 for explanation of CSGW function). The published DACT1 shRNA sequence used contained an *Eco*RI site (Table 2.4), so this oligonucleotide was not cloned into CSGW, and the adipogenic screen for this gene was performed using the pSilR-H1 construct as described in section 5.3.

5.4.2 Optimisation of lentiviral infection technique

Competent viral particles are produced from CSGW by co-transfection of the 293T cell line with this vector and three other packaging plasmids, pLP1, pLP2 and pLP/VSVG (Fig. 2.2). These vectors cumulatively contribute all the proteins required to produce competent viral particles, but their presence on 3 separate plasmids reduces the probability of recombination producing an active, replication-competent virus. Furthermore, 293T cells allow production of large amounts of viral particles. Accordingly, viral supernatant for each construct shRNA-CSGW was produced (2.5.2). Additionally, virus was produced from empty CSGW, plus CSGW containing a shRNA targeting the v-cyclin gene of KSHV, under the control of the U6 promoter, for use as a control for off-target hairpin effects. Viral supernatant was filtered, aliquoted and stored at -80°C.

Initially, the minimum amount of virus that could transduce the maximum number of cells was determined. A high (~80% or more) transduction rate renders further selection steps unnecessary, and transduced cells can be identified through GFP expression. Furthermore, it is preferable to use the minimum viral titre necessary to avoid cytotoxic effects. To optimise cell transduction, the CSGW constructs were first titrated. Thus, 5×10^4 hMSCs per condition were transduced in suspension with serial dilutions of virus produced from

each CSGW construct (2.5.4), and after 48h the proportion of GFP$^+$ cells was analysed using FACS (Fig 5.12). Initially, a control hairpin targeting the KSHV LANA gene was employed, but was found be of insufficient quality for use, so the later adoption of the v-cyc-shRNA precluded it's addition in these analyses. Representative data for the empty CSGW construct are shown, and indicate that the lowest dilution assessed (1:2), was required to give the maximum transduction rate; higher dilutions resulted in less cells expressing GFP (Fig. 5.12). Thus, a saturation stage, where excess virus was being used, had not been reached in this dilution series.

When the transduction rate at this dilution was compared, it was found that CSGW transduction resulted in more than 80% of cells expressing GFP, whereas the shRNA-CSGW constructs yielded a 25-57% transduction rate (Fig. 5.13A). This suggests that the presence of a hairpin may decrease the efficiency of virus production or transduction of target cells. The viral copy number per cell was next determined using Q-PCR. Genomic DNA was extracted (2.5.6) from hMSCs transduced as described with a 1:2 dilution of viral supernatant from each of the constructs. Primers detecting the viral packaging signal (ψ) in CSGW, plus GAPDH to assess the number of cells present in the sample, were used to calculate the average number of copies of virus present in each cell (2.5.6). It was found that empty CSGW transduction led to an average of 12.2 copies per cell, whereas this number ranged from 2.8-6.1 for shRNA-CSGW constructs (Fig. 5.13B). The pattern of viral copy number per cell was comparable to the pattern of number of GFP$^+$ cells for each construct (Fig 5.13A and B). It could be expected that a higher average viral copy number per cell would correspond to a higher number of GFP$^+$ cells, thus indicating the correspondence between these observations for each construct. This finding again highlighted the fact that the presence of shRNA hairpins either yields a lower viral titre after 293T transfection, or that viral particles containing shRNA hairpins are less able to transduce target cells. Furthermore, it was noted that GFP expression appeared stronger in empty CSGW-transduced cells than in those transduced with shRNA-CSGW (Fig. 5.13C). A number of possibilities could explain this observation, including that the presence of a shRNA may have affected the level of gene expression from the lentivirus, or may have had a wider effect on gene expression. Due to the significant differences in transduction rates between empty and hairpin-containing CSGW, it was decided to use neat viral
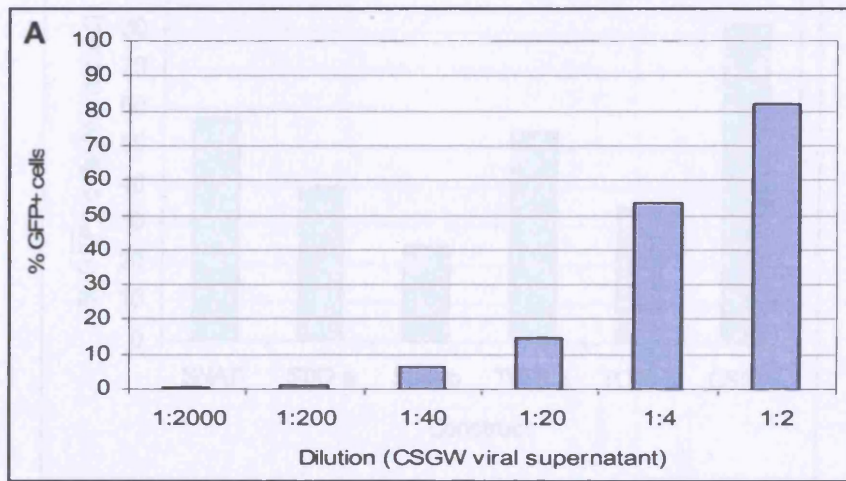
Figure 5.12. Titration of the hMSC transduction rate for each lentiviral construct. 293T cells were transfected with lentiviral construct for each gene plus packaging plasmids (2.5.2). After 60h the viral supernatant was harvested, filtered and added to hMSCs at different dilutions (2.5.4); A, FACS analysis (2.3.2) of GFP expression after hMSC transduction with decreasing dilutions of CSGW (empty lentiviral vector, shown as a representative titration). The pattern of transduction rate at different dilutions is representative of all other constructs, but actual numbers vary (see Fig. 5.13A).
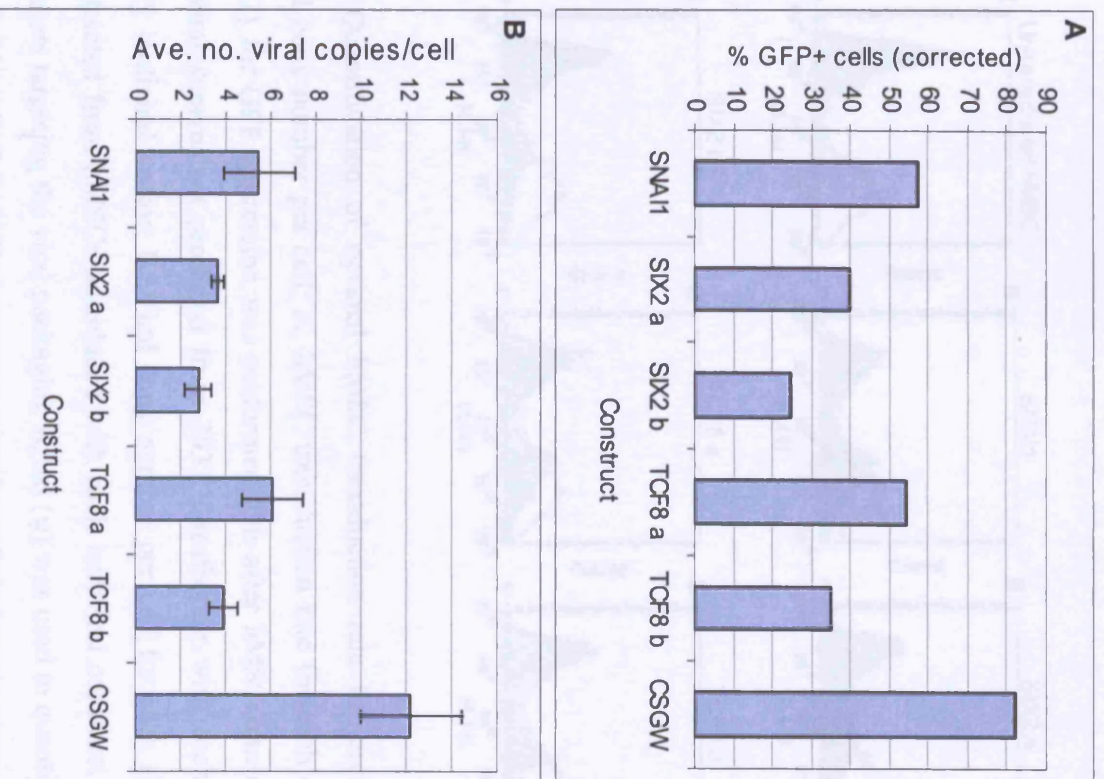
Figure 5.13. Continued overleaf.

229

C

Untransduced hMSC | SNAI1 | SIX2 a
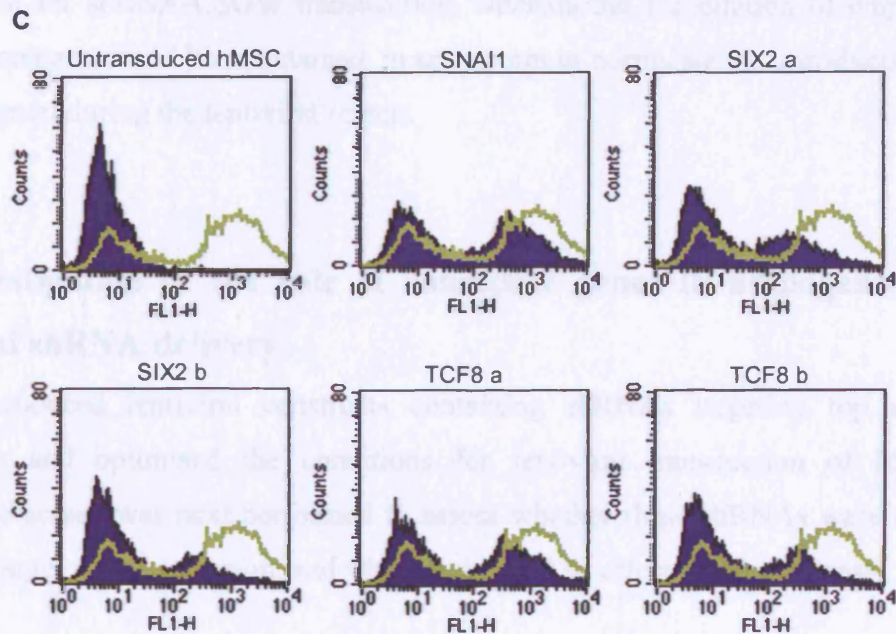
SIX2 b | TCF8 a | TCF8 b

Figure 5.13. Quantification of optimal hMSC transduction rate for each construct and resulting viral copy number per cell; A, hMSC transduction rate for each construct. FACS analysis (2.3.2) for GFP expression was performed 48h after hMSC transduction with 1:2 dilutions of viral supernatant produced from 293T transfection with each viral construct. CSGW: empty lentiviral vector; B, Viral copy number per cell for each shRNA. Genomic DNA was extracted from hMSCs transduced with each lentiviral construct and quantitative PCR with primers targeting the viral packaging signal (ψ) was used to quantify the number of integrated virus per cell (2.5.6). Data represent number of viral copies per cell and are mean ±S.D. This was performed in collaboration with R. Vart; C, FACS histogram plots of GFP fluorescence after hMSC transduction with 1:2 dilutions of viral supernatant. Purple filled histograms indicate GFP fluorescence produced by the lentiviral construct indicated, and the green line indicates GFP fluorescence in hMSCs transduced with empty CSGW.

supernatant for shRNA-CSGW transduction, whereas the 1:2 dilution of empty CSGW viral supernatant would be maintained, in an attempt to normalise the transduction rate for each construct during the lentiviral screen.

5.5 Investigation of the role of candidate genes in adipogenesis using lentiviral shRNA delivery

Having produced lentiviral constructs containing shRNAs targeting top adipogenic candidates and optimised the conditions for lentiviral transduction of hMSCs, an adipogenic screen was next performed to assess whether these shRNAs were capable of reducing target gene expression, and whether this had an effect on adipogenesis.

According to the optimisation described above, $5x10^4$ passage 6 hMSCs per condition were transduced in suspension with undiluted viral supernatant for shRNA-CSGW constructs, and a 1:2 dilution of empty CSGW viral supernatant was used. Viral supernatant from the v-cyc-shRNA construct was used neat; although the transduction efficiency or viral copy number had not been determined for this construct, it was deemed probable that the presence of a hairpin would cause the same reduction in transduction efficiency that was observed for the other hairpin constructs. After 60h, GFP expression in transduced hMSCs was determined. Cells were also expanded for 3 passages to assess whether GFP expression was maintained over this time. Figure 5.14A shows that for most constructs, near 100% of cells were transduced under these conditions, and the level of GFP expression was maintained after expansion. These results suggest that the use of neat viral supernatant had had the desired effect of increasing transduction efficiency for shRNA-CSGW constructs. Furthermore, there was no need for further selection of stably transduced cells, as GFP expression showed that the majority of cells would be expressing the shRNA. With the shRNA construct SIX2-b, a slight reduction in GFP$^+$ cells was observed after expansion, from 97.1 to 88.5%, possibly indicating a growth disadvantage conferred by this shRNA (if it also had an effect on target gene expression). However, this rate of transduction was deemed sufficiently high for use in the differentiation screen.
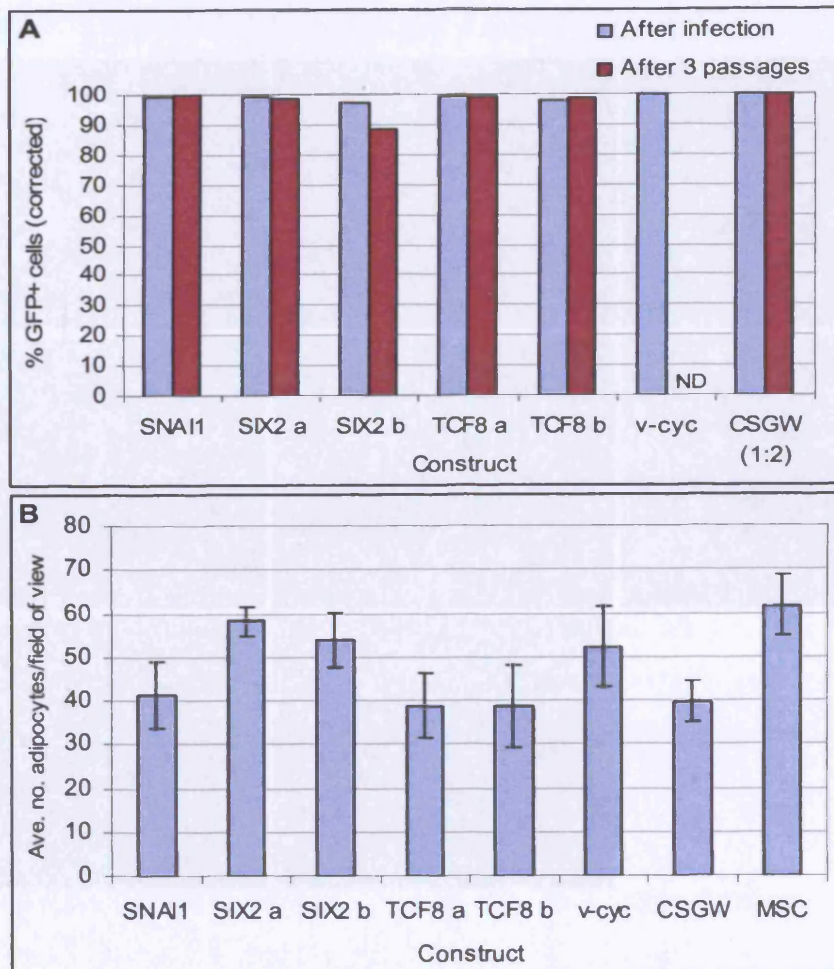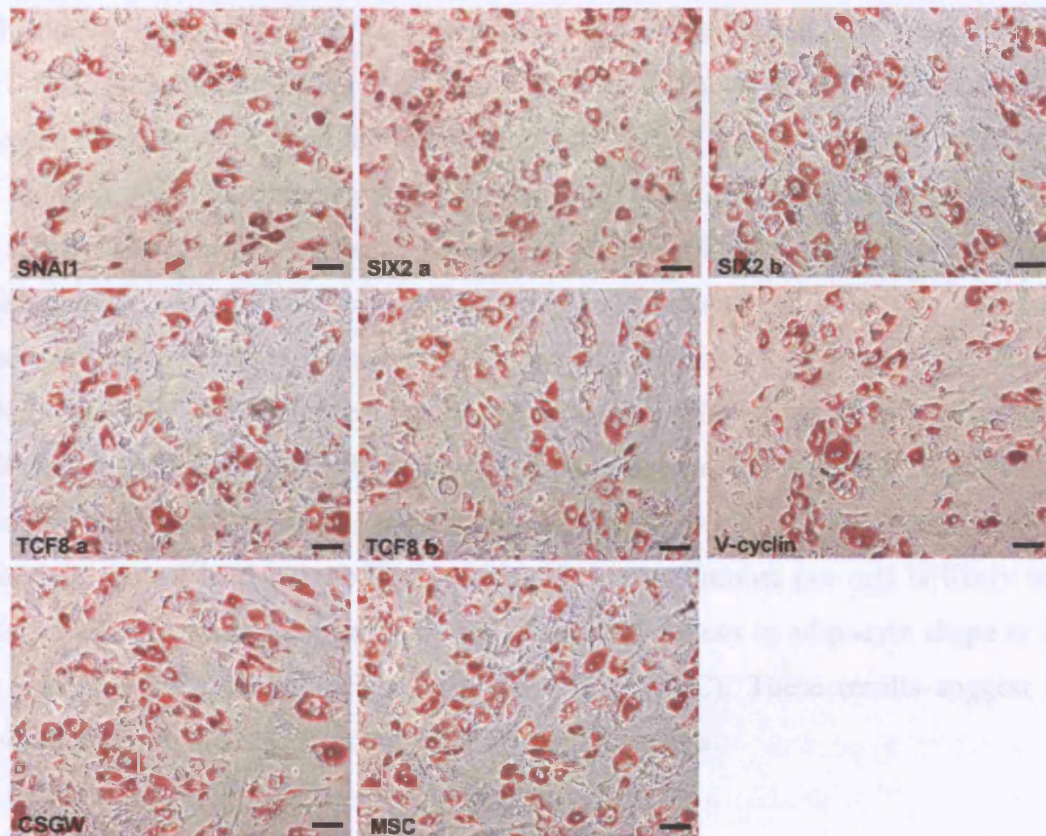
Figure 5.14. Continued overleaf.

Figure 5.14. Transduction rate and extent of differentiation during lentiviral shRNA screen of top GOIs; A, hMSCs were transduced with neat viral supernatant (apart from empty vector – 1:2 dilution) in order to maximise transduction rates (2.5.4). Cells were analysed for GFP expression using FACS analysis (2.3.2) 48h after transduction (blue bars), and after 3 passages (purple bars), in order to assess whether a growth disadvantage was conferred on transduced cells. CSGW: empty lentiviral vector; v-cyc: CSGW containing shRNA hairpin targeting v-cyclin gene of KSHV; MSC: untransduced hMSCs. ND, not determined; B, extent of differentiation in cells transduced with each construct. Transduced cells were grown to confluence then induced to differentiate to adipocytes (2.2.2.1). At day 14 of differentiation, cells were fixed and stained with Oil Red O (2.2.3.1), and the number of adipocytes per field of view was counted for 4 fields of view per condition. Data represent number of adipocytes per field of view and are mean ± SD. Experiments were performed in duplicate. Two-sample t-tests (assuming equal variance) identified no significant differences (p>0.05) when the mean of each condition was compared to the v-cyc shRNA control; C, representative field of view for each condition. Scale bars represent 200μm.

233

Transduced hMSCs were therefore passaged once at a ratio of 1:3 and then grown to confluence in triplicate wells for each condition. RNA was extracted from one well for use in RT-PCR, and the remaining confluent cells were induced to differentiate using the adipogenic induction cocktail. Differentiation was carried out for 14 days, after which time the cells were fixed and stained with Oil Red O, and differentiation was quantified (Fig 5.14B). The average number of adipocytes per field of view ranged from 40-62; the highest number was observed in untransduced hMSCs (average: 62), but this was similar to the number observed in cells expressing the SIX2-a shRNA (average: 58). The two shRNAs targeting TCF8 showed the lowest extent of differentiation, but this was not significantly different from the rate of differentiation observed in v-cyc-shRNA transduced cells (p>0.05, two-sample Student's t-test), which was considered the most appropriate control in this experiment as the viral copy number per cell is likely to be similar in all cells containing shRNAs. No visible differences in adipocyte shape or size were observed between any of the conditions (Fig 5.14C). These results suggest that shRNA expression did not affect the process of adipogenesis.

As with previous retroviral screens, it was crucial to next determine whether any expression knock-downs had been produced by shRNA introduction. Therefore, cDNA was synthesised from RNA extracted from transduced hMSCs, and was used as a PCR template using gene-specific primers. The expression of each gene was determined in cells transduced with each shRNA targeting that GOI, plus v-cyc-shRNA, empty vector-transduced and untransduced cells (Fig. 5.15). As with the pooled shRNA retroviral screen, expression of each GOI seemed slightly higher in untransduced hMSCs, and a drop in expression was seen after transduction even with empty virus. However, no significant difference was seen in the extent of adipogenesis between these conditions, meaning that the process of differentiation was unaffected by the process of transduction. Interestingly, a consistent observation was the lower expression of SIX2 and TCF8 in cells transduced with SIX2 b-shRNA and TCF8 a-shRNA, respectively (Fig. 5.15). This could indicate that the SIX2-b and TCF8-a shRNAs had induced some extent of transcriptional silencing of their respective target genes. However, there was no significant increase or decrease in adipogenesis with these shRNAs.
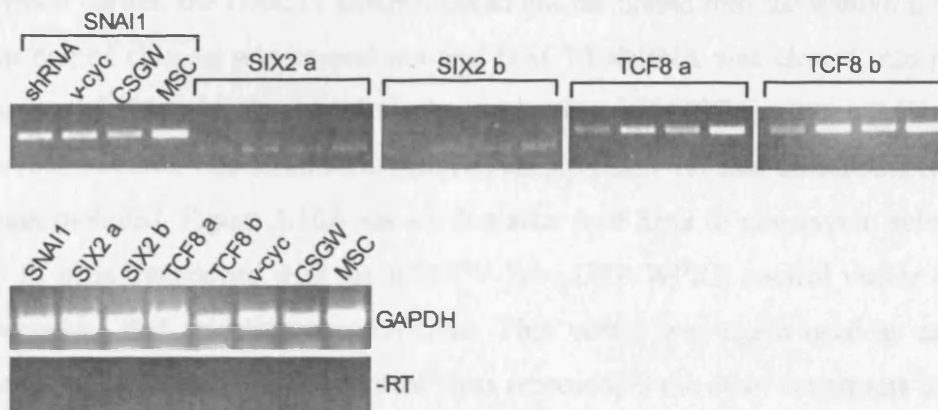
Figure 5.15. Effect of lentiviral-mediated shRNA expression on candidate gene mRNA levels. RNA was extracted (2.6.1) from confluent hMSCs expressing shRNA(s) for each GOI. cDNA was synthesised (2.6.2) and used as a PCR template with gene-specific primers (2.6.3.2). The expression of each GOI was compared in 4 conditions: hMSCs expressing each shRNA targeting the GOI (shRNA), shRNA hairpin targeting KSHV viral cyclin (v-cyc), empty CSGW vector (CSGW) or untransduced hMSCs (MSC). GAPDH: internal cDNA loading control and –RT, no reverse transcription control (using adipsin primers). PCR cycle numbers: GAPDH: 25, all other genes: 28.

As described earlier, the DACT1 shRNA could not be cloned into the lentiviral vector, so only step one of cloning was carried out and DACT1-shRNA was cloned into pSilR-H1. Transfection of 293gps followed by viral transduction of hMSCs was carried out as with the retroviral screens. The scrambled-shRNA, empty pSilR-H1 and untransduced controls were again included. Figure 5.16A shows that after four days of puromycin selection, the majority of cells transduced with the pMSCV-Puro-GFP-WPRE control vector expressed GFP, indicating that selection was efficient. This vector was again used as an internal control as it was assumed that the control virus represented the other constructs in terms of the efficiency of each experimental step. After drug-selected cells were grown to confluency, differentiation was induced and continued for 15d, then cells were Oil Red O-stained and differentiation was quantified. Average numbers of adipocytes per field of view ranged from 27-34 in the four conditions and there were no statistically significant differences between these observations (p>0.05, Fig. 5.16B). No major morphological differences were seen in adipocytes from each condition (Fig. 5.16C). When the expression of DACT1 was examined, no reduction could be seen in DACT1 shRNA-transduced cells compared with controls (Fig. 5.16D). These results suggest that this shRNA did not lead to reduced DACT1 expression during adipogenesis, and consequently differentiation was unaffected.
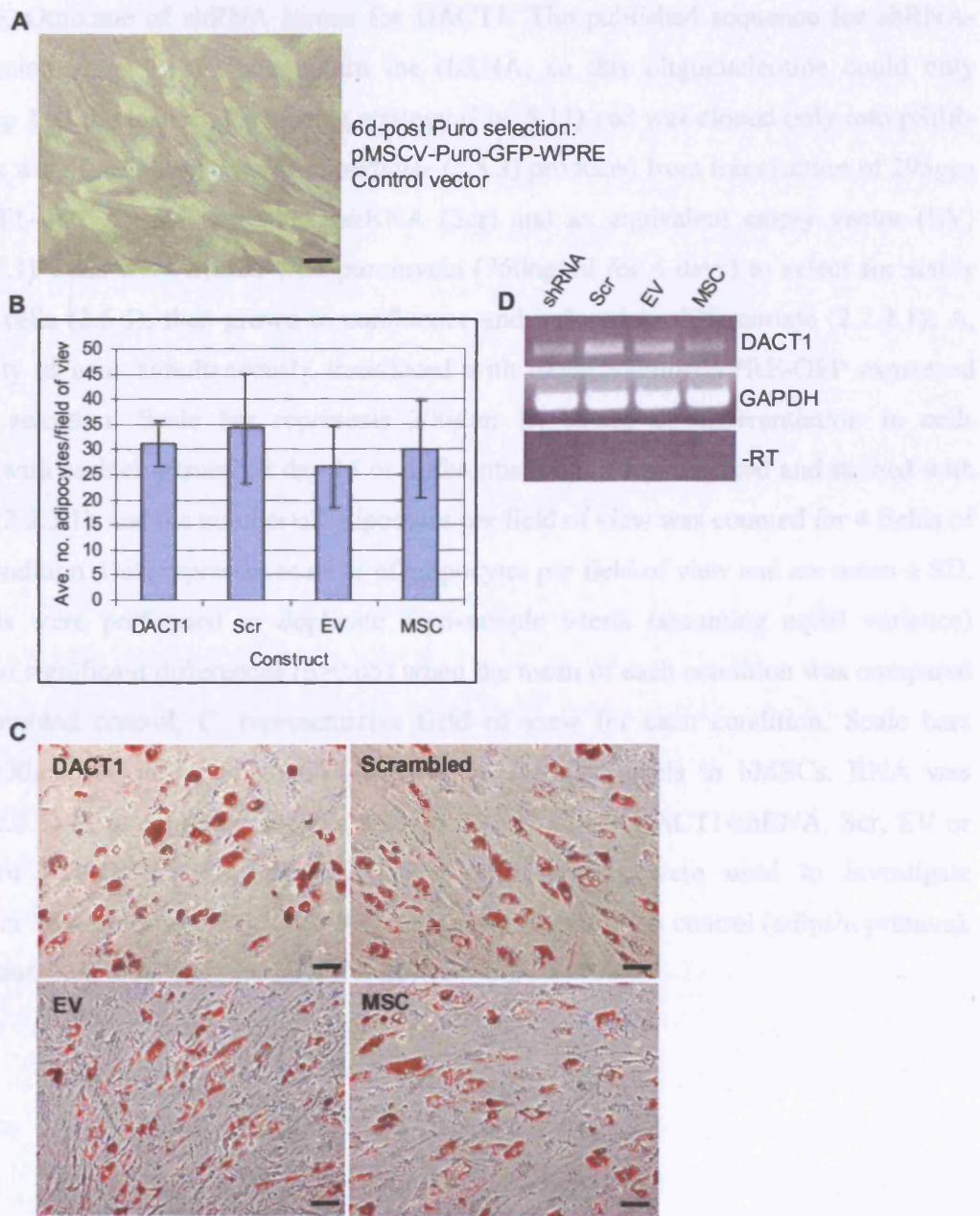
Figure 5.16. Continued overleaf.

Figure 5.16. Outcome of shRNA screen for DACT1. The published sequence for shRNA-DACT1 contained an *Eco*RI site within the shRNA, so this oligonucleotide could only undergo step 1 of the lentiviral screening strategy (Fig. 5.11) and was cloned only into pSilR-H1. hMSCs were transduced with viral particles (2.5.3) produced from transfection of 293gps with DACT1-shRNA, plus scrambled shRNA (Scr) and an equivalent empty vector (EV) control (2.5.1) Cells were treated with puromycin (750ng/ml for 4 days) to select for stably transduced cells (2.5.5), then grown to confluence and induced to differentiate (2.2.2.1); A, The majority of cells simultaneously transduced with pMSCV-Puro-WPRE-GFP expressed GFP after selection. Scale bar represents 200μm; B, extent of differentiation in cells transduced with each construct. At day 15 of differentiation, cells were fixed and stained with Oil Red O (2.2.3.1), and the number of adipocytes per field of view was counted for 4 fields of view per condition. Data represent number of adipocytes per field of view and are mean ± SD. Experiments were performed in duplicate Two-sample t-tests (assuming equal variance) identified no significant differences (p>0.05) when the mean of each condition was compared to the Scrambled control; C, representative field of view for each condition. Scale bars represent 200μm; D, effect of DACT1-shRNA on DACT1 levels in hMSCs. RNA was extracted (2.6.1) from confluent hMSCs transduced with either DACT1-shRNA, Scr, EV or untransduced hMSCs (MSC), and DACT1-specific primers were used to investigate expression in each condition (2.6.3.2). –RT: no reverse transcription control (adipsin primers). PCR cycle numbers: GAPDH: 25, DACT1: 28.

5.6 Discussion

The results presented in this chapter describe the development, optimisation and performance of two targeted shRNA screening techniques aimed at investigating the function of candidate genes in hMSC adipogenesis. Although shRNAs targeting a selection of candidate genes were isolated and efficiently delivered to hMSCs via retroviral or lentiviral vectors, no significant silencing effect was observed for the majority of target genes, and therefore no effect on adipogenesis was apparent. However, two lentivirally delivered shRNAs may have induced a knock-down of their target genes, indicating that this methodology is functional, but that further work would be needed to determine whether these genes play a role during adipogenesis.

The initial approach undertaken was to use the NKi RNAi library to perform a shRNA screen targeting 83 novel candidate regulators, and investigate their role in adipogenesis. However, attempts at recovering shRNAs for these genes revealed a high level of recombination inherent in the library (Fig 5.2), and resulted in recovery of shRNAs for only 10 GOIs (Table 5.2). Recombination has been documented as a major problem by the developers of this library, and also affects other vector-based RNAi libraries (Downward 2004). This suggests that large-scale screens with this library are only possible when recombination is absent, as in the master copy held by the library developers (Berns et al 2004, Kolfshoten et al 2005, and personal communication from NKi library distributors). The issue of recombination is being addressed in modified versions of vector-based libraries via the use of recombination-deficient bacterial strains, or the inclusion of selectable markers close to the shRNA insert (Echeverri and Perrimon 2006). However, in this study, strategies were developed that allowed use of the RNAi library on a smaller scale, through sub-cloning of shRNA inserts into an alternative, non-recombining vector. In addition, a lentiviral-based shRNA screen was also employed which resulted in the inclusion of 2 further GOIs, plus an additional analysis of 2 GOIs that were included in the retroviral library screen, resulting in a thorough analysis of 12 GOIs. Although this number is the maximum that could be achieved in this study, it may have limited the chance of identifying a candidate gene that plays an essential role in the early stages of hMSC adipogenesis. Further work could be designed to use newer, non-recombining

vectors in a more high-throughput format to study a larger number of genes; possibly all genes identified through the microarray studies of Chapter 4.

A trend resulting from each shRNA screen was that adipogenesis was not significantly reduced (or a reliable significant decrease could not be detected as there was a level of variation in control samples) after transduction with any shRNA construct (or pool), compared to the appropriate controls (either control hairpin or empty vector, where present, Figs 5.7, 5.9, 5.14, 5.16). RT-PCR analysis revealed that the majority of candidate genes did not appear to have been silenced by their shRNAs, providing an explanation for the lack of effect on adipogenesis (Fig 5.8, 5.10, 5.15, 5.16). Possible reasons for this lack of silencing are discussed below. An exception to this was TBX2 X-shRNA in the retroviral individual screen, where adipogenesis was significantly decreased compared to both the scrambled shRNA and empty vector controls (Fig. 5.9). Although this shRNA gave the lowest rate of adipogenesis in the pooled screen, it was not significantly less than the other shRNAs. Furthermore, no decrease in TBX2 expression was detected using RT-PCR in either screen, suggesting that this result may have been due to experimental variation. It is possible that a knock-down had occurred which was not identified by the RT-PCR methodology. The use of more sensitive quantitative RT-PCR could address this issue, but was beyond the time-frame of this study.

RT-PCR analysis of gene expression in the lentiviral screen indicated a possible decrease in SIX2 and TCF8 expression with the shRNA SIX2-b and TCF8-b (Fig 5.15). This effect was consistently observed. These genes were included in both the retroviral library screen and lentiviral screen, which suggests that the custom-designed shRNA sequences for these genes were more effective at silencing gene expression than those isolated from the RNAi library. Alternatively, lentiviral transduction may have resulted in more efficient shRNA expression within cells than with retroviral transduction. Lentiviruses have been reported to induce gene knock-down in hMSCs (Cho *et al* 2006, Clements *et al* 2006). There is less literature describing the use of retrovirally-delivered shRNA in hMSCs, although several studies have used a retroviral vector to deliver shRNA into murine MSCs (Hong *et al* 2005, Okuyama *et al* 2006). Lentiviral vectors have several advantages over retroviral vectors, including that they can be easily and accurately titred (see below), and can

transduce non-dividing as well as dividing cells. Thus, it could be suggested that lentiviral vectors, where possible, should be the focus of future work involving shRNA in hMSCs.

Overall, the observations suggest that these shRNAs were functional and the method of delivery into hMSCs allowed them to exert a silencing effect on gene expression. However, as adipogenesis was unaffected, it is probable that either the knock-down of TCF8 and SIX2 was too small to produce an effect on differentiation, or that it was large enough but these genes play a redundant role, or have no role, in adipogenesis. The percentage of transduced cells was found to decrease slightly over three passages in SIX2-b-transduced hMSCs (Fig. 5.14A). This could possibly indicate, if the knock-down did cause significant inhibition of SIX2 function, that SIX2 may play a role in hMSC biology rather than specifically in differentiation. Western blots would be required to assess SIX2 and TCF8 protein levels, but were not performed in this thesis due to time constraints.

The most probable explanation for the fact that few shRNAs produced an observable decrease in target gene expression is that they were non-functional. As previously mentioned, si(sh)RNA design strategies are not yet 100% efficient, so there is no guarantee that every shRNA will be able to silence gene expression. Furthermore, it was stated that only 70% of NKi library targets were silenced in initial work (Berns *et al* 2004), and this was when pools of 3 shRNAs were used; for many genes targeted in this thesis only 1 or 2 shRNAs could be isolated, reducing the chances of finding at least 1 functional shRNA. Both the SNAI1 and DACT1 shRNAs were expressed from the U6 promoter in published work (Kajita *et al* 2004, Yau *et al* 2005), whereas the H1 promoter was used here, providing a possible explanation for the absence of knock-down by these functionally-confirmed shRNAs. It is also possible that hMSCs are somewhat resistant to the effects of RNAi: cell type-specific differences in efficiency of shRNA and siRNA-mediated gene silencing have been reported, which were hypothesised to be due to differences in expression of intermediates of the RNAi pathway in certain cells, or differences in an interferon response elicited by the viral shRNA constructs in each cell type (Bantounas *et al* 2004). Importantly, several additional observations indicate that in some cases, other factors during experimentation may have affected the extent of gene knock-downs, and/or the ability to detect them.

During the pooled retroviral screen, it was noted that target gene expression was lower in cells transduced with both scrambled shRNA as well as GOI-shRNA, compared with untransduced hMSCs (Fig. 5.8A). This was accompanied by a significant decrease in the extent of adipogenesis in these cells, compared with untransduced hMSCs. This effect was also seen, but to a lesser (non-significant) extent, in the lentiviral screen (Fig. 5.15). Gene expression seemed higher in untransduced hMSCs than in cells containing hairpin constructs; even cells containing empty vector only showed GOI expression more similar to v-cyc-shRNA control cells than to untransduced hMSCs. However, in the lentiviral screen this decrease in gene expression after hairpin (or possibly vector) introduction was less, and was not accompanied by a decrease in adipogenesis. Moreover, such effects were not observed in the individual library screen or the DACT1 experiment, suggesting that the cause of these observations is variable in its extent and occurrence.

It has been reported that some siRNA sequences can induce the silencing of genes other than their target gene due to regions of partial homology, possibly as small as 11 nucleotides (Jackson *et al* 2003). It was therefore possible that silencing of the GOIs had been induced in cells transduced with the non-human shRNA controls (scrambled and v-cyc). However, this was not likely to be the case as GOI expression was also reduced in PPARγ-shRNA-transduced cells (Fig. 5.8C), indicating that the effect was more widespread and not sequence-specific. Furthermore, GOI expression was also slightly reduced in empty CSGW-transduced cells. Therefore, the observations of decreased gene expression and occasionally decreased adipogenesis are likely to have been due to sequence-independent effects of either shRNA introduction, viral introduction, or both.

Double-stranded RNA, in the form of shRNA or siRNA, has been shown to produce sequence-independent (termed non-specific) stimulation and suppression of non-target genes (Persengiev *et al* 2004, Kariko *et al* 2004). It could be hypothesised therefore that the expression of shRNAs (control or GOI-shRNA) induced non-specific silencing of a subset of genes, including the GOIs. However, it seems unlikely that the effect of this phenomenon would be broad enough to affect every GOI, or to affect enough genes to impair adipogenesis (as in the pooled retroviral screen).

A further, and more probable, explanation for the observations is the induction of an anti-viral interferon response. Interferon responses can be triggered by dsRNA in the form of both shRNA and siRNA (Bridge *et al* 2003, Sledz *et al* 2003), and have also been shown to occur after the introduction of some vectors into cultured cells (Akusjärvi *et al* 1987, Bridge *et al* 2003). This response results in upregulation of interferon-stimulated genes, mRNA degradation and a general reduction in protein synthesis (Williams and Haque 1997). Thus, it is possible that either shRNA or vector introduction triggered such a response, which led to the reductions in gene expression and adipogenesis specific to some of the screens. This would possibly make observing shRNA-induced knock-downs more difficult, as control as well as GOI-targeted cells may already have some extent of decrease in gene expression and/or differentiation capability. Confirmation that these observations were specifically due to an interferon-mediated response would require analysis of the expression of interferon-related genes such as 2'5'-oligoadenylate synthetase in these experiments. This analysis is beyond the scope of this thesis, but an interferon response is currently considered hypothetically as the most viable option to explain the cumulative findings.

It is not possible in the pooled shRNA screen to determine whether such an effect may have been caused by the shRNA or the virus itself; the Scr-shRNA was initially deemed a sufficient control so an empty vector was not included. However, the slight reduction in gene expression after empty CSGW transduction would indicate that the virus itself is, at least in part, responsible for such an effect. Furthermore, in the pooled shRNA screen, puromycin-selected control cells (which therefore all contained the pMSCV-Puro-WPRE-GFP control vector) mostly expressed GFP at a low level (Fig. 5.7A), which could indicate that during this experiment, a response had been triggered that resulted in lower GFP expression. However, in subsequent experiments this effect was not seen and GFP expression was high, and in previous work in Chapter 3 retroviruses had been used to transduce hMSCs with the PPARγ gene with no effects on hMSC biology (other than those caused by PPARγ, Fig. 3.9).

A possible reason for this observed variability in response to virus/shRNA expression was that the viral titre or viral copy number per cell may have varied between experiments. Retroviruses were not titred as stocks of the viral supernatant must be frozen whilst titring via Q-PCR. As retroviral particles are relatively unstable when frozen, the viral titre is reduced significantly by this process, so Q-PCR results would not represent the titre of frozen stocks. Furthermore, the maximum virus concentration possible is always used during experimentation to avoid low cell densities after drug selection. MSCs, and stem cells in general, can be difficult to transduce (Godfrey 2006), and low cell densities produced by low transduction rates lead to cell senescence and an inability to expand cells after drug selection. It is therefore possible that viral titre was higher in the pooled shRNA screen than in the individual shRNA screen, resulting in variation in the response to viral transduction or hairpin introduction. Alternatively (or additionally), if fewer hMSCs were dividing, or were dividing more slowly during the transduction process for the pooled screen, then those cells that were dividing may have become highly transduced (as retroviruses can only transduce dividing cells). This may have resulted in a population of highly transduced cells specifically in this experiment that then induced a response.

The overall effect of a general decrease in gene expression or differentiation capability in all cells (i.e. including control cells thus the effect is not caused by the target gene shRNA) could be that any expression knock-down or phenotype produced by functional shRNAs would be masked. This could have occurred in the pooled shRNA screen. However, the individual shRNA screen was in many ways a replication of the pooled screen, and provided more information as each shRNA was assayed individually. As no non-specific effects were observed, the experimental readout is likely to be accurate. Furthermore, effects in the lentiviral screen were minimal, and gene knock-downs were detected. Thus, a comprehensive study of the effect of shRNAs targeting 12 GOIs on gene expression and hMSC adipogenesis has been performed, although knock-down of expression was observed for only 2 genes.

There have been several papers documenting the use of siRNA oligonucleotides to study gene function in hMSC differentiation (Luo *et al* 2004, Xu *et al* 2006) and in other stem cell populations (Oliveira and Goodell 2003, Matin *et al* 2004). It is possible therefore that

siRNAs are more effective at silencing gene expression in stem cells, although they are also capable of inducing non-specific cellular responses (Sledz *et al* 2003). However, I consider that their use to study processes such as hMSC differentiation may not be ideal as they have only a transient effect on gene expression, and as differentiation was continued in this study for 14 days, continued reapplication of siRNA would be necessary and knock-down effect may vary. In summary, this study has utilised a novel approach to the study of early hMSC differentiation, as vector-mediated shRNA was employed and multiple genes were investigated simultaneously; previous work has often focused on individual genes. Potential problems of such an approach have also been documented, specifically involving variable non-specific effects of virus or hairpin introduction into hMSCs, which have highlighted possible areas for future optimisation.

CHAPTER 6. THE EFFECT OF TRANSFORMING
MUTATIONS ON hMSC DIFFERENTIATION *IN VITRO*.

6.1 Introduction and aims

The existence of cancer stem cells (1.8.1) was first demonstrated in acute myeloid leukaemia (Bonnet and Dick 1997). Further cancer stem cell populations have since been prospectively isolated from solid tumours of the breast and brain (Al-Hajj et al 2003, Singh et al 2003, Singh et al 2004). Cancer stem cells have the ability to reconstitute the full and specific range of cell types in their cancer of origin after transplantation into an animal model, and also to give rise to further cancer stem cells that have this same ability after serial transplantation (Bonnet and Dick 1997, Al-Hajj et al 2003, Singh et al 2004). Thus, cancer stem cells have the potential for self-renewal as well as (aberrant) differentiation. The striking resemblance of cancer stem cells to normal stem cells has led to the suggestion that cancer may arise from the aberrant functioning of normal stem cell attributes (1.7.2, Reya and Clevers 2005). In support of this, numerous genetic alterations that lead to tumour progression also control normal stem cell function (1.8.2, Pardal et al 2003, Reya and Clevers 2005, Polyak et al 2006). Based on such observations, it has been suggested that tumourigenic mutations that give rise to these abnormal properties may arise in adult stem cells, and evidence in support of this hypothesis has been provided (1.8.2, Bonnet and Dick 1997, Marx 2003, Pardal et al 2003). Thus, studying the effects of tumourigenic events on aspects of stem cell function such as differentiation may provide further understanding of the mechanisms of cancer progression.

The progression of normal human cells into cancer cells is termed transformation, and results from the stepwise accumulation of genetic alterations that lead to loss of tumour suppressor gene function and abnormal activation of proto-oncogenes (Hahn and Wienberg 2002, Vogelstein and Kinzler 2004). A large body of research has described the cellular pathways perturbed during the progression of many cancers (1.8.3, reviewed in Hahahan and Weinberg 2000). As a result of such studies, in vitro models of cellular transformation have been developed (Hahn et al 1999, Rangarajan et al 2004). In vitro transformation involves the introduction of a defined set of genetic elements, termed here transforming "hits", into recipient cells. These include hTERT, oncogenic H-Ras (H-Ras^{V12}), viral oncogenes such as HPV16 E6 and E7 which inhibit the functions of p53 and pRb respectively, and SV40 ST antigen which inhibits c-Myc inactivation. The presence

(or mutation) of such genes is observed in numerous cancer sub-types: by way of example, p53 mutations are seen in more than 50% of all human cancers (Harris 1996), and more than 90% of human tumours exhibit telomerase activity (Shay and Bacchetti 1997). Therefore, such systems are likely to be accurate models of the molecular pathways involved in human cancer, and provide the opportunity to study molecular aspects of cancer progression *in vitro*.

Sarcomas are a group of soft tissue tumours that often resemble their tissue of origin, for example bone, adipose and cartilage (1.8.4, Helman and Meltzer 2003). As hMSCs are capable of differentiating into such tissue types (Pittenger *et al* 1999), it is a possibility that the accumulation of tumourigenic mutations in hMSCs may be the cause of sarcoma formation, associated with aberrant hMSC differentiation (Helman and Meltzer 2003). Thus, studies aimed at investigating whether *in vitro* transformation of these cells results in changes in their differentiation capability may provide answers to this question.

Several groups have demonstrated that long term culture of both murine and human MSCs can lead to spontaneous transformation and tumour formation in immunodeficient mice (1.8.4, Serakinci *et al* 2004, Burns *et al* 2005, Rubio *et al* 2005, Miura *et al* 2006). One study found that the resulting tumours (fibrosarcomas) contained a sub-population of cells that could form cell clusters in culture and further tumours after re-transplantation, strikingly similar to cancer stem cells, although this study was not performed at the single-cell level (Miura *et al* 2006). The genetic alterations resulting in this spontaneous transformation were found to involve elevated telomerase activity, c-Myc over-expression, and loss of some cell cycle regulators including the INK4A/ARF locus; these findings were consistent between different studies (Serakinci *et al* 2004, Rubio *et al* 2005, Miura *et al* 2006). Serakinci and co-workers also described a mutation causing activation of K-Ras (Serakinci *et al* 2004). This suggests that the use of step-wise transformation models outlined above would be a relatively accurate representation of the processes of hMSC transformation. These studies however did not give a full assessment of the differentiation capabilities of the transformed MSCs, although one group found that transformed murine MSCs exhibited some bone formation within tumours (Miura *et al* 2006).

There have been several studies that have investigated the effect of stepwise introduction of transforming genes on MSC immortalisation, transformation and differentiation (1.8.4). In two studies, the introduction of hTERT was found to immortalise hMSCs without reducing their differentiation potential even at high passages (Simonsen *et al* 2002, Shi *et al* 2002); in fact, osteogenesis was enhanced in one case (Shi *et al* 2002). The effect of 2 (E6 and E7, Hung *et al* 2004b) or 3 (hTERT, E6 and E7, Okamoto *et al* 2002) hits on hMSCs have been investigated: it was found that osteogenic, chondrogenic and adipogenic differentiation were unaffected. However, these cells were only immortalised and not transformed. It may be that further hits that lead to hMSC transformation could be associated with changes in differentiation potential, but this theory is as yet unexplored.

Aims

The main aim of this chapter is to determine the effect of transformation on hMSC differentiation. Human MSCs containing increasing numbers of transforming genes to a maximum of 5 hits will be employed. To achieve this aim:

- The morphology and growth properties of these cell lines will be investigated in order to optimise conditions for differentiation experiments

- The ability of each cell line to undergo both adipogenesis and osteogenesis will be determined using several techniques.

6.2 Characterisation of hMSCs after the introduction of transforming hits

In vitro cell transformation involves the ectopic expression of a limited selection of genes which induce the deregulation of major cellular pathways and the ultimate production of cancer-like cells (Hahn *et al* 1999, Rangarajan *et al* 2004). The introduction of such genes into hMSCs would, in all probability, have significant effects on properties such as proliferation and morphology. The aim of these experiments was therefore to examine the effects of the addition of transforming hits on several cellular properties, in order to define experimental conditions for differentiation studies.

Transformation of hMSCs was accomplished using retroviral vectors separately encoding five genes: hTERT, HPV16 E6 and E7, SV40 ST and H-Ras^{V12}. Through disruption of the function of cellular genes, the expression of these exogenous genes, or transforming "hits", results in cellular immortalisation, inhibition of apoptosis, deregulated cell cycle and increased growth, motility and proliferation (Fig 1.7 and Fig. 6.1). Each transforming hit was introduced successively into hMSCs in a defined order through retroviral infection and drug selection. The expression of each transforming hit, or their downstream targets, was confirmed via Western blotting and microarray analysis of gene expression; this work was undertaken by J. Funes (Funes *et al* 2006, submitted). The resulting cell lines (the word "line" will subsequently be used for ease of explanation, but it should be noted that these populations are not true cell lines as they are not clonally derived) contained several combinations and increasing numbers of transforming hits, ranging from 0-5 (Fig. 6.1). In work also performed by J. Funes and others, it was found that the expression of hTERT (1 hit) was sufficient to immortalise hMSCs in this system. Additionally, 5 hit cells showed robust transformed characteristics, giving rise to colonies in soft agarose and tumours in immunodeficient mice. Other cell lines showed partial characteristics of transformation: all E7-containing cell lines formed colonies in soft agar with 4 hit (hTERT, E6, E7, ST) cells showing colonies soonest (Funes *et al* 2006, submitted). The 4(E7) cells (hTERT, E7, ST, H-Ras) gave rise to large colonies in soft agar and tumours in 2 out of 6 mice, indicating that they expressed the expected array of transforming hits, although this has not yet been confirmed via Western blotting (Funes *et al*, unpublished observations).
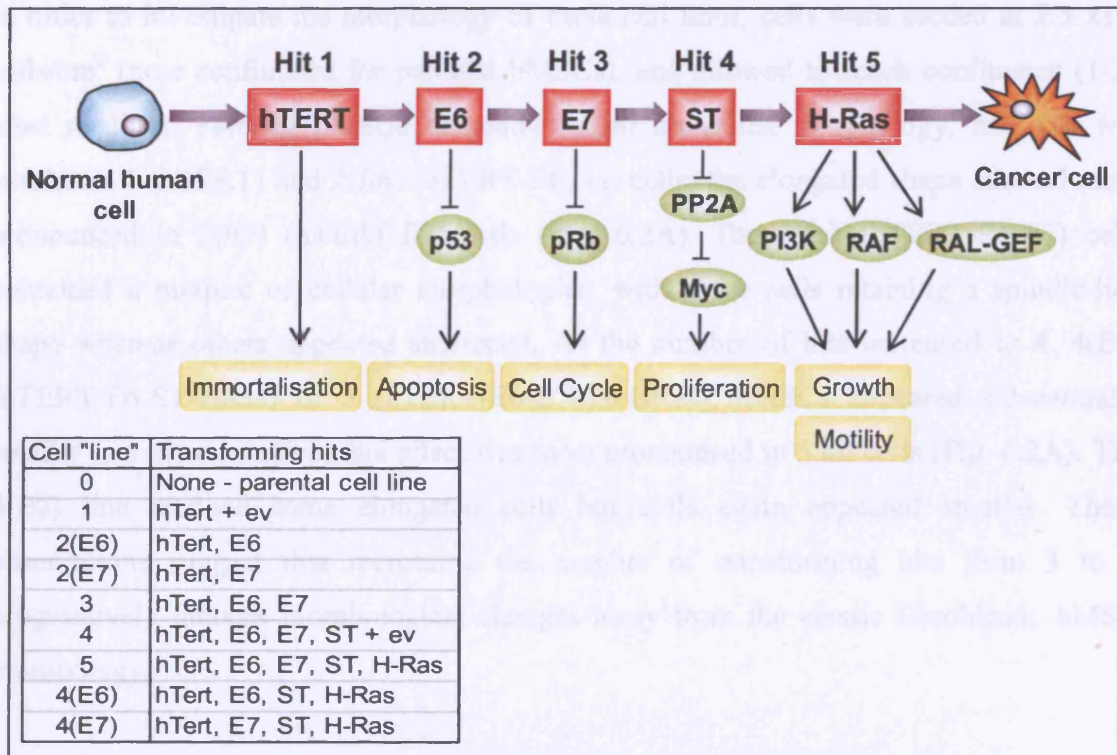
Figure 6.1 Schematic view of stepwise *in vitro* transformation system and cell "lines" created for study. Successive introduction of viral or human proteins, known as transforming "hits", results in the inhibition of tumour suppressor gene function or constitutive activation of oncogenes, leading to activation of pathways required for cellular transformation and ultimately the production of a cancer cell. hTERT introduction leads to telomere stabilisation and immortalisation. pRb and p53 are inactivated by viral oncoproteins E7 and E6 respectively, leading to bypass of apoptosis and cell cycle pathways. SV40 ST inactivates PP2A, leading to stabilisation of Myc and increased proliferation and angiogenic properties. Oncogenic RAS (V12 H-Ras) transmits constitutive signals through several pathways, leading to increased growth, motility, and angiogenesis. Defined cell populations containing increasing numbers or different combinations of "hits" were created for functional investigation. ev: empty vector. These cell "lines" were produced by Dr. J. Funes.

In order to investigate the morphology of these cell lines, cells were seeded at 2.5 x10^4 cells/cm^2 (near confluence for parental hMSCs), and allowed to reach confluence (1-2d after seeding). Parental hMSCs maintained their elongated morphology, and this was similar in 1 (hTERT) and 2(E6) (hTERT-E6) hit cells; the elongated shape seemed more pronounced in 2(E7) (hTERT-E7) cells (Fig. 6.2A). The 3 hit (hTERT-E6-E7) cells contained a mixture of cellular morphologies, with some cells retaining a spindle-like shape whereas others appeared shortened. As the number of hits increased to 4, 4(E6) (hTERT-E6-ST-HRas) or 5 (hTERT-E6-E7-ST-HRas), hMSCs appeared substantially smaller and more rounded; this effect was most pronounced in 5 hit cells (Fig. 6.2A). The 4(E7) line retained some elongated cells but cells again appeared smaller. These observations suggest that increasing the number of transforming hits from 3 to 5 progressively induces morphological changes away from the classic fibroblastic hMSC morphology.

Cell counts were next performed at confluence and after incubation of confluent cells without passaging for 7 days (2.2.6), in order to assess proliferation rates and the extent of contact inhibition in each cell line. Parental hMSCs and 1 and 2(E6) hit cells showed similar cell numbers at confluence, ranging from 1.9-2.8 x10^4 cells/cm^2 (Fig. 6.2B), which could in part be due to their similar morphology and size. The 4(E6) population also showed similar density at confluence (2.9 x10^4 cells/cm^2). Cell numbers for the parental hMSC, 1 hit and 4(E6) cell lines did not increase significantly after 7d (p>0.05, two-sample Student's t-test), indicating that these cells were sensitive to contact inhibition, and consequently growth arrest was induced after confluence. Interestingly, 2(E7) cells yielded 4.6x10^4 cells/cm^2 at confluence, double the number of their hTERT-E6 counterparts and also more than 3 hit cells, indicating that the presence of E7 (thus loss of Rb) may induce a higher proliferation rate and/or smaller cell size (Fig. 6.2B, and Funes *et al* 2006, submitted). In both 2(E7) and 3 hit hMSCs, cell numbers increased significantly from confluence to 7d (p<0.05 and p<0.01, respectively), indicating a lower responsiveness to contact inhibition. More than twice as many 4 hit cells (4.7 x10^4 cells/cm^2) and more than three times as many 5 hit cells (6.0 x10^4 cells/cm^2) were present at confluence compared with parental hMSCs (1.9 x10^4 cells/cm^2), which may be due to their smaller size allowing a higher number of cells to be accommodated in a monolayer,
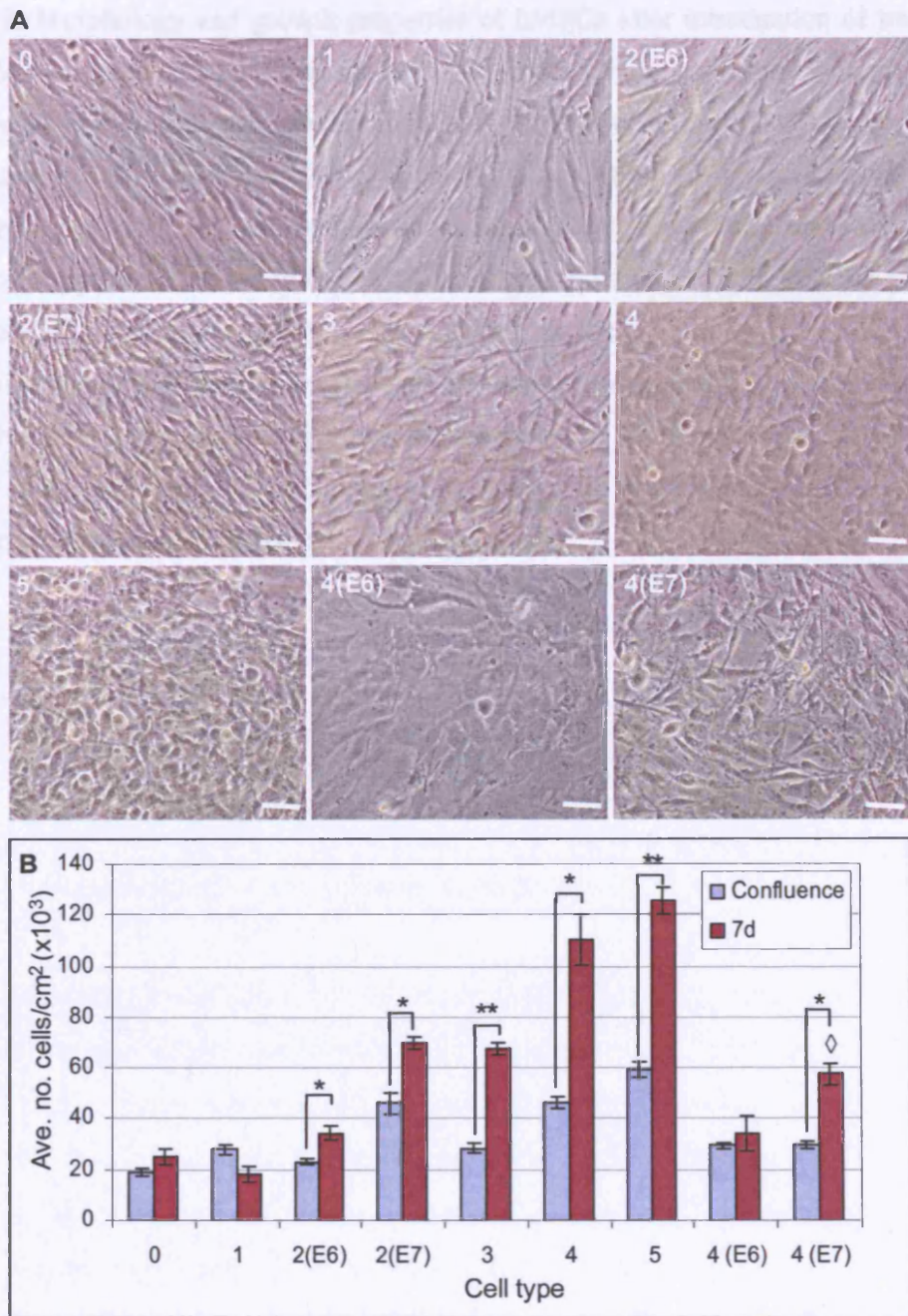
Figure 6.2. Continued overleaf.

Figure 6.2. Morphology and growth properties of hMSCs after introduction of transforming hits. hMSCs were successively transduced with virus carrying each transforming gene then drug-selected, producing cell populations with different numbers and combinations of transforming hits (J. Funes *et al* 2006, submitted). A, Cells were grown to confluence and their morphology analysed via phase contrast microscopy. Scale bars represent 150μm; B, cells of each type were seeded in 6-well plate format and counted at confluence and 7d after confluence (7d). Counts were performed in duplicate for each condition (2.2.6). Data represent number of cells/cm^2 and are mean ± standard deviation. *, $p<0.05$ and **, $p<0.01$ using a two-sample Student's t-test (assuming equal variance). ◊, at this stage, some 4(E7) cells had detached from the plate but remaining cells were still proliferating, and the 5 hit cell monolayer were beginning to detach.

but could also indicate a higher proliferation rate to achieve this cell number in this time. Furthermore, significant increases in cell number of 4 (p<0.05) and 5 (p<0.01) hit cells was observed at 7d, suggesting that cell-to-cell contact did not inhibit proliferation after confluence (Fig. 6.2B). Additionally, 4(E7) cells were observed to proliferate rapidly during this time, and by 6-7d the cell monolayer had detached from the plastic surface. Remaining adherent 4(E7) cells continued to proliferate, but detached cells may have been lost during harvesting for cell counts, resulting in unrepresentatively low counts. However, the difference between the number of cells present at 0d and 7d was still statistically significant (p<0.05), even after loss of the cell monolayer. At this time 5 hit cells were also beginning to detach (see also Fig. 6.3, arrowheads). These results suggest that 2(E7), 3, 4, 5 and 4(E7) cell lines may have an increased proliferation rate and decreased responsiveness to contact inhibition; this would correlate with the transformed state of 5 and possibly 4(E7) cells. In contrast to this, 1 hit and 4(E6) cells, and also 2(E6) cells (but only in terms in terms of their morphology, as their cell number increased significantly between 0 and 7d (p<0.05)) remain more similar to parental hMSCs, indicating that the number or combination of transforming hits had not produced a large effect on their morphology or loss of contact inhibition.

During cell culture, it was also noted that the colour of the growth media, which contained phenol red, became progressively more yellow as the number of hits increased (Fig 6.3). This indicates a higher media acidity in cells with more transforming hits, which was likely to be due to increased metabolic products (e.g. lactate) caused by higher cell numbers, rather than from increased metabolic activity of each cell, as it was shown that glucose uptake and lactate production (markers of the rate of glycolysis) are not significantly higher in transformed hMSCs that in parental hMSCs (Funes *et al* 2006, submitted). Interestingly, this effect was less pronounced in cells incubated in osteogenic media, and even less so in adipogenic media, suggesting that components of these media cause a reduced rate of metabolism or proliferation.

As a cumulative result of the observations of a higher ability to continue proliferating after confluence and high media acidity produced by 4 and 5 hit cells, it was decided that for differentiation experiments, media would be changed every other day instead of every 3

Figure 6.3. Media pH for each cell line at day 9 after seeding, indicated by phenol red colour. Media colour was visualised in normal growth conditions or after adipogenic or osteogenic induction. Yellow colours indicate an acidic pH, whereas red indicates a higher pH. Arrowheads indicate detached cell monolayers at day 9. X: Cells surplus to this study.

days as in all previous experiments. This was performed in an attempt to prevent metabolic by-products from accumulating in the media and becoming toxic, or mutagenic, to the cells. Furthermore, differentiation would be continued for as short a time as necessary in an attempt to avoid cell detachment during the process.

6.3 The effect of transforming hits on hMSC adipogenesis

Having established some of the properties of hMSCs containing different numbers of transforming hits, experiments were set up with the aim of investigating the ability of each cell line to differentiate into adipocytes.

Each cell type containing different transforming hits was seeded at 2.5 $\times 10^4$ cells/cm^2 in 12-well plate format and then left for 24h to achieve full confluence. Parental hMSCs were at passage 8 whereas the passage number of other cell lines ranged from 21-29 (Fig. 6.4); this reflected the fact that several rounds of viral transduction and drug selection were required to produce cells containing higher numbers of hits, inevitably resulting in higher passage numbers for cells with more hits. Furthermore, the faster proliferation rate of 3-5 hit cells entails a need for more frequent passaging. The experimental format was as follows: a duplicate set of confluent cells (i.e. cells containing the full range of transforming hits) was induced to undergo differentiation using adipogenic media (2.2.2.1), whereas two further sets of cells were incubated in normal growth media (without bFGF) for the duration of the experiment. Media was changed every other day. RNA was extracted from one set of control cells immediately after confluence, and was extracted from one set of adipocytes when differentiation was terminated (2.6.1). This format was repeated in separate experiments, and the results of a single representative experiment are shown here. During adipogenic differentiation no cell detachment was observed, whereas 4(E7), 5 and 2(E7) hit cells detached during this time in control medium. This indicates that the adipogenic cocktail has potent growth inhibitory properties, a process required for adipogenesis (Rosen and Spiegelman 2000).
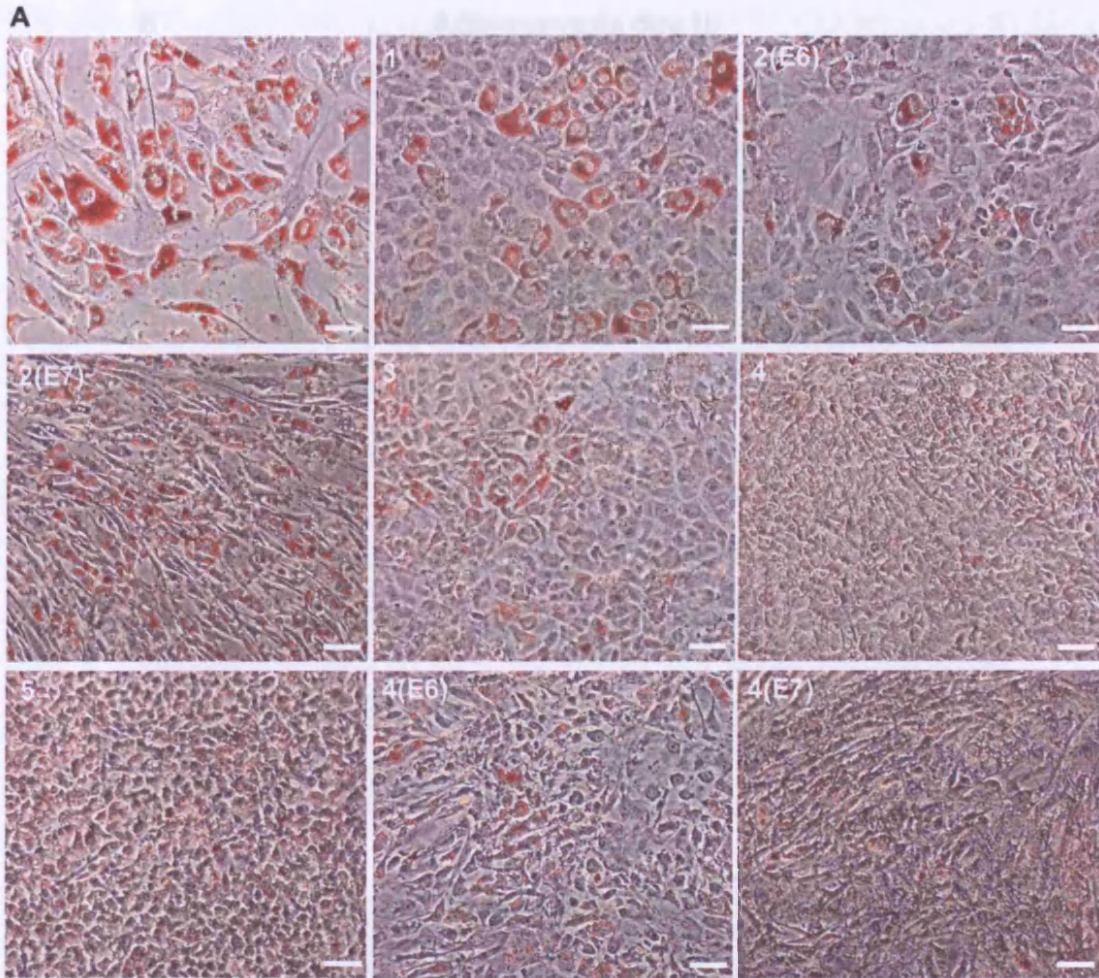
Figure 6.4. Effect of successive transforming hits on hMSC adipogenesis. Post-confluent cells of each type were induced to differentiate into adipocytes using the classic induction cocktail (2.2.2.1). Differentiation was continued for 10 days, with media changes every other day; parallel cells were supplemented every other day with normal growth medium. At the end of differentiation cells were fixed and stained with Oil Red O and visualised via phase contrast microscopy (2.2.3.1). Passage numbers: MSCs, p8; 1 hit, p21; 2(E6), p21; 2(E7), p25; 3, p24, 4, p27; 5, p29, 4(E6), p22; 4(E7), p29; A, a representative field of view indicating the extent of differentiation in each cell line. Scale bars represent 150μm; B (overleaf), High magnification images of differentiated cells. Scale bars represent 150μm; C (overleaf), Control cells incubated in hMSC growth media (without bFGF) showed no detectable differentiation. Scale bars represent 150μm. *, cell monolayers had detached in the duration of the experiment, but remaining adherent cells continued proliferating.

B   Adipogenesis day 10
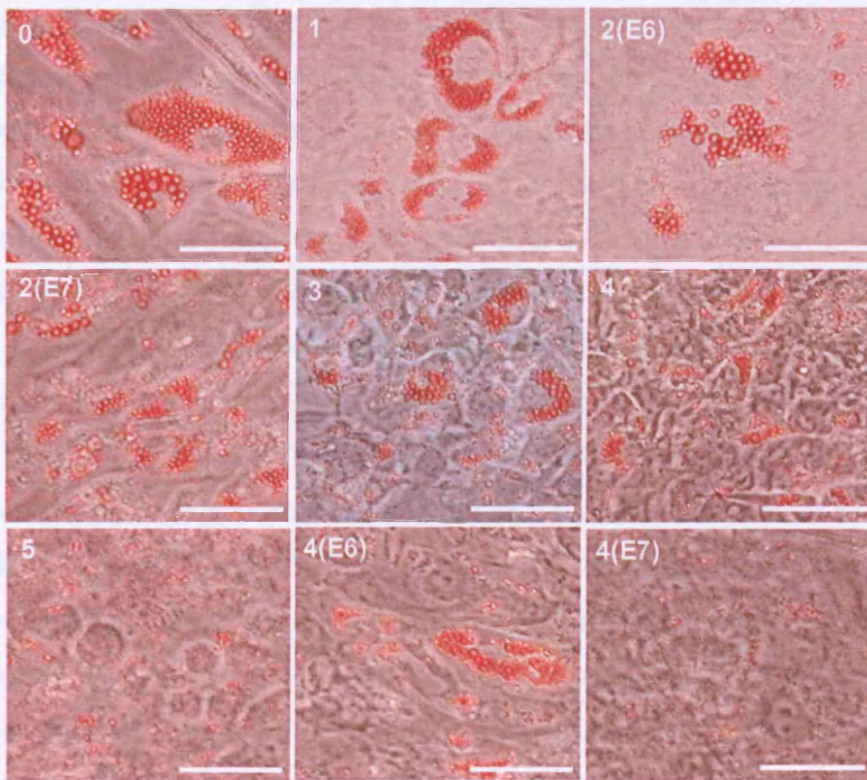
C   No differentiation
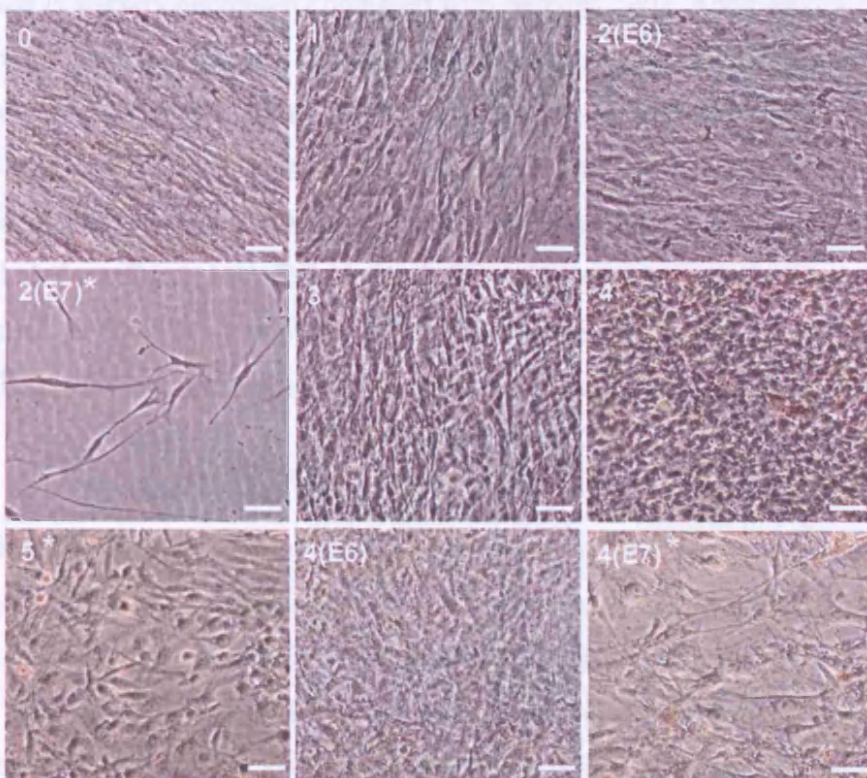
At day 10, differentiation was terminated and cells that were not used for RNA extraction were stained with Oil Red O in order to visualise lipid accumulation (2.2.3.1, Fig. 6.4). Accurate quantification of adipogenesis was not possible due to the morphology of the 4 and 5 hit cell lines causing inaccurate quantification of the proportion of cells containing lipid. Furthermore, the number of lipid droplets per cell, as well as the number of cells that differentiated, was a major factor for assessment, so a qualitative assessment was deemed necessary, and was able to account for both of these factors. Whereas no differentiation was observed in any of the control cell populations incubated in growth media only (Fig. 6.4C), parental hMSCs induced to differentiate accumulated multiple lipid droplets in the majority of cells (Fig. 6.4A and B). Similar lipid-laden adipocytes were seen in 1 hit cells, though they were fewer in number than with parental hMSCs. The number of adipocytes decreased further in the 2 and 3 hit cell lines, as did the extent of lipid accumulation, although multiple lipid droplets were still present in differentiated cells. Occasional lipid-filled adipocytes were also visible in the 4(E6) cell line, although they were not common. These results indicate that cells containing 1-3 transforming hits are still capable of forming mature adipocytes containing multiple lipid droplets, but the extent of this differentiation is slightly decreased compared with parental cells. Additionally, 4(E6) cells are still capable of differentiating into adipocytes containing multiple lipid droplets, and the extent of differentiation is similar to that of the 2 or 3 hit cells, indicating that E7 (Rb) may have a more potent inhibitory affect than E6 (p53) on adipogenic ability. In the 4 hit cell line, cells containing several lipid droplets were visible, but the lipid often only partially filled the cells and no fully laden adipocytes were observed. A more dramatic effect was seen in the 5 hit and 4(E7) population, where the majority of cells contained only a few lipid droplets, indicating that their ability to accumulate lipid was significantly impaired.

To further investigate the effect of hMSC transformation on adipogenesis, RT-PCR was employed to analyse the expression of aP2, PPARγ and C/EBPα (adipogenic markers, section 1.6.2) in control and differentiated cells containing different hits (2.6.3.1, Fig 6.5A). Low but detectable expression of all three genes was seen in most undifferentiated cell lines, and PPARγ mRNA expression showed a slight increase in undifferentiated control cells with 2 or more hits, excluding 4(E7) cells. Expression of these genes was

260
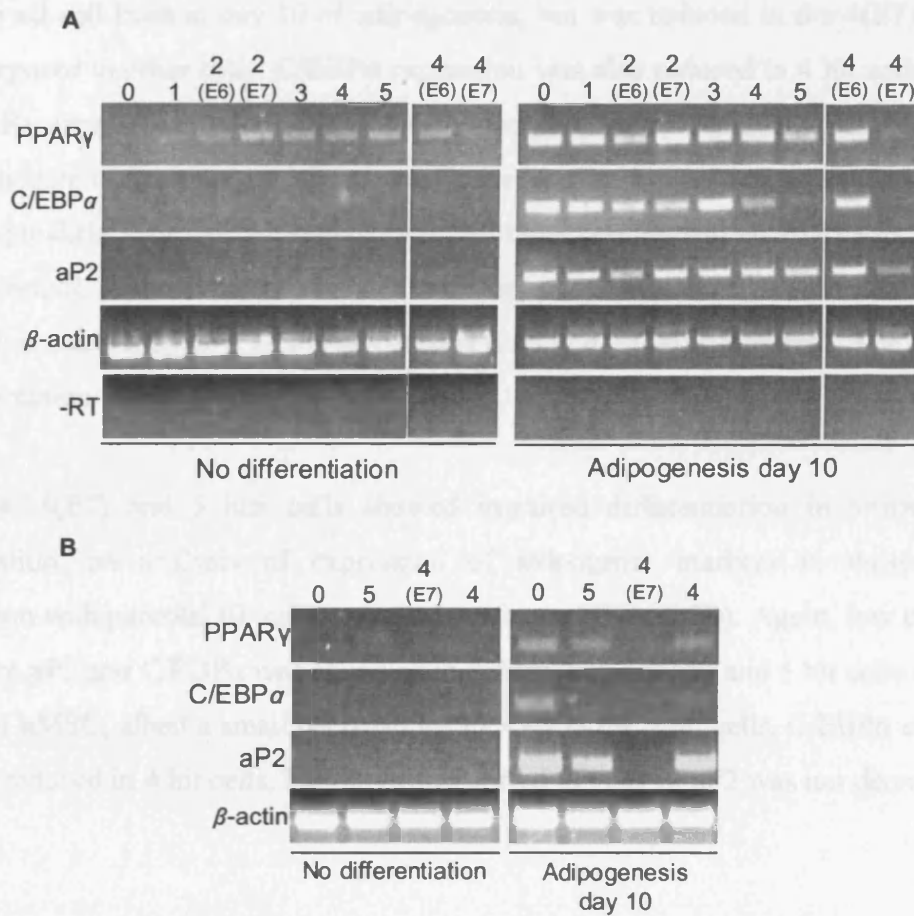
Figure 6.5. Effect of successive transforming hits on mRNA levels of adipogenic markers. RNA was collected from day 10 differentiated cells, and control cells one day after reaching confluence (2.6.1). cDNA was synthesised (2.6.2) and gene-specific primers were used for PCR to determine adipogenic marker expression in each cell type after differentiation (2.6.3.1). A, expression of adipogenic genes in all cell lines under investigation. β-actin, internal control; -RT: no reverse transcription control (osteocalcin primers); B, focused analysis of gene expression in 4, 4(E7), and 5 hit lines compared with parental cells. PCR cycle numbers used: β-actin: 25, all other genes: 28.

higher in all cell lines at day 10 of adipogenesis, but was reduced in the 4(E7) and 5 hit cells compared to other cells. C/EBPα expression was also reduced in 4 hit cells, and aP2 and PPARγ expression appeared lessened in the 2(E6) hit line, indicating that loss of p53 may contribute to the negative effect on adipogenic gene expression. In these analyses β-actin (a cytoskeletal protein present in most cell types) rather than GAPDH was used as an internal control, as it had been found that transformation affected metabolic pathways such as glycolysis, in which GAPDH is involved (Funes et al 2006, submitted). The level of β-actin expression was comparable in all cell lines.

As the 4, 4(E7) and 5 hits cells showed impaired differentiation in terms of lipid accumulation, an analysis of expression of adipogenic markers in these cells in comparison with parental (0) cells was next performed (Fig. 6.5B). Again, low expression of PPARγ, aP2 and C/EBPα was observed in differentiated 4(E7) and 5 hit cells compared to normal hMSC, albeit a small decrease for PPARγ in the 5 hit cells. C/EBPα expression was also reduced in 4 hit cells, but the expression of PPARγ or aP2 was not decreased in 4 hit cells.

Taken together, these results suggest that the introduction of four transforming hits, including hTERT, E7, ST and H-Ras, or the additional introduction of E6, does not abolish hMSC adipogenesis, but impairs the process, as evidenced by decreased lipid accumulation and adipogenic marker gene expression.

An important factor to be taken into account in these experiments was the passage number of the cells used. As previously stated, whereas hMSCs were used at around passage 8, all cells containing transforming hits had undergone further rounds of passaging. Untransformed hMSCs lose their ability to differentiate at higher passage numbers, although the exact passage number at which this occurs is variable between reports (Conget and Minguell 1999, Tsutsumi et al 2001, Fig. 3.5). It is therefore possible that passage number, as well as the introduction of genes required for transformation, may affect differentiation. To address this issue, it was necessary to perform differentiations of parental and transformed cells of similar passages, and compare this with differentiation as higher passages, to determine whether the process is affected. The only cell line available

at a low passage was hMSC-hTERT (not containing empty vector). As stated, it was not possible to expand hMSCs to a high passage, so differentiation of hMSCs and hMSC-hTERT was carried out at passage 6 and compared to 1 hit cells differentiated at higher passages (p19, p21 and p22, Fig. 6.6). Low passage hMSCs and hMSC-hTERT differentiated to a similar extent at passage 6, and this was comparable to differentiation of 1 hit cells at higher passages ($p > 0.05$, two sample Student's t-test), although here differentiation and possibly also the rate of apoptosis was more variable between samples indicating that passage number may have a slight effect. Indeed, 1 hit cells shown in Figure 6.4 appeared to have a reduced rate of differentiation compared to parental hMSCs, and less apoptosis appeared to have occurred, indicating that the spontaneous accumulation of further mutations could not be ruled out in some experiments.

6.4 The effect of transforming hits on hMSC osteogenesis

hMSCs are multipotent, and hence are capable of differentiating into several cell types (Pittenger *et al* 1999, Fig. 3.3). Therefore, the effect of hMSC transformation on osteogenesis was also investigated in addition to their differentiation to adipocytes.

Experiments were set up in an identical manner to adipogenesis as follows: six sets of cells containing each combination of transforming hits were seeded and grown to confluence. Passage number was identical to adipogenesis experiments (Fig. 6.7). One day post-confluence, osteogenesis was induced in three set of cells (2.2.2.2) whereas remaining cells (3 sets) were incubated in normal growth media (without bFGF). RNA was extracted from one set of control cells immediately after confluence, and was extracted from one set of hMSCs differentiated to osteoblasts when differentiation was terminated. The remaining sets of cells (2 per condition) were used to assess osteogenesis via histochemical staining. This format was repeated in separate experiments, and the results of a representative experiment are shown. Media was changed every other day, and initially osteogenesis was continued for 14d. The normal duration of osteogenic differentiation is 21d (Pittenger *et al* 1999), but it was found that by this stage the 2(E7)

263
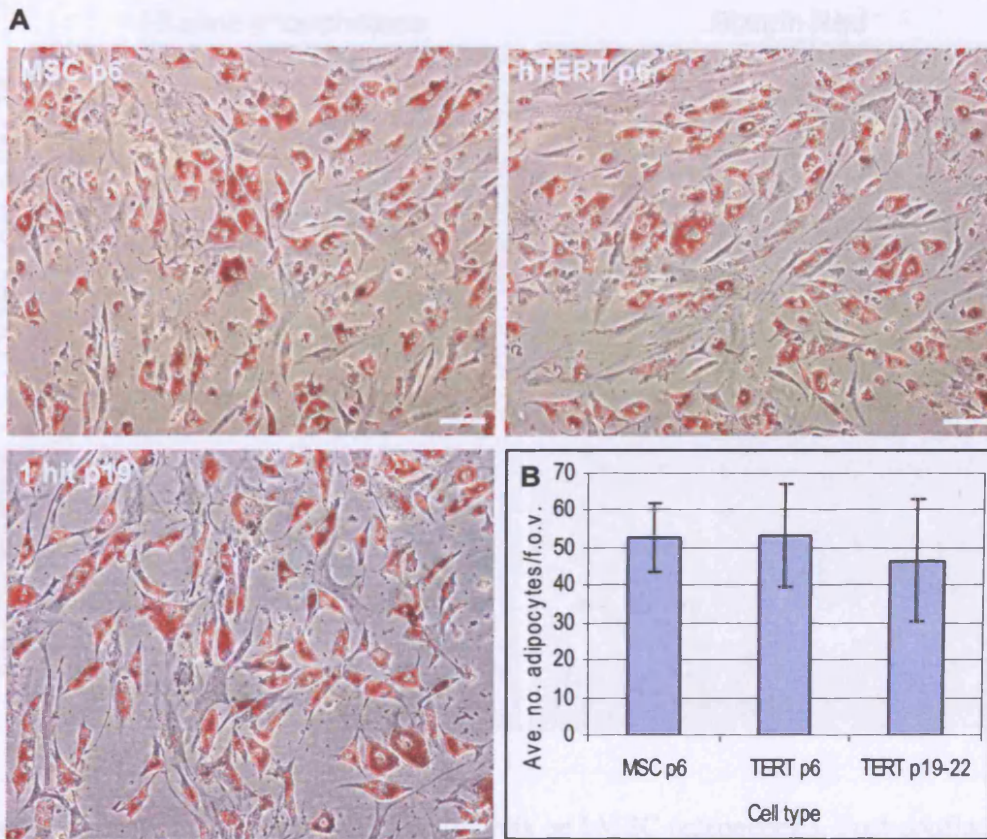
Figure 6.6. Comparison of differentiation rates of parental hMSCs and hMSC-hTERT cells at low and high passage numbers. hMSCs, with or without hTERT, were induced to differentiate to adipocytes at passage 6, as were 1 hit cells at higher passages (2.2.2.1); A, At day 10 of differentiation cells were fixed and stained with Oil Red O (2.2.3.1). Scale bars represent 200μm; B, Differentiation was quantified by counting the number of adipocytes per field of view for 3 fields of view per condition. Data represent number of adipocytes per field of view and are mean ± SD.
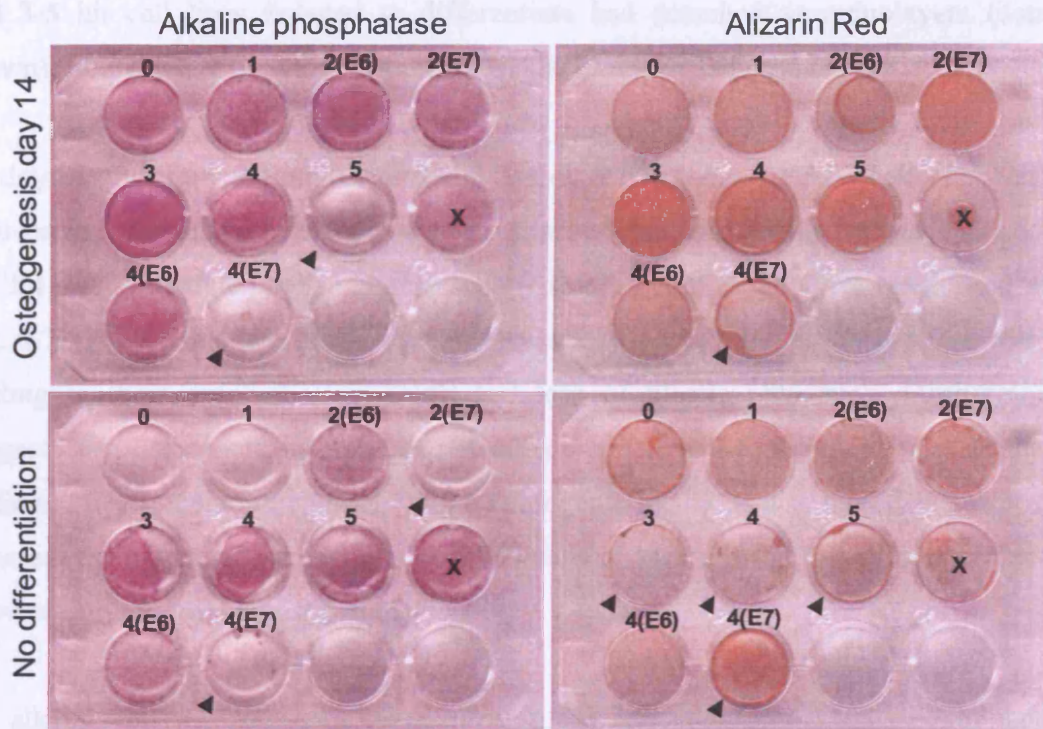
Figure 6.7. Effect of successive transforming hits on hMSC osteogenesis. Post-confluent cells containing each combination of transforming hits were induced to differentiate into osteoblasts using a classic induction cocktail (2.2.2.2). Differentiation was continued for 14 days, with media changes every other day; parallel cells were supplemented every other day with normal growth medium. At the end of differentiation cells were fixed and stained with alizarin red (2.2.3.3) to visualise calcium deposition, or assayed for alkaline phosphatase activity (2.2.3.2). Passage numbers: hMSCs, p8; 1 hit, p21; 2(E6), p22; 2(E7), p25; 3, p24, 4, p27; 5, p29, 4(E6), p22; 4(E7), p29; Black arrowheads indicate cell lines where monolayers were lost during experimentation. X indicates surplus wells not required for this study.

and 3-5 hit cell lines induced to differentiate had detached as monolayers (data not shown), likely due to higher growth rates and lack of growth arrest in such cells.

By day 14 of differentiation, it was found that all 4(E7) and some 5 hit cells that had been induced to differentiate to osteoblasts had detached from the plastic surface, as had some 3-5 hit undifferentiated cells, as assessed by floating cell monolayers (excluding 4(E6), Fig. 6.7). Again, this is probably due to the lack of contact inhibition in these cell lines leading to their eventual overgrowth and loss of plastic adherence. Furthermore, it suggests that osteogenic media does not induce strong anti-proliferative signals that are sufficient to overcome this effect, unlike adipogenic media. It was therefore not possible to assess the extent of osteogenesis in 4(E7) cells at 14d, and the loss of some 5 hit cells allowed only a partial investigation.

An alkaline phosphatase activity assay (2.2.3.2) and alizarin red staining of calcium deposition (2.2.3.3) were used to assess osteogenesis where possible. Undifferentiated cells stained for alkaline phosphatase activity showed a background level of staining that became more intense as the number of hits increased (Fig. 6.7). This may have been due to some spontaneous differentiation to osteoblasts in these cells, but it should be noted that these staining techniques are not controlled for cell number, and it is not possible to accurately quantify these osteogenic stains at a single-cell level. Therefore, increased staining intensity may be due to increased cell number, which occurs as the number of hits increases (Fig. 6.2B), making comparisons between cell lines in each condition inaccurate. However, alkaline phosphatase activity was higher for each cell line when differentiation was induced than in undifferentiated cells, indicating that a greater level of osteogenesis had occurred in these cells. It was not possible to assess calcium deposition using alizarin red staining in 3-5 hit control cells (Fig. 6.7) in the experiment illustrated, but in equivalent experiments where 3-4 hit control cells could be stained with alizarin red at day 14, an increase in staining intensity was seen with increasing hits, similar to the observations with alkaline phosphatase activity (data not shown). Where possible, it was observed that alizarin red staining was more intense in differentiated cell lines compared to their non-differentiated counterparts, apart from parental cells and 1 hit cells where there was minimal difference (Fig. 6.7).

RT-PCR analysis of osteogenic markers was next performed to assess the extent of osteogenesis on the mRNA level. As the amount of RNA used in each reaction is normalised, this allows comparison between different cell lines. The genes studied were: the main osteogenic transcription factor CBFA1 (Ducy *et al* 1997); PLZF, a transcription factor that is expressed upstream of CBFA1 during osteogenesis (Ikeda *et al* 2005, Fig. 6.8A) and COL1A1, alkaline phosphatase (tissue non-specific, also known as the bone/liver/kidney isoform) and osteocalcin, which are expressed at high levels in differentiated osteoblasts (Ducy *et al* 2000). It was not possible to assess gene expression in day 14 4(E7) differentiated cells due to their consistent detachment, but RNA was collected from 5 hit differentiated cells.

Expression of the internal control, β-actin, was similar in all cell lines. CBFA1 expression was observed in most undifferentiated cell lines excluding parental hMSCs and 4(E7) cells (Fig. 6.8B). Expression of this gene increased after differentiation in parental cells, but no significant increase in CBFA1 expression was seen in other cell lines. Similarly, osteocalcin expression was observed in some undifferentiated cells, specifically in the 2(E6) and 3 hit cell lines, and slight increases in mRNA levels were observed in parental and 2(E7) hMSCs but not cells containing 4 (all combinations) or 5 hits. COL1A1 expression was detected in all undifferentiated cells and did not increase greatly after differentiation in any cell line. Both alkaline phosphatase and PLZF showed higher expression in each differentiated cell line compared to undifferentiated controls, with the exception of alkaline phosphatase expression in 5 hit cells which appeared similar in both conditions. In summary, a single pattern does not describe the expression of all of the genes analysed, either after differentiation or between cell lines. Increases in CBFA1, PLZF and ALP expression occurred during differentiation of parental hMSCs; osteocalcin and COL1A1 showed a slight increase in gene expression. Furthermore, alkaline phosphatase activity was higher in differentiated than undifferentiated hMSCs, although calcium deposition was not detectably stronger. The observations of increased alkaline phosphatase staining and expression of several osteogenic markers in parental hMSCs is similar to previous reports of hMSC osteogenesis (Pittenger *et al* 1999, Gronthos *et al* 2003a), which supports this system as a reliable model of differentiation; the lack of a large increase in calcium deposition seen here during differentiation compared to definite
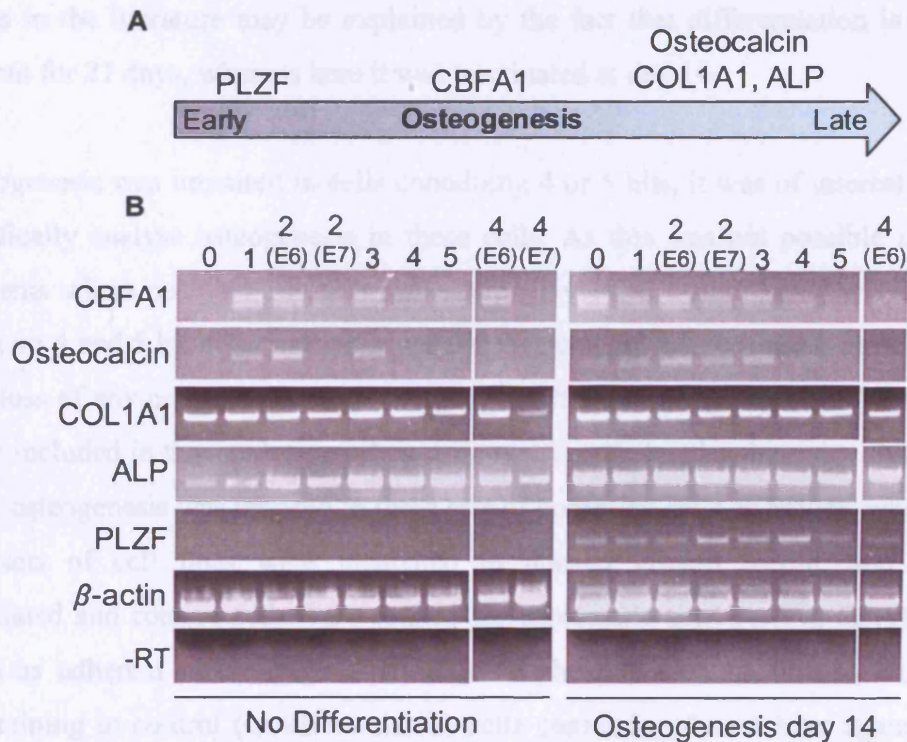
Figure 6.8. Effect of successive transforming hits on mRNA levels of osteogenic markers. A, Schematic of osteogenesis highlighting the genes analysed in this study and their time of expression during differentiation; B, RNA was collected (2.6.1) from day 14 differentiated cells and control cells one day after reaching confluence. cDNA was synthesised (2.6.2) and gene-specific primers for CBFA1, osteocalcin, COL1A1, ALP (alkaline phosphatase, tissue non-specific) and PLZF were used in PCR to determine osteogenic marker expression in each cell line after differentiation (2.6.3.1). –RT: no reverse transcription control (osteocalcin primers). PCR cycle numbers used: β-actin, COL1A1: 25; ALP, PLZF: 28; osteocalcin, CBFA1: 30.

increases in the literature may be explained by the fact that differentiation is classically carried out for 21 days, whereas here it was terminated at day 14.

As adipogenesis was impaired in cells containing 4 or 5 hits, it was of interest to attempt to specifically analyse osteogenesis in these cells. As this was not possible in previous experiments where osteogenesis was assessed at day 14, a further study was undertaken focusing on 4 and 5 hit cells, and osteogenesis was carried out for only 6 days in order to prevent loss of any cell lines through overgrowth after continued culture. The 4(E7) cells were not included in this analysis as they were continually lost by the 6 day timepoint. To this end, osteogenesis was induced in three sets of confluent cells as before, whereas three further sets of cell lines were incubated in normal growth media, and at day 6 differentiated and control cells were stained for osteogenic activity (Fig 6.9A). All cells survived as adherent monolayers. Both alkaline phosphatase and alizarin red revealed strong staining in control (undifferentiated) cells containing 4 or 5 hits; again this may have been due to an increase in cell number between lines causing differences in overall staining intensity. Nevertheless, the extent of osteogenesis can theoretically be assessed through comparison of induced and control cells of a specific line. However, no difference in staining intensity was seen in any condition between control and differentiated cells, apart from in the parental hMSCs where alkaline phosphatase activity seemed greater in differentiated cells (Fig 6.9A). This was possibly due to the time at which cells were stained; as osteogenesis is normally carried out for 21 days, it is likely that at 6 days osteogenesis is still in it's early stages and would not have progressed enough to produce differences in induced cells compared to control cells. Thus, due to the growth properties of transformed cells, it was not possible to conclusively assess osteogenesis in this system through staining techniques.

As before, RNA was extracted from one set each of control and day 6 differentiated cells (Fig. 6.9B), and a representative selection of three osteogenic markers were analysed for their expression using RT-PCR. In this set of experiments, expression of all three genes was detected in undifferentiated parental hMSCs, as well as cells containing 4 or 5 hits, indicating that these genes are normally expressed to some extent in undifferentiated hMSCs. However, expression of CBFA1, osteocalcin and COL1A1 was comparable in

269
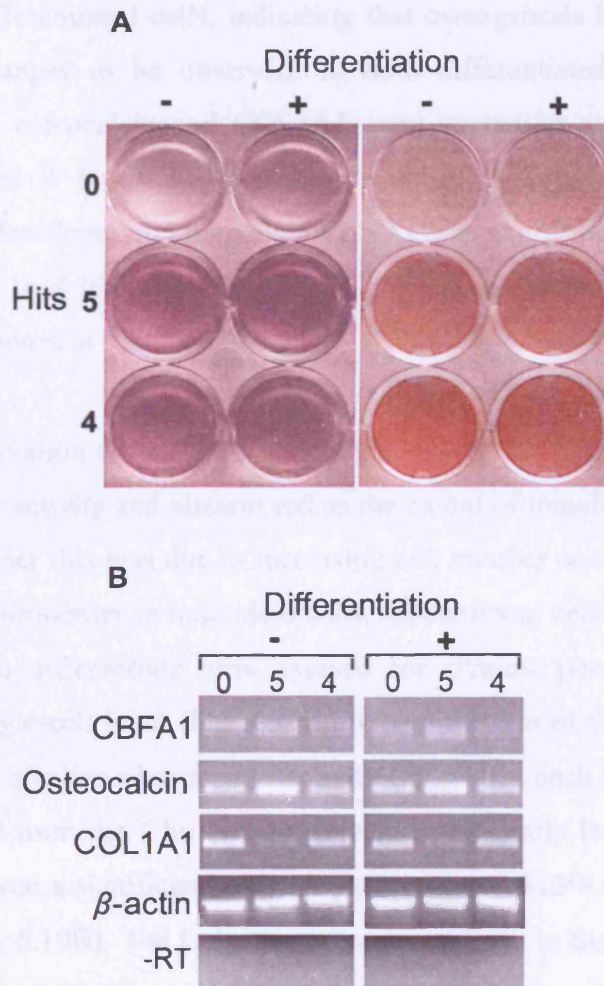
Figure 6.9. Focused analysis of the effect of four or five transforming hits on osteogenesis. Differentiation was carried out for 6 days in an attempt to minimise cell over-proliferation and loss of adherence, after which time control and differentiated cells were stained and RNA was extracted. A, alkaline phosphatase activity assay (2.2.3.2) and alizarin red stain for calcium deposition (2.2.3.3) on undifferentiated cells (-) and cells induced to differentiate (+) (2.2.2.2); B, RT-PCR analysis (2.6.3.1) of expression of a selection of osteogenic marker genes and –RT control (osteocalcin primers). PCR cycle numbers used: β-actin: 25; COL1A1: 20; osteocalcin, CBFA1: 32.

both control and differentiated cells, indicating that osteogenesis had not progressed far enough for any changes to be observed. In both differentiated and undifferentiated conditions, CBFA1, osteocalcin and COL1A1 were expressed at comparable levels in parental hMSCs and 4 hit cells. However, whereas CBFA1 and osteocalcin were expressed to the highest level in 5 hit cells, COL1A1 was expressed at lower levels in the 5 hit line compared to 4 hit or parental hMSCs. Thus, no single pattern described the expression of these genes in each cell line.

One consistent observation during experimentation was an increase in staining with both alkaline phosphatase activity and alizarin red as the extent of transformation increased. In order to assess whether this was due to increasing cell number or an inherent increase in levels of osteogenic properties in individual cells, subconfluent cells of each type that had not been induced to differentiate were assayed for alkaline phosphatase activity and visualised at the single-cell level (Fig. 6.10A). Quantification of the percentage of cells staining positive for alkaline phosphatase activity (ALP+) in each cell line revealed that every cell line, apart from the 1 hit (which showed significantly less ALP+ cells), 2(E7) and 4(E7) cells showed a significantly higher percentage of ALP+ cells than the parental hMSC cell line (Fig. 6.10B). The largest increases were seen in the 2(E6), 3, 4 and 5 hit cells. This correlates with the higher expression levels of certain osteogenic genes observed in the undifferentiated 2(E6) and 3 hit cell lines (Fig. 6.8.B), and also with the staining that was seen in all four of these undifferentiated cell lines in previous experiments (Fig. 6.7, Fig. 6.9A). This may indicate that these cell lines possess more osteogenic properties than other cell lines, but it should be noted that alkaline phosphatase is not a marker that is fully specific to osteogenesis.
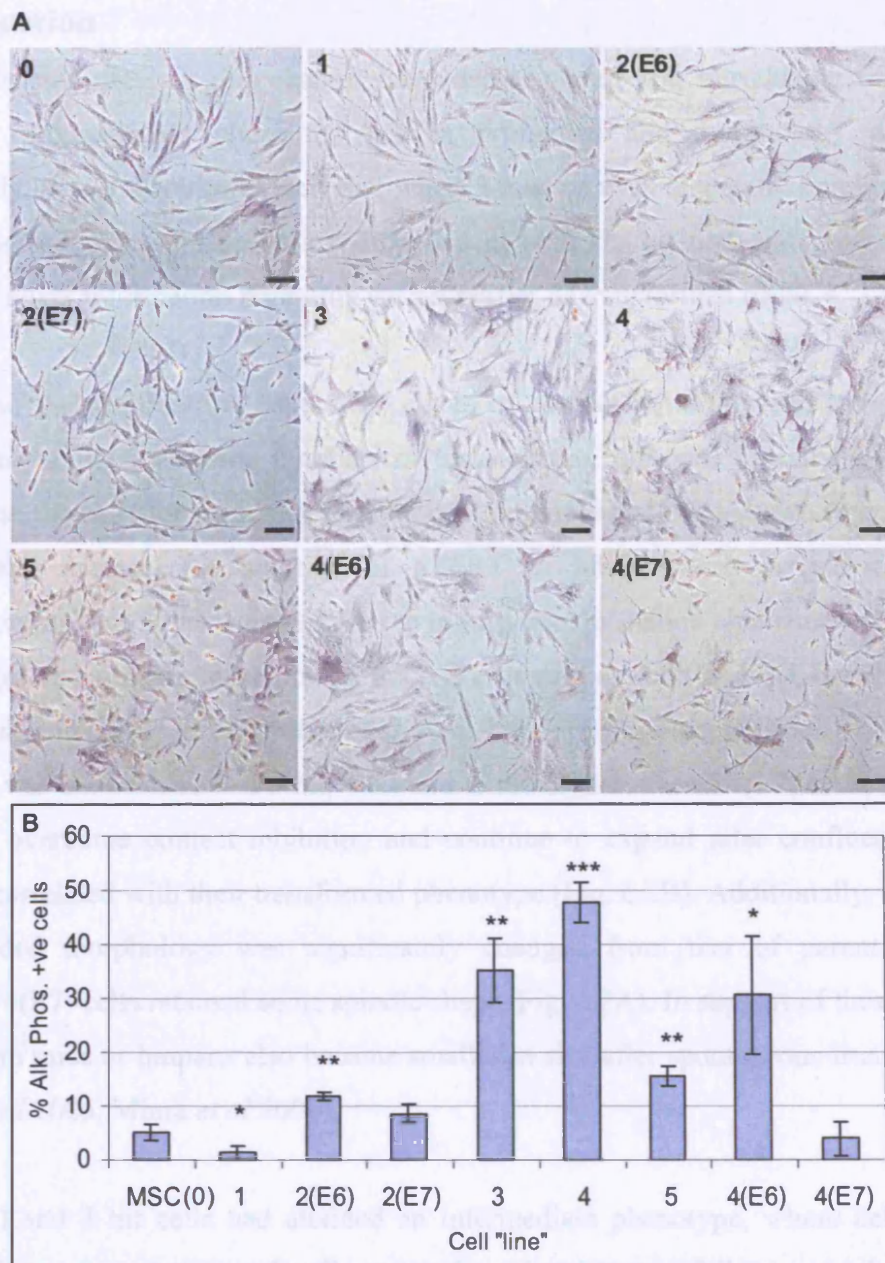
Figure 6.10. Background level of alkaline phosphatase activity in undifferentiated, subconfluent cells of each type. Cells undergoing normal passaging were kept at similar densities then stained for alkaline phosphatase activity (2.2.3.2). A, representative field of view for each cell line after staining. Scale bars represent 200μm; B, quantification of alkaline phosphatase activity in subconfluent cells of each cell line. Three fields of view were counted per cell line. Data represent % alkaline phosphatase positive cells per cell line and are mean ± standard deviation. *, p<0.05; **, p<0.01 and ***p<0.0001 using a two-sample Student's t-test (assuming equal variance), when values were compared to the parental hMSC cell line.

6.5 Discussion

The results described in this chapter have demonstrated that introducing transforming mutations progressively alters the growth properties and morphology of hMSCs. Importantly, it was also discovered that when 5 hits, or even a specific combination of 4 hits, are introduced, adipogenesis is impeded at both the cellular and molecular level, providing novel information regarding the properties of transformed hMSCs.

Analysis of the proliferation capability, extent of contact inhibition and morphology of hMSCs containing increasing numbers of transforming hits was initially performed to characterise the cells for further differentiation experiments. It was shown in related work that, in this system, the addition of hTERT to hMSCs was sufficient for their immortalisation, but 5 hits were required to induce transformation characterised by growth in soft agar and tumour formation in mice. Furthermore, 4 hit and 4(E7) cells attained some transformed properties. (Funes *et al* 2006, submitted and unpublished observations). Indeed, it was shown here that 5 hit, 4 hit and to the highest extent 4(E7) cells, showed an ability to overcome contact inhibition and continue to expand after confluence, which could be correlated with their transformed phenotype (Fig. 6.2B). Additionally, their small and rounded morphology was significantly changed from that of parental hMSCs (although 4(E7) cells retained some spindle shape, Fig. 6.2A). In support of these findings, MSCs from mice or humans also became smaller in size after spontaneous transformation (Rubio *et al* 2005, Miura *et al* 2006).

The 2(E7) and 3 hit cells had attained an intermediate phenotype, where cell shape is slightly changed from parental cell morphology but has not fully acquired the small, rounded shape of 4 or 5 hit cells. Furthermore, some proliferation occurred after confluence indicating that, to some extent, the cells were capable of overcoming contact inhibition. This would fit with the findings in this system that these cells are not transformed, as transformation entails loss of contact inhibition, and with previous studies that indicate that hMSCs containing up to 3 hits do not display transformed properties (Okamoto *et al* 2002, Hung *et al* 2004b). The fact that 2(E7) cells seem to be smaller, and/or can proliferate more after confluence than 2(E6) cells indicates that loss of Rb

plays a significant role in transformation; indeed, another study showed that hMSCs transduced with combinations of up to 3 hits including E7 had lower population doubling times than those lacking E7 (Mori *et al* 2005b).

Finally, it was found that cells containing 1 or 2(E6) hits, and surprisingly the 4(E6) cells, behave similarly to parental cells in that they are sensitive to contact inhibition (Fig 6.2). Further, 1 and 2(E6) hits cells retain morphologies similar to normal hMSCs. The observation that cells containing 4 hits but lacking E7 are not as aggressive as cells containing only hTERT and E7 again indicates that the loss of Rb may play a large contribution to the aggressive phenotype observed when E7 is present, and that loss of p53 is not sufficient to induce this phenotype. Similar to these findings, other studies of hMSCs immortalised with telomerase have not reported morphological changes or changes in proliferation rate after introduction of hTERT (Shi *et al* 2002), but the effects of the combinations of transforming hits exemplified by the 4(E6) cells used here, or even the 2(E6) cells, on growth and morphology of hMSCs or other cell types, have not been reported.

In summary, as well as providing an experimental basis and set-up for subsequent differentiation experiments, these observations of hMSC characteristics would allow an understanding of the relationship between differentiation capability and the progression towards, or attainment of, transformation in each cell line.

The effect of transformation on hMSC osteogenesis

Each of the cell lines containing 0-5 transforming hits was assessed for its' ability to undergo differentiation to both osteoblasts and adipocytes; a well-documented ability of untransformed hMSCs (Pittenger *et al* 1999). For osteogenesis, differentiation was initially assessed at 14d. Alkaline phosphatase activity and calcium deposition assays suggested that cells containing 0-3 hits, and also 4(E6) cells, were capable of some level of osteogenesis after induction, compared with their un-induced controls (Fig. 6.7), although calcium deposition was not extensive in parental and 1 hit differentiated cells. RT-PCR studies were not conclusive; a trend observed for ALP and PLZF was an increase in gene expression after osteogenic induction in some cell lines, which is in line with the

274

results of histochemical staining, but not all genes analysed showed this pattern so this conclusion is not fully supported (Fig. 6.8). However, previous work has reported similar findings; the introduction of hTERT maintained (Simonsen *et al* 2002) or increased (Shi *et al* 2002) hMSC osteogenic capacity, and Okamoto *et al* showed that hMSCs containing hTERT, E6 and E7 are still capable of undergoing osteogenesis, associated with increases in ALP and COL1A1 (Okamoto *et al* 2002). However, in that study osteogenesis was terminated at 21 days, so the endpoint of 14 days used here may have been too early to see some gene expression changes, or large increases in osteogenic staining intensity.

At day 14 it was not possible to comprehensively assess osteogenesis in 4, 4(E7) or 5 hit cells. Their aggressive proliferation led to the instability and subsequent detachment of the monolayer of all 4(E7) cells and some 4 and 5 hit cells between around days 7 and 14 of differentiation. In an attempt to combat this problem, osteogenesis was terminated after 6 days, at which time all cell lines were still attached as monolayers. However, when each differentiated cell line was stained and compared with it's undifferentiated counterpart, no significant differences in staining intensity, and therefore the extent of osteogenesis, were seen (Fig. 6.9A). Similar to this, RT-PCRs showed no particular pattern of expression of osteogenic marker genes between cell lines, and no visible differences in expression after differentiation (Fig. 6.9B). Thus, using this system it was not possible to conclusively assess osteogenesis in these cell lines. As previous studies have not assessed osteogenesis in 4 or 5 hit lines or after spontaneous differentiation, such problems have not previously been described. Due to the linked nature of transformation and differentiation pathways, it would be difficult to overcome this problem *in vitro*. Possible solutions include differentiation on a single-cell level in conditions that do not allow further cell growth, as in the system used by McBeath *et al* (2004), which assessed differentiation of hMSCs on fibronectin "islands" that allowed attachment of only a single cell, or after hMSC treatment with mitomycin-C, which arrests cell proliferation (McBeath *et al* 2004).

An observation common to 6 and 14 day experiments was a background level of alkaline phosphatase activity and calcium deposition in undifferentiated cells that was proportional to the number of hits. This may have been due either to increased osteogenic capacity of these cell lines, or a higher background level of staining due to the inevitable increase in

cell number. When this was assessed on the single cell level using the alkaline phosphatase assay, it was found that in the 2(E6), 3, 4, 4(E6) and 5 hit cell lines, a higher proportion of cells were positive for alkaline phosphatase activity than in the parental hMSCs (Fig. 6.10). This correlated with previous observations (Fig. 6.7, 6.8B, 6.9A), and may indicate that these cells have accumulated osteogenic properties as a result of the transforming hits. However, alkaline phosphatase activity is induced under conditions other than osteogenesis (the alkaline phosphatase isoform expressed in bone can also be expressed in other tissues such as liver and kidney), so in itself is not a definitive osteogenic marker. As a further indication of enhanced osteogenic properties in some undifferentiated cell lines, background expression level of CBFA1, osteocalcin and ALP was seen in some of these undifferentiated cell lines, specifically 2(E6) and 3 hit lines, which was higher than in parental hMSCs. It has been reported that the introduction of hTERT causes upregulation of CBFA1 and osteocalcin expression (Gronthos et al 2003b), so this may, at least in part, explain these observations, as gene expression is also increased in the 1 hit cell line (Fig. 6.8B). Surprisingly however, the 1 hit line showed significantly less alkaline phosphatase activity than parental hMSCs, at the single cell level. Heterozygous loss of Rb results in a 500-fold increase in the incidence of osteosarcoma (1.8.2, Thomas et al 2006). It could therefore be hypothesised that the introduction of E7, and therefore loss of Rb, could affect background levels of osteogenic markers in undifferentiated hMSCs; however, the 2(E7) and 4(E7) cells did not have significantly higher alkaline phosphatase activity, or background osteogenic gene expression levels, than parental hMSCs. It is therefore possible that this effect is context-dependent; i.e. it is affected by the other transforming mutations that are also present. Patients with Li-Fraumeni syndrome (characterised by inherited heterozygous loss of p53) are also more susceptible to osteosarcoma, as well as other tumour types, than the general population (Malkin et al 1990). This suggests that this gene is intrinsically linked to the osteogenic differentiation pathway, and hence the introduction of E6, and therefore the loss of p53 and Rb function, may alter the propensity for hMSCs to differentiate to osteoblasts. All of the cells lines in which E6 was present showed increased alkaline phosphatase activity in sub-confluent undifferentiated cells (Fig. 6.10), indicating that this transforming mutation may play a major role in altering alkaline phosphatase activity in hMSCs. However, it must again be stated that in itself, this observation is not sufficient to

276

draw conclusions regarding osteogenesis, and a more comprehensive analysis of differentiation, at the single-cell level if possible, would be needed. Of note, tumours formed by 5 hit cells in immunodeficient mice did not exhibit characteristics of osteosarcomas, but rather resembled spindle cell tumours (Funes *et al* 2006, submitted), which does not support the conclusion that 5 hit cells have a higher propensity to undergo osteogenesis.

In summary, although several attempts were made to assess osteogenesis after hMSC transformation and both staining and RT-PCR techniques were employed, it was not possible to conclusively assess osteogenesis in 4 (excluding 4(E6)) or 5 hit cell lines in this system. The growth properties of transformed hMSCs precluded the study of osteogenesis at it's classic 21 day endpoint or even a full analysis at 14 days, and insufficient differentiation at the early 6 day timepoint rendered the experiments inconclusive. However, these studies suggest that cells containing with 1-3 transforming hits, and also the 4(E6) cell lines, were capable or some extent of osteogenesis, similar to previous findings (Okamoto *et al* 2002, Simonsen *et al* 2002, Shi *et al* 2002). Furthermore, the introduction of some transforming genes may enhance spontaneous hMSC osteogenic capacity.

The effect of transformation on hMSC adipogenesis

The adipogenic differentiation system was found to have a major advantage in that adipogenic induction resulted in significant growth arrest, even in the more transformed cell lines (Fig. 6.3). This may have been due the adipogenic induction cocktail causing increased expression of PPARγ and C/EBPα, which are known to bring about growth arrest during adipogenesis through induction of several CDK inhibitors (1.6.2.1). Additionally, adipogenic differentiation needs a shorter time period than osteogenesis to be accomplished. Thus, adipogenesis could be assessed in all cell lines and a comprehensive analysis of the effect of transformation was possible.

It was found that adipogenesis was not significantly impaired in all cell lines with less than 3 hits, and also in the 4(E6) line, as lipid-laden cells were visible (Fig. 6.4), although not as numerous as with differentiated parental hMSCs. It is likely that this was not

caused, in the most part, by differences in passage number between cell lines, as low passage hMSCs expressing hTERT were as capable of adipogenesis as untransduced hMSCs, and there was no significant difference when compared with high passage hMSCs, although differentiation was more variable (Fig. 6.6). In line with these observations, hMSCs expressing hTERT were still capable of forming adipocytes after 40 or 60 population doublings (Shi *et al* 2002). Additionally, adipogenic gene expression was similar in these cell lines to parental cells (Fig. 6.5A). At the mRNA level, 2(E6) hit cells showed a slightly lower ability to differentiate, as expression of PPARγ and aP2 was reduced compared to other cell lines. There was a further decrease in lipid accumulation as the number of hits increased to 3, but no drop in the expression of adipogenic markers, which would suggest that mature adipocytes are still being formed. These results are similar to other observations where equivalent 3 hit cells, or cells containing just E6 and E7, were capable of forming mature adipocytes (Okamoto *et al* 2002, Hung *et al* 2004b). However, in the results reported here, adipogenesis appears less robust in 2 and 3 hit lines than previously reported. This could be due to differences in induction conditions, or due differences in the extent of transformation between the studies; the study by Okamoto *et al* required 3 hits (hTERT, E6 and E7) to immortalise hMSCs (Okamoto *et al* 2002), whereas here only hTERT was required, suggesting that the 3 hit cells used in this thesis may more closely resemble transformed cells than those used previously, possibly due to higher expression of the transforming genes in the cells described here, or the unknown source or passage number of the hMSCs used in the Okamoto study.

When 4 (including 4(E7)) or 5 transforming hits were introduced into hMSCs, a more dramatic effect was observed. Individual cells contained few, small lipid droplets in 5 hit and 4(E7) cells; slightly more lipid accumulation was observed in 4 hit cells (Fig. 6.4). It could be hypothesised that a decrease in lipid accumulation may be a result of a higher metabolic rate rather than decreased differentiation; however this decrease was associated by a drop in expression of adipogenic genes, most prominently in the 4(E7) cells, and to a lesser extent in 4 hit cells (Fig. 6.5B). It can therefore be concluded that although adipogenesis still occurs in these cell lines, it is impaired. This effect is most pronounced in the 4(E7) and 5 hit cell lines, thus hMSC transformation can be associated with impaired adipogenesis. This observation has not previously been reported. It remains

formally possible that the additional hits specific to 4(E7) or 5 hit cells (stabilisation of c-Myc and constitutive Ras activity) are the cause of this loss of differentiation, rather than it being a direct effect of transformation. To address this possibility, these hits would have to be introduced individually into hMSCs to investigate whether adipogenesis is similarly affected. However, introduction of c-Myc or oncogenic Ras individually into primary cells can result in apoptosis or senescence (Rangarajan *et al* 2004). It could therefore be concluded that, in this system, the additional introduction of hTERT and loss of p53 and Rb activity play an important role in the observations. Inducing hMSC transformation via different means (as discussed below) and determining the effect on differentiation would also be of use in resolving this issue.

As previously described (1.8.2), the molecular pathways involved in transformation and stem cell functions such as differentiation are highly inter-linked. This is indeed the case in adipogenesis, and could go some way to explaining the molecular basis behind the impairment of differentiation in cells that display transformed characteristics. The presence of E7, and thus the loss of Rb function, seems to be required for this effect, as 4(E6) cells behave in a similar manner to cells containing less transforming hits. Although Rb is known to inhibit late-stage adipogenesis involving PPARγ (Fajas *et al* 2002), it enhances early adipogenesis through the C/EBP proteins (Chen *et al* 1996, Cole *et al* 2004, 1.6.2.1); thus, loss of Rb function may block, or impede, the early stages of differentiation. Overexpression of Ras has been shown to increase adipogenesis in 3T3L1 cells, which could be expected as it is a downstream effector of insulin signalling (Porras *et al* 1994). This pro-adipogenic signal from Ras can be transmitted through Raf-1 kinase; however, Ras signalling can also inhibit adipogenesis if signals are transmitted through the MAPK pathway (Font de Mora *et al* 1997). This occurs, at least in part, through phosphorylation and inactivation of PPARγ mediated by ERK1 and ERK2 (Hu *et al* 1996, Font de Mora *et al* 1997). Furthermore, c-Myc is known to be an inhibitor of the terminal stages of 3T3-L1 differentiation (Heath *et al* 2000). Interestingly, a study in Rb-deficient MEFs demonstrated that Rb is capable of down-modulating ERK1/2 activity, possibly through suppression of Ras (Hansen *et al* 2002). It was hypothesised that the block in differentiation observed in these Rb-deficient cells could be caused by increased Ras activity leading to MAPK-mediated c-Myc stabilisation. It is notable that, in this system,

279

the loss of Rb on its own (plus hTERT expression, 2(E7) line) does not result in decreased adipogenesis. However, the results of the Hansen study could provide a theory regarding how the loss of Rb, activation of Ras, and stabilisation of c-Myc (as in the 5 hit line), or combinations thereof (4 or 4(E7) lines), could act in concert (or in synergy) to cause a defect to adipogenesis.

This work provides evidence that *in vitro* transformation of hMSCs can be associated with impaired adipogenic differentiation. This finding is in agreement with the fact that tumours, and more specifically sarcomas, are constituted mainly by cells that partially resemble mature cell types, and that cancer stem cells are capable of forming all cell types in their tumour of origin, so it would not be expected that differentiation would be entirely abolished. Human myxoid liposarcomas, a malignant sarcoma of adipose tissue containing cells partially resembling adipocytes, are defined by a chromosomal translocation that fuses the basic helix-loop-helix domain of the CHOP10 gene (an inhibitor of adipogenesis under normal circumstances, 1.6.2.3 and Fig. 4.6) with the RNA binding domain of the FUS gene, or rarely the EWS gene (Pérez-Mancera and Sánchez-García 2005). This results in deregulated protein-binding activities of CHOP10, which may block adipogenesis by interfering with C/EBPβ function (Adelmant *et al* 1998). The mechanisms underlying the transforming ability of this fusion protein are not clear, but require the FUS/EWS RNA binding domain (Zinszner *et al* 1994). In relation to the theory that cancer stem cells may arise via mutations in a normal stem cell population, it has been suggested that the CHOP10-FUS translocation may occur in hMSCs, causing a lineage-specific block in differentiation at a stage along the adipogenic pathway (Pérez-Mancera and Sánchez-García 2005). However, this remains a theory, and it is also possible that the translocation may occur in adipogenic precursors. It would therefore be of interest to investigate the effect on hMSC transformation and differentiation of other transforming hits; specifically the introduction of the FUS-CHOP10 fusion gene into hMSCs *in vitro*. Such work may lead to further understanding of the mechanisms by which transformation affects adipogenesis.

CHAPTER 7. GENERAL DISCUSSION.

7.1 Conclusions and discussion

The research described in this thesis has resulted in the following main novel observations:

- hMSC adipogenesis can be stimulated via the overexpression of PPARγ alone, without the need for addition of differentiation inducers.

- Distinct temporal patterns of transcriptional regulation define the earliest stages of hMSC adipogenesis. This involves over 4200 genes, many of which were described as being regulated during hMSC adipogenesis for the first time in this thesis.

- Specific biological processes such as the up- and downregulation of subsets of developmental genes and temporal regulation of Wnt, TGFβ and BMP pathway components are involved in early hMSC adipogenesis

- A novel functional genomics strategy combining gene expression profiling, comprehensive global annotation and functional screens was developed and applied to investigate early hMSC adipogenesis

- Finally, transformation of hMSCs results in impaired adipogenesis.

Elucidation of the molecular mechanisms underlying stem cell differentiation would extend our knowledge of the processes of early development, and may also provide a means to exploit this unique attribute of stem cells in regenerative medicine. The properties of MSCs render this population and ideal system with which to achieve such aims. Their potential to regenerate damaged tissues has already been highlighted by animal studies and even some clinical trials (Toma *et al* 2002, Horwitz *et al* 1999, 2002, 1.4.6). Furthermore, they are a well-defined cell population, and their isolation from adult human bone marrow, or other tissue sources, means that a large numbers of hMSCs can be obtained easily and without confounding ethical issues. Crucially, hMSCs are easily manipulated in culture and can robustly differentiate into several cell lineages *in vitro*, including adipocytes. Despite this, little is known regarding the earliest molecular events of hMSC differentiation to adipocytes, as much research into adipogenesis has used committed precursor cell lines. The earliest stages of stem cell differentiation involve processes such as lineage commitment, and it is possible that the molecular events that regulate these processes may be similar between different stem cell populations, or within the various lineages of one stem cell type. Thus, whereas molecular regulation at the later

stages of adipogenesis is often specific to the individual lineage (PPARγ, for example), an understanding of the early molecular events of adipogenesis may have broader applications in other stem cell types. It was therefore aim of this thesis to investigate the hypothesis that hMSC adipogenesis involves important molecular events that occur at an early stage of differentiation. The major strategy to achieve this aim was the use of a combination of functional genomics techniques, the results of which revealed much detail regarding the molecular processes at the early stages of hMSC adipogenesis.

In Chapter 3 optimal hMSC isolation, growth and differentiation to adipocytes was established as the first step towards this goal. The properties of the hMSCs used, in terms of their surface marker profile, growth kinetics and ability to undergo differentiation, were found to be similar to previous reports (Pittenger *et al* 1999, Tsutsumi *et al* 2001), and hence it was concluded that accurate comparisons could be made between this work and the work of other researchers. Subsequently, it was shown that gene expression changes could be detected in this system within the first 9h of hMSC adipogenesis that were likely to represent crucial early molecular events. Furthermore, it was concluded that this stage of differentiation had not previously been characterised by other work. This research thus provided a basis for later studies that were aimed at detecting and defining these events. Although it was found in Chapter 4 that several known regulators of preadipocyte differentiation, such as C/EBPβ and KLF2, were also transcriptionally regulated during early hMSC differentiation, this was concluded to be possibly due to the similar methods used to induce differentiation of preadipocyte cell lines and hMSCs. Additional genes that showed similar expression patterns to these known regulators had not previously been shown to be regulated during adipogenesis, and therefore presented a unique opportunity to identify novel regulators of early hMSC adipogenesis.

The microarray expression profiles obtained during differentiation (Chapter 4) revealed that the first 9h of hMSC adipogenesis involved expression changes in 6319 probesets, equivalent to over 4200 known genes (given the data processing methods and significance cut-off employed). Although other microarray time-course studies have investigated this process, arrays covering only a subset of human genes were used (Nakamura *et al* 2003, Sekiya *et al* 2004). This thesis was the first study to use genome-wide microarrays to

investigate hMSC differentiation, and as a result demonstrated that transcriptional regulation at the early stages of hMSC adipogenesis is more complex than previously known. In a similar fashion to the regulation of preadipocyte differentiation, some genes described in this thesis had been identified in other studies of hMSC adipogenesis, but many more genes were not known to be regulated during this process.

This thesis employed hierarchical clustering of genes showing significant expression changes, combined with a number of global annotation methods, with the aim of obtaining an overview of the pathways and biological processes that function at specific stages of early hMSC adipogenesis. The work described here represents a step beyond the analytical methods used in other microarray studies of stem cell function. Frequently, gene expression changes during stem cell differentiation have been represented in simple lists indicating up-regulated or down-regulated genes (for example, Hung *et al* 2004a, Sekiya *et al* 2002, 2004), but this does not give an indication as to the interacting gene networks that may play vital roles. It is becoming more common as the technology advances to employ global annotation methodologies to perform microarray data analyses, which can be used to interpret the function of the regulated genes and thus identify biological processes, signalling pathways or gene families (as a non-exhaustive list) that may play a role in the process under investigation. This has been the case in 3T3-L1 adipogenesis (Hackl *et al* 2005), a recent study of hMSC differentiation to adipocytes and chondrocytes (Yloslato *et al* 2006), and also, for example, in work concerning mESC pluripotency (Palmqvist *et al* 2005). However, if small microarrays are used (as in all these studies), an incomplete view of regulatory pathways and processes with be obtained as all components will not have been analysed. Additionally, some annotation methods can neglect important information; GO is not yet fully comprehensive and hence some known human genes do not possess GO annotations, and pathway maps for visualising microarray data may not contain all members of a certain signalling pathway. This problem can be addressed by employing several annotation techniques, as demonstrated by the studies by Hackl and Palmqvist, but in the field of hMSC adipogenesis this has yet to be achieved. This thesis made use of several different annotation methods, as well as genome-wide arrays, and thus presents the most comprehensive view of the pathways and processes involved in hMSC adipogenesis. By way of example, although previous work had indicated that

developmental genes become downregulated during hMSC adipogenesis (Hung *et al* 2004a), the use of Gene Ontology as well as manual annotation of specific expression clusters demonstrated further detail regarding the full range of developmental genes involved. Furthermore, the involvement of the Wnt, TGFβ and BMP signalling pathways during adipogenesis in hMSC or C3H10T1/2 cells (a multipotent murine cell line with differentiation capabilities similar to MSCs, 1.6.1) has previously been described (Locklin *et al* 1999, De Boer *et al* 2004, Etheridge *et al* 2004, Tang *et al* 2004), but through the use of global pathway analysis techniques, the exact pathway components involved as well as their specific timing of expression was revealed by work in this thesis.

The automated and manual annotation techniques identified a substantial number of candidate regulators of the early stages of hMSC adipogenesis. A shRNA library screening strategy was chosen to investigate the roles of these candidates (Chapter 5), the reasons for this decision being twofold.

Firstly, it was shown in Chapter 3 that retroviral transduction of hMSCs with a vector expressing PPARγ isoforms permitted evaluation of the effect of exogenous genes on differentiation. Interestingly, this is the first report of the ability of PPARγ to induce spontaneous adipogenesis in hMSC without the need for induction reagents. Taken together with other studies reporting the retroviral over-expression of genes in MSCs of human or murine origin (Shi *et al* 2002, McBeath *et al* 2004), or RNAi-mediated gene silencing in these cells (Hong *et al* 2005, Xu *et al* 2006), this work confirmed that the use of such tools would allow thorough investigation of candidate gene function during hMSC differentiation.

Secondly, shRNA library screens are documented as an effective high-throughput method of identifying factors from a large initial gene set (i.e. all genes covered by the library, or a chosen subset of genes) that are involved in a specific process (Berns *et al* 2004, Downward 2004). It appears from the literature that high-throughput methods are rarely used to follow on from microarray experiments, and instead only one or few genes are the focus of further research (Ross *et al* 2002, Ikeda *et al* 2005). This is perhaps because other microarray studies, on the whole, yield smaller lists of significantly changed genes than

described here, rendering high-throughput techniques less of a requirement. However, in the context of this thesis, and perhaps as an example for future functional genomics research, shRNA screening was considered a particularly relevant technique for the downstream functional analysis of data produced by microarray expression profiling. The combination of these two approaches would maximise the output of results; performing a shRNA screen alone has the potential to identify novel factors involved in early hMSC adipogenesis, but no information regarding their mechanism of action would be obtained. In combination with microarray gene expression analysis, information regarding co-regulated genes, common pathways members or genes with a similar function would be provided, allowing for more efficient characterisation of the role of specific genes during adipogenesis. Thus, this thesis has described the use of a novel strategy designed to maximise the output of results from microarray studies that has not been employed in previous work. Future work (discussed in 7.2) would aim to use modified techniques of shRNA screening to realise the full potential of this strategy.

A further undertaking of this thesis was to investigate the effect of the stepwise introduction of genes involved in neoplastic transformation on hMSC differentiation. It is a strong possibility that cancer-initiating stem cells may arise from transforming mutations in normal stem cell populations. In this thesis it was therefore hypothesised that modelling this situation *in vitro*, for which hMSCs were again an ideal system, and investigating the effect of transformation on a fundamental property of stem cells, i.e. differentiation, may increase our knowledge of the processes involved in cancer progression. In the results described in Chapter 6, it was found that when hMSCs were transformed adipogenesis could still occur, but in an impaired fashion. When fewer transforming genes were introduced into hMSCs that were not sufficient in number or type to induce transformation, such a pronounced effect on differentiation was not observed. It could be hypothesised that this result would be expected, as the process of transformation often involves perturbation of pathways linked to stem cell self-renewal or differentiation. Several studies have assessed the ability of hMSCs to become spontaneously transformed, but did not address their differentiation ability after transformation (Rubio *et al* 2005, Serakinci *et al* 2004). In addition, although the step-wise model of introduction of transforming hits has been applied to hMSCs, no study has achieved full transformation of

286

these cells (Okamoto *et al* 2002, Hung *et al* 2004b). Thus, this is the first documentation of the effect of the *in vitro* differentiation capability of transformed hMSCs. Step-wise *in vitro* transformation has not been described in other adult stem cell populations, so the work described here may be a useful basis for further studies.

7.2 Future directions

The work described in this thesis presents a number of interesting possibilities for future research which, given more time and enhanced technologies, could elucidate further the molecular regulation of early hMSC adipogenesis.

The largest potential for adding to the results described in this thesis would come from developing the shRNA screening strategy to allow a more thorough analysis of the wealth of information provided by the microarray studies. A major technical difficulty encountered during the shRNA screens described in this thesis was recombination of the retroviral vector leading to the loss of shRNA inserts. Therefore, future work could involve using advanced shRNA libraries that utilise vectors with selectable markers to exclude recombined plasmids, or a shRNA library that was amplified in a recombination-deficient bacterial strain (Echeverri and Perrimon 2006). This would lead to a great increase in the efficiency of the strategy, as experimental procedures would be more straightforward and a higher number of genes could be analysed simultaneously. Indeed, the first shRNA isolation method described in Chapter 5 (5.2.1) could be employed, and its relative simplicity highlights the potential rapidity of shRNA isolation, if recombination does not introduce a complicating factor. The availability of high-throughput laboratory equipment could allow the screen to be performed in 96-well plate format, as the loss-of-function readout in hMSC differentiation is clearly visible, and enough cells would likely be present in each well for sufficient quantification of a knock-down effect. Finally, to complement such work (but on a smaller scale) the availability of inducible shRNA libraries (Ngo *et al* 2006) could allow gene knock-downs to be induced at specific stages of differentiation, thus allowing a more in-depth analysis of whether a gene is required only at the early stages of hMSC adipogenesis.

The results described in Chapter 5 revealed that two custom-designed shRNAs produced small reductions in the expression of TCF8 and SIX2, but no corresponding reduction in adipogenesis was observed. Given more time, it would be interesting to investigate further these observations, such as whether these shRNAs lead to a noticeable reduction in protein levels, or whether modified shRNAs (i.e. using the same 19nt shRNA sequence but different hairpin loop sequences or the U6 promoter) could achieve a more robust knock-down of expression of these genes. It was also found in Chapter 5 that lentiviral vectors may be more efficient than retroviral vectors at achieving gene-knockdowns in hMSCs, as the shRNAs that produced these observed gene expression knock-downs were delivered into hMSCs using a lentiviral vector. Lentivirus-based shRNA libraries do not yet exist, so it would be difficult to perform large-scale screens using these vectors, but to study a small number of genes using lentivirally-mediated shRNA would be possible. Further to this, a recurring observation was that vectors, or hairpin-containing vectors, induced what was possibly an interferon response, leading to a reduction in differentiation or gene expression. However, the occurrence of an interferon response is currently speculation and it would therefore be useful, in terms of optimising the use of vector-based shRNA in hMSCs, to investigate whether the expression of interferon-related genes is induced after hMSC transduction.

An exciting possibility is now arising from advances in global gene expression analysis technology as well as the knowledge of the entire human genome sequence (Lander et al 2001, Venter et al 2001). This is the identification of transcription factor binding sites in the upstream promoter elements of large gene groups selected from microarray studies, with the aim of elucidating transcriptional networks controlling the process under investigation (Karanam et al 2004). Recently, this technique was used to follow up a microarray study of hMSC adipogenesis (Yloslato et al 2006). Several novel downstream targets of PPARγ and C/EBPα were described via analysis of the upstream promoter elements of genes that were co-regulated with known PPARγ or C/EBPα targets, and subsequent identification of PPARγ and/or C/EBPα binding sites in these promoter regions. This strategy could be applied directly to the microarray study described in this thesis. Transcription factors with known binding sites that were up-regulated at an early

time during differentiation could be used as a starting point, and then genes in the same expression cluster as this factor, or showing expression changes (whether it be up- or downregulation) at the same timepoint, or with known targets of the transcription factor, could be analysed for the presence of the specific binding site. Thus, putative transcriptional networks could be compiled that define the earliest transcriptional events of hMSC adipogenesis, which could subsequently be functionally tested using the strategy described above. Several limitations to this method exist, such as the incomplete knowledge of the DNA binding sequences for all transcription factors, as well as the possibility that a binding site may lie outside of the region being analysed. However, this form of analysis would provide a further opportunity to extract biologically relevant information for the microarray data.

Some interesting avenues for further investigation arose from the studies of hMSC transformation and the effect on differentiation. As discussed in Chapter 6, it was not possible with the system used to accurately assess the ability of hMSCs to undergo osteogenesis. Methods that allow osteogenesis to be assessed on a single-cell level such as seeding cells on fibronectin "islands", and inhibiting their ability to proliferate (which would not be likely to affect differentiation as growth arrest rather than proliferation is a required step for differentiation) could be developed. In combination with studies of adipogenesis, this would allow investigation of whether hMSC transformation causes a bias to differentiation into a particular lineage. Finally, it could also be of clinical relevance to investigate the effect on hMSC differentiation of other methods of transformation. An intriguing candidate for this would be the FUS-CHOP10 fusion protein that is commonly observed in liposarcomas, tumours of mesenchymal origin. This has not yet been expressed in hMSCs, but such research has the potential to discover the mechanisms by which some mesenchymal tumours originate. It is thought that the FUS-CHOP10 translocation results in a block in differentiation at a point prior to PPARγ expression (Adelmant et al 1998). If novel regulators of early adipogenesis were to be identified as a result of the work described in this thesis, their expression could be studied during differentiation of hMSCs expressing the FUS-CHOP10 translocation in order to understand the mechanisms or the precise stage at which this translocation may prevent normal hMSC adipogenesis.

REFERENCES

Abbas-Terki,T., Blanco-Bose,W., Deglon,N., Pralong,W., and Aebischer,P. (2002). Lentiviral-mediated RNA interference. Hum. Gene Ther. *13*, 2197-2201.

Adelmant,G., Gilbert,J.D., and Freytag,S.O. (1998). Human translocation liposarcoma-CCAAT/enhancer binding protein (C/EBP) homologous protein (TLS-CHOP) oncoprotein prevents adipocyte differentiation by directly interfering with C/EBPbeta function. J. Biol. Chem. *273*, 15574-15581.

Affymetrix (2001) Microarray Suite User Guide, Version 5. Available at: http://www.affymetrix.com/support/technical/manuals.affx.

Affymetrix (2003) Datasheet: GeneChip® Human Genome Arrays. Available at: http://www.affymetrix.com/support/technical/datasheets/human_datasheet.pdf

Affymetrix (2004) GeneChip Expression Analysis Technical Manual. Available at: http://www.affymetrix.com/support/technical/manual/expression_manual.affx

Ahrens,M., Ankenbauer,T., Schroder,D., Hollnagel,A., Mayer,H., and Gross,G. (1993). Expression of human bone morphogenetic proteins-2 or -4 in murine mesenchymal progenitor C3H10T1/2 cells induces differentiation into distinct mesenchymal cell lineages. DNA Cell Biol. *12*, 871-880.

Ailhaud,G., Grimaldi,P., and Negrel,R. (1992). Cellular and molecular aspects of adipose tissue development. Annu. Rev. Nutr. *12*, 207-233.

Ailhaud G. (2001) Development of white adipose tissue and adipocyte differentiation. In Adipose Tissue, S. Klaus, ed. (Georgetown, Texas: Landes Bioscience), pp 27-55

Akashi,K., Traver,D., Miyamoto,T., and Weissman,I.L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature *404*, 193-197.

Akerblad,P., Lind,U., Liberg,D., Bamberg,K., and Sigvardsson,M. (2002). Early B-cell factor (O/E-1) is a promoter of adipogenesis and involved in control of genes important for terminal adipocyte differentiation. Mol. Cell Biol. *22*, 8015-8025.

Akusjarvi,G., Svensson,C., and Nygard,O. (1987). A mechanism by which adenovirus virus-associated RNAI controls translation in a transient expression assay. Mol. Cell Biol. *7*, 549-551.

Al-Hajj,M., Wicha,M.S., ito-Hernandez,A., Morrison,S.J., and Clarke,M.F. (2003). Prospective identification of tumorigenic breast cancer cells. Proc. Natl. Acad. Sci. U.S.A *100*, 3983-3988.

Al-Hajj,M., and Clarke,M.F. (2004). Self-renewal and solid tumor stem cells. (2004). Oncogene *23*, 7274-7282.

Albright,A.L. and Stern,J.S. (1998). Adipose Tissue. In Encyclopedia of Sports Medicine and Science, T.D.Fahey, ed. Internet Society for Sport Science).

Allison,D.B., Cui,X., Page,G.P., and Sabripour,M. (2006). Microarray data analysis: from disarray to consolidation and consensus. Nat. Rev. Genet. 7, 55-65.

Altiok,S., Xu,M., and Spiegelman,B.M. (1997). PPARgamma induces cell cycle withdrawal: inhibition of E2F/DP DNA-binding activity via down-regulation of PP2A. Genes Dev. 11, 1987-1998.

Ambion (2005) pSilencer™ 5.1 Retro Vectors Technical Manual (version 0501). Available at: http://www.ambion.com/techlib/prot/fm_5782.pdf

Anjos-Afonso,F., Siapati,E.K., and Bonnet,D. (2004). In vivo contribution of murine mesenchymal stem cells into multiple cell-types under minimal damage conditions. J. Cell Sci. 117, 5655-5664.

Arango,N.A., Szotek,P.P., Manganaro,T.F., Oliva,E., Donahoe,P.K., and Teixeira,J. (2005). Conditional deletion of beta-catenin in the mesenchyme of the developing mouse uterus results in a switch to adipogenesis in the myometrium. Dev. Biol. 288, 276-283.

Asahina,I., Sampath,T.K., and Hauschka,P.V. (1996). Human osteogenic protein-1 induces chondroblastic, osteoblastic, and/or adipocytic differentiation of clonal murine target cells. Exp. Cell Res. 222, 38-47.

Ashburner,M., Ball,C.A., Blake,J.A., Botstein,D., Butler,H., Cherry,J.M., Davis,A.P., Dolinski,K., Dwight,S.S., Eppig,J.T., Harris,M.A., Hill,D.P., Issel-Tarver,L., Kasarskis,A., Lewis,S., Matese,J.C., Richardson,J.E., Ringwald,M., Rubin,G.M., and Sherlock,G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 25 , 25-29.

Ashrafi,K., Chang,F.Y., Watts,J.L., Fraser,A.G., Kamath,R.S., Ahringer,J., and Ruvkun,G. (2003). Genome-wide RNAi analysis of Caenorhabditis elegans fat regulatory genes. Nature 421, 268-272.

Ashton,B.A., Allen,T.D., Howlett,C.R., Eaglesom,C.C., Hattori,A., and Owen,M. (1980). Formation of bone and cartilage by marrow stromal cells in diffusion chambers in vivo. Clin. Orthop. 294-307.

Aubert,J., Dessolin,S., Belmonte,N., Li,M., McKenzie,F.R., Staccini,L., Villageois,P., Barhanin,B., Vernallis,A., Smith,A.G., Ailhaud,G., and Dani,C. (1999). Leukemia inhibitory factor and its receptor promote adipocyte differentiation via the mitogen-activated protein kinase cascade. J. Biol. Chem. 274, 24965-24972.

Aubert,J., Saint-Marc,P., Belmonte,N., Dani,C., Negrel,R., and Ailhaud,G. (2000). Prostacyclin IP receptor up-regulates the early expression of C/EBPbeta and C/EBPdelta in preadipose cells. Mol. Cell Endocrinol. 160, 149-156.

Austin,S., Medvedev,A., Yan,Z.H., Adachi,H., Hirose,T., and Jetten,A.M. (1998). Induction of the nuclear orphan receptor RORgamma during adipocyte differentiation of D1 and 3T3-L1 cells. Cell Growth Differ. *9*, 267-276.

Aza-Blanc,P., Cooper,C.L., Wagner,K., Batalov,S., Deveraux,Q.L., and Cooke,M.P. (2003). Identification of modulators of TRAIL-induced apoptosis via RNAi-based phenotypic screening. Mol. Cell *12*, 627-637.

Baksh,D., Song,L., and Tuan,R.S. (2004). Adult mesenchymal stem cells: characterization, differentiation, and application in cell and gene therapy. J. Cell Mol. Med. *8*, 301-316.

Banerjee,S.S., Feinberg,M.W., Watanabe,M., Gray,S., Haspel,R.L., Denkinger,D.J., Kawahara,R., Hauner,H., and Jain,M.K. (2003). The Kruppel-like factor KLF2 inhibits peroxisome proliferator-activated receptor-gamma expression and adipogenesis. J. Biol. Chem. *278*, 2581-2584.

Bantounas,I., Phylactou,L.A., and Uney,J.B. (2004). RNA interference and the use of small interfering RNA to study gene function in mammalian systems. J. Mol. Endocrinol. *33*, 545-557.

Barak,Y., Nelson,M.C., Ong,E.S., Jones,Y.Z., Ruiz-Lozano,P., Chien,K.R., Koder,A., and Evans,R.M. (1999). PPAR gamma is required for placental, cardiac, and adipose tissue development. Mol. Cell *4*, 585-595.

Baum,C.M., Weissman,I.L., Tsukamoto,A.S., Buckle,A.M., and Peault,B. (1992). Isolation of a candidate human hematopoietic stem-cell population. Proc. Natl. Acad. Sci. U.S.A *89*, 2804-2808.

Bays,H., Mandarino,L., and DeFronzo,R.A. (2004). Role of the adipocyte, free fatty acids, and ectopic fat in pathogenesis of type 2 diabetes mellitus: peroxisomal proliferator-activated receptor agonists provide a rational therapeutic approach. J. Clin. Endocrinol. Metab *89*, 463-478.

Belmonte,N., Phillips,B.W., Massiera,F., Villageois,P., Wdziekonski,B., Saint-Marc,P., Nichols,J., Aubert,J., Saeki,K., Yuo,A., Narumiya,S., Ailhaud,G., and Dani,C. (2001). Activation of extracellular signal-regulated kinases and CREB/ATF-1 mediate the expression of CCAAT/enhancer binding proteins beta and -delta in preadipocytes. Mol. Endocrinol. *15*, 2037-2049.

Beltrami,A.P., Barlucchi,L., Torella,D., Baker,M., Limana,F., Chimenti,S., Kasahara,H., Rota,M., Musso,E., Urbanek,K., Leri,A., Kajstura,J., Nadal-Ginard,B., and Anversa,P. (2003). Adult cardiac stem cells are multipotent and support myocardial regeneration. Cell *114*, 763-776.

Benjamini Y. and Hochberg Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Statist. Soc. B, 57(1) 289-300

Bennett,C.N., Ross,S.E., Longo,K.A., Bajnok,L., Hemati,N., Johnson,K.W., Harrison,S.D., and MacDougald,O.A. (2002). Regulation of Wnt signaling during adipogenesis. J. Biol. Chem. *277*, 30998-31004.

Berns,K., Hijmans,E.M., Mullenders,J., Brummelkamp,T.R., Velds,A., Heimerikx,M., Kerkhoven,R.M., Madiredjo,M., Nijkamp,W., Weigelt,B., Agami,R., Ge,W., Cavet,G., Linsley,P.S., Beijersbergen,R.L., and Bernards,R. (2004). A large-scale RNAi screen in human cells identifies new components of the p53 pathway. Nature *428*, 431-437.

Bhatia,M., Wang,J.C., Kapp,U., Bonnet,D., and Dick,J.E. (1997). Purification of primitive human hematopoietic cells capable of repopulating immune-deficient mice. Proc. Natl. Acad. Sci. U.S.A *94*, 5320-5325.

Bhatia,M., Bonnet,D., Murdoch,B., Gan,O.I., and Dick,J.E. (1998). A newly discovered class of human hematopoietic cells with SCID-repopulating activity. Nat. Med. *4*, 1038-1045.

Bjerkvig,R., Tysnes,B.B., Aboody,K.S., Najbauer,J., and Terzis,A.J. (2005). Opinion: the origin of the cancer stem cell: current controversies and new insights. Nat. Rev. Cancer *5*, 899-904.

Blanpain,C., Lowry,W.E., Geoghegan,A., Polak,L., and Fuchs,E. (2004). Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. Cell *118*, 635-648.

Boiret,N., Rapatel,C., Veyrat-Masson,R., Guillouard,L., Guerin,J.J., Pigeon,P., Descamps,S., Boisgard,S., and Berger,M.G. (2005). Characterization of nonexpanded mesenchymal progenitor cells from normal adult human bone marrow. Exp. Hematol. *33*, 219-225.

Bolstad,B.M., Irizarry,R.A., Astrand,M., and Speed,T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics. *19*, 185-193.

Bonnet,D. and Dick,J.E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat. Med. *3*, 730-737.

Bonnet,D. (2003). Biology of human bone marrow stem cells. Clin. Exp. Med. *3*, 140-149.

Boschmann M. (2001) Heterogeneity of adipose tissue metabolism. In Adipose Tissue, S. Klaus, ed. (Georgetown, Texas: Landes Bioscience), pp 131-157

Bradley,A., Evans,M., Kaufman,M.H., and Robertson,E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature *309*, 255-256.

Brazelton,T.R., Rossi,F.M., Keshet,G.I., and Blau,H.M. (2000). From marrow to brain: expression of neuronal phenotypes in adult mice. Science *290*, 1775-1779.

Bridge,A.J., Pebernard,S., Ducraux,A., Nicoulaz,A.L., and Iggo,R. (2003). Induction of an interferon response by RNAi vectors in mammalian cells. Nat. Genet. *34*, 263-264.

Bruder,S.P., Jaiswal,N., and Haynesworth,S.E. (1997). Growth kinetics, self-renewal, and the osteogenic potential of purified human mesenchymal stem cells during extensive subcultivation and following cryopreservation. J. Cell Biochem. *64*, 278-294.

Brummelkamp,T.R., Bernards,R., and Agami,R. (2002a). A system for stable expression of short interfering RNAs in mammalian cells. Science *296*, 550-553.

Brummelkamp,T.R., Bernards,R., and Agami,R. (2002b). Stable suppression of tumorigenicity by virus-mediated RNA interference. Cancer Cell *2*, 243-247.

Brummelkamp,T.R., Nijman,S.M., Dirac,A.M., and Bernards,R. (2003). Loss of the cylindromatosis tumour suppressor inhibits apoptosis by activating NF-kappaB. Nature *424*, 797-801.

Burns,J.S., Abdallah,B.M., Guldberg,P., Rygaard,J., Schroder,H.D., and Kassem,M. (2005). Tumorigenic heterogeneity in cancer stem cells evolved from long-term cultures of telomerase-immortalized human mesenchymal stem cells. Cancer Res. *65*, 3126-3135.

Burton,G.R., Guan,Y., Nagarajan,R., and McGehee,R.E., Jr. (2002). Microarray analysis of gene expression during early adipocyte differentiation. Gene *293*, 21-31.

Burton,G.R. and McGehee,R.E., Jr. (2004). Identification of candidate genes involved in the regulation of adipocyte differentiation using microarray-based gene expression profiling. Nutrition *20*, 109-114.

Burton,G.R., Nagarajan,R., Peterson,C.A., and McGehee,R.E., Jr. (2004). Microarray analysis of differentiation-specific gene expression during 3T3-L1 adipogenesis. Gene *329*, 167-185.

Butte,A. (2002). The use and analysis of microarray data. Nat. Rev. Drug Discov. *1*, 951-960.

Cao,Z., Umek,R.M., and McKnight,S.L. (1991). Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. Genes Dev. *5*, 1538-1552.

Caplan,A.I. (1991). Mesenchymal stem cells. J. Orthop. Res. *9*, 641-650.

Castro-Malaspina,H., Gay,R.E., Resnick,G., Kapoor,N., Meyers,P., Chiarieri,D., McKenzie,S., Broxmeyer,H.E., and Moore,M.A. (1980). Characterization of human bone marrow fibroblast colony-forming cells (CFU-F) and their progeny. Blood *56*, 289-301.

Champigny,O. and Ricquier,D. (1996). Evidence from in vitro differentiating cells that adrenoceptor agonists can increase uncoupling protein mRNA level in adipocytes of adult humans: an RT-PCR study. J. Lipid Res. *37*, 1907-1914.

Chen,D., Ji,X., Harris,M.A., Feng,J.Q., Karsenty,G., Celeste,A.J., Rosen,V., Mundy,G.R., and Harris,S.E. (1998). Differential roles for bone morphogenetic protein (BMP) receptor

type IB and IA in differentiation and specification of mesenchymal precursor cells to osteoblast and adipocyte lineages. J. Cell Biol. *142*, 295-305.

Chen,F., Ogawa,K., Nagarajan,R.P., Zhang,M., Kuang,C., and Chen,Y. (2003). Regulation of TG-interacting factor by transforming growth factor-beta. Biochem. J. *371*, 257-263.

Chen,H., Charlat,O., Tartaglia,L.A., Woolf,E.A., Weng,X., Ellis,S.J., Lakey,N.D., Culpepper,J., Moore,K.J., Breitbart,R.E., Duyk,G.M., Tepper,R.I., and Morgenstern,J.P. (1996a). Evidence that the diabetes gene encodes the leptin receptor: identification of a mutation in the leptin receptor gene in db/db mice. Cell *84*, 491-495.

Chen,P.L., Riley,D.J., Chen,Y., and Lee,W.H. (1996b). Retinoblastoma protein positively regulates terminal adipocyte differentiation through direct interaction with C/EBPs. Genes Dev. *10*, 2794-2804.

Chen,Z., Torrens,J.I., Anand,A., Spiegelman,B.M., and Friedman,J.M. (2005). Krox20 stimulates adipogenesis via C/EBPbeta-dependent and -independent mechanisms. Cell Metab *1*, 93-106.

Cheyette,B.N., Waxman,J.S., Miller,J.R., Takemaru,K., Sheldahl,L.C., Khlebtsova,N., Fox,E.P., Earnest,T., and Moon,R.T. (2002). Dapper, a Dishevelled-associated antagonist of beta-catenin and JNK signaling, is required for notochord formation. Dev. Cell *2*, 449-461.

Cho,H.H., Kim,Y.J., Kim,S.J., Kim,J.H., Bae,Y.C., Ba,B., and Jung,J.S. (2006). Endogenous Wnt signaling promotes proliferation and suppresses osteogenic differentiation in human adipose derived stromal cells. Tissue Eng *12*, 111-121.

Choudhuri,S. (2004). Microarrays in biology and medicine. J. Biochem. Mol. Toxicol. *18*, 171-179.

Choy,L., Skillington,J., and Derynck,R. (2000). Roles of autocrine TGF-beta receptor and Smad signaling in adipocyte differentiation. J. Cell Biol. *149*, 667-682.

Choy,L. and Derynck,R. (2003). Transforming growth factor-beta inhibits adipocyte differentiation by Smad3 interacting with CCAAT/enhancer-binding protein (C/EBP) and repressing C/EBP transactivation function. J. Biol. Chem. *278*, 9609-9619.

Cinti S. (2001) Morphology of the adipose organ. In Adipose Tissue, S. Klaus, ed. (Georgetown, Texas: Landes Bioscience), pp 11-26

Claudinot,S., Nicolas,M., Oshima,H., Rochat,A., and Barrandon,Y. (2005). Long-term renewal of hair follicles from clonogenic multipotent stem cells. Proc. Natl. Acad. Sci. U.S.A *102*, 14677-14682.

Clements,M.O., Godfrey,A., Crossley,J., Wilson,S.J., Takeuchi,Y., and Boshoff,C. (2006) Lentiviral manipulation of gene expression in human adult and embryonic stem cells. Tissue Eng. May 1.

Clouthier,D.E., Comerford,S.A., and Hammer,R.E. (1997). Hepatic fibrosis, glomerulosclerosis, and a lipodystrophy-like syndrome in PEPCK-TGF-beta1 transgenic mice. J. Clin. Invest *100*, 2697-2713.

Cole,K.A., Harmon,A.W., Harp,J.B., and Patel,Y.M. (2004). Rb regulates C/EBPbeta-DNA-binding activity during 3T3-L1 adipogenesis. Am. J. Physiol Cell Physiol *286*, C349-C354.

Conget,P.A. and Minguell,J.J. (1999). Phenotypical and functional properties of human bone marrow mesenchymal progenitor cells. J. Cell Physiol *181*, 67-73.

Cotsarelis,G., Sun,T.T., and Lavker,R.M. (1990). Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis. Cell *61*, 1329-1337.

Crozat,A., Aman,P., Mandahl,N., and Ron,D. (1993). Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma. Nature *363*, 640-644.

Cullen,L.M. and Arndt,G.M. (2005). Genome-wide screening for gene function using RNAi in mammalian cells. Immunol. Cell Biol. *83*, 217-223.

Curtis,R.K., Oresic,M., and Vidal-Puig,A. (2005). Pathways to the analysis of microarray data. Trends Biotechnol. *23*, 429-435.

Daheron,L., Opitz,S.L., Zaehres,H., Lensch,W.M., Andrews,P.W., Itskovitz-Eldor,J., and Daley,G.Q. (2004). LIF/STAT3 signaling fails to maintain self-renewal of human embryonic stem cells. Stem Cells *22*, 770-778.

Damelin,M., Sun,Y.E., Brundula Sodja,V., and Bestor,T.H. (2005). Decatenation checkpoint deficiency in stem and progenitor cells. Cancer Cell *8*, 479-484.

Davis,K.E., Moldes,M., and Farmer,S.R. (2004). The forkhead transcription factor FoxC2 inhibits white adipocyte differentiation. J. Biol. Chem. *279*, 42453-42461.

De Boer,J., Wang,H.J., and Van Blitterswijk,C. (2004). Effects of Wnt signaling on proliferation and differentiation of human mesenchymal stem cells. Tissue Eng *10*, 393-401.

Dennis,J.E., Merriam,A., Awadallah,A., Yoo,J.U., Johnstone,B., and Caplan,A.I. (1999). A quadripotential mesenchymal progenitor cell isolated from the marrow of an adult mouse. J. Bone Miner. Res. *14*, 700-709.

Dennis,J.E., Carbillet,J.P., Caplan,A.I., and Charbord,P. (2002). The STRO-1+ marrow cell population is multipotential. Cells Tissues. Organs *170*, 73-82.

Dennis,J.E. and Caplan,A.I. (2004). Bone marrow mesenchymal stem cells. In Stem Cells Handbook, S.Sell, ed. (Totowa, NJ: Humana Press), pp. 107-117.

Dicker,A., Le,B.K., Astrom,G., van,H., V, Gotherstrom,C., Blomqvist,L., Arner,P., and Ryden,M. (2005). Functional studies of mesenchymal stem cells derived from adult human adipose tissue. Exp. Cell Res. *308*, 283-290.

Downward,J. (2004). Use of RNA interference libraries to investigate oncogenic signalling in mammalian cells. Oncogene *23*, 8376-8383.

Draper,J.S., Moore,H.D., Ruban,L.N., Gokhale,P.J., and Andrews,P.W. (2004a). Culture and characterization of human embryonic stem cells. Stem Cells Dev. *13*, 325-336.

Draper,J.S., Smith,K., Gokhale,P., Moore,H.D., Maltby,E., Johnson,J., Meisner,L., Zwaka,T.P., Thomson,J.A., and Andrews,P.W. (2004b). Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. Nat. Biotechnol. *22*, 53-54.

Ducy,P., Zhang,R., Geoffroy,V., Ridall,A.L., and Karsenty,G. (1997). Osf2/Cbfa1: a transcriptional activator of osteoblast differentiation. Cell *89*, 747-754.

Ducy,P. (2000). Cbfa1: a molecular switch in osteoblast biology. Dev. Dyn. *219*, 461-471.

Dunne,J., Hanby,A.M., Poulsom,R., Jones,T.A., Sheer,D., Chin,W.G., Da,S.M., Zhao,Q., Beverley,P.C., and Owen,M.J. (1995). Molecular cloning and tissue expression of FAT, the human homologue of the Drosophila fat gene that is located on chromosome 4q34-q35 and encodes a putative adhesion molecule. Genomics *30*, 207-223.

Echeverri,C.J. and Perrimon,N. (2006). High-throughput RNAi screening in cultured cells: a user's guide. Nat. Rev. Genet. *7*, 373-384.

Eisen,M.B., Spellman,P.T., Brown,P.O., and Botstein,D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. U.S.A *95*, 14863-14868.

Ekins,R. and Chu,F.W. (1999). Microarrays: their origins and applications. Trends Biotechnol. *17*, 217-218.

Ekins,R.P. and Chu,F.W. (1991). Multianalyte microspot immunoassay--microanalytical "compact disk" of the future. Clin. Chem. *37*, 1955-1967.

Elbashir,S.M., Harborth,J., Lendeckel,W., Yalcin,A., Weber,K., and Tuschl,T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature *411*, 494-498.

Erices,A., Conget,P., and Minguell,J.J. (2000). Mesenchymal progenitor cells in human umbilical cord blood. Br. J. Haematol. *109*, 235-242.

Etheridge,S.L., Spencer,G.J., Heath,D.J., and Genever,P.G. (2004). Expression profiling and functional analysis of wnt signaling mechanisms in mesenchymal stem cells. Stem Cells *22*, 849-860.

Evans,M.J. and Kaufman,M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154-156.

Fajas,L., Schoonjans,K., Gelman,L., Kim,J.B., Najib,J., Martin,G., Fruchart,J.C., Briggs,M., Spiegelman,B.M., and Auwerx,J. (1999). Regulation of peroxisome proliferator-activated receptor gamma expression by adipocyte differentiation and determination factor 1/sterol regulatory element binding protein 1: implications for adipocyte differentiation and metabolism. Mol. Cell Biol. *19*, 5495-5503.

Fajas,L., Egler,V., Reiter,R., Hansen,J., Kristiansen,K., Debril,M.B., Miard,S., and Auwerx,J. (2002). The retinoblastoma-histone deacetylase 3 complex inhibits PPARgamma and adipocyte differentiation. Dev. Cell *3*, 903-910.

Ferrari,G., Cusella-De Angelis,G., Coletta,M., Paolucci,E., Stornaiuolo,A., Cossu,G., and Mavilio,F. (1998). Muscle regeneration by bone marrow-derived myogenic progenitors. Science *279*, 1528-1530.

Fire,A., Xu,S., Montgomery,M.K., Kostas,S.A., Driver,S.E., and Mello,C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Floyd,Z.E. and Stephens,J.M. (2003). STAT5A promotes adipogenesis in nonprecursor cells and associates with the glucocorticoid receptor during adipocyte differentiation. Diabetes *52*, 308-314.

Font de Mora,J., Porras,A., Ahn,N., and Santos,E. (1997). Mitogen-activated protein kinase activation is not necessary for, but antagonizes, 3T3-L1 adipocytic differentiation. Mol. Cell Biol. *17*, 6068-6075.

Fontaine,C., Dubois,G., Duguay,Y., Helledie,T., Vu-Dac,N., Gervois,P., Soncin,F., Mandrup,S., Fruchart,J.C., Fruchart-Najib,J., and Staels,B. (2003). The orphan nuclear receptor Rev-Erbalpha is a peroxisome proliferator-activated receptor (PPAR) gamma target gene and promotes PPARgamma-induced adipocyte differentiation. J. Biol. Chem. *278*, 37672-37680.

Forman,B.M., Tontonoz,P., Chen,J., Brun,R.P., Spiegelman,B.M., and Evans,R.M. (1995). 15-Deoxy-delta 12, 14-prostaglandin J2 is a ligand for the adipocyte determination factor PPAR gamma. Cell *83*, 803-812.

Friedenstein,A.J., Piatetzky-Shapiro,I.I., and Petrakova,K.V. (1966). Osteogenesis in transplants of bone marrow cells. J. Embryol. Exp. Morphol. *16*, 381-390.

Friedenstein,A.J., Chailakhjan,R.K., and Lalykina,K.S. (1970). The development of fibroblast colonies in monolayer cultures of guinea-pig bone marrow and spleen cells. Cell Tissue Kinet. *3*, 393-403.

Friedenstein,A.J. (1976). Precursor cells of mechanocytes. Int. Rev. Cytol. *47*, 327-359.

Friedenstein,A.J. (1980). Stromal mechanisms of bone marrow: cloning in vitro and retransplantation in vivo. Haematol. Blood Transfus. *25*, 19-29.

Fu,Y., Luo,N., Klein,R.L., and Garvey,W.T. (2005). Adiponectin promotes adipocyte differentiation, insulin sensitivity, and lipid accumulation. J. Lipid Res. *46*, 1369-1379.

Fukuhara,A., Matsuda,M., Nishizawa,M., Segawa,K., Tanaka,M., Kishimoto,K., Matsuki,Y., Murakami,M., Ichisaka,T., Murakami,H., Watanabe,E., Takagi,T., Akiyoshi,M., Ohtsubo,T., Kihara,S., Yamashita,S., Makishima,M., Funahashi,T., Yamanaka,S., Hiramatsu,R., Matsuzawa,Y., and Shimomura,I. (2005). Visfatin: a protein secreted by visceral fat that mimics the effects of insulin. Science *307*, 426-430.

Funes,J.M., Quintero,M., Henderson,S., Martinez,D., Qureschi,U., Westwood,C., Clements,M., Pedley,B., Moncada,S. and Boshoff,C. (2006). Increased glycolysis is a feature of *in vivo* tumor growth, but not *in vitro* transformation of adult stem cells. Submitted (Nat. Med.).

Gang,E.J., Jeong,J.A., Hong,S.H., Hwang,S.H., Kim,S.W., Yang,I.H., Ahn,C., Han,H., and Kim,H. (2004). Skeletal myogenic differentiation of mesenchymal stem cells isolated from human umbilical cord blood. Stem Cells *22*, 617-624.

Gavalas,A., Ruhrberg,C., Livet,J., Henderson,C.E., and Krumlauf,R. (2003). Neuronal defects in the hindbrain of Hoxa1, Hoxb1 and Hoxb2 mutants reflect regulatory interactions among these Hox genes. Development *130*, 5663-5679.

Gentleman,R.C., Carey,V.J., Bates,D.M., Bolstad,B., Dettling,M., Dudoit,S., Ellis,B., Gautier,L., Ge,Y., Gentry,J., Hornik,K., Hothorn,T., Huber,W., Iacus,S., Irizarry,R., Leisch,F., Li,C., Maechler,M., Rossini,A.J., Sawitzki,G., Smith,C., Smyth,G., Tierney,L., Yang,J.Y., and Zhang,J. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. *5*, R80.

Gerhold,D.L., Liu,F., Jiang,G., Li,Z., Xu,J., Lu,M., Sachs,J.R., Bagchi,A., Fridman,A., Holder,D.J., Doebber,T.W., Berger,J., Elbrecht,A., Moller,D.E., and Zhang,B.B. (2002). Gene expression profile of adipocyte differentiation and its regulation by peroxisome proliferator-activated receptor-gamma agonists. Endocrinology *143*, 2106-2118.

Gesta,S., Bluher,M., Yamamoto,Y., Norris,A.W., Berndt,J., Kralisch,S., Boucher,J., Lewis,C., and Kahn,C.R. (2006). Evidence for a role of developmental genes in the origin of obesity and body fat distribution. Proc. Natl. Acad. Sci. U.S.A *103*, 6676-6681.

Gibbs,C.P., Kukekov,V.G., Reith,J.D., Tchigrinova,O., Suslov,O.N., Scott,E.W., Ghivizzani,S.C., Ignatova,T.N., and Steindler,D.A. (2005). Stem-like cells in bone sarcomas: implications for tumorigenesis. Neoplasia. *7*, 967-976.

Gimble,J.M., Morgan,C., Kelly,K., Wu,X., Dandapani,V., Wang,C.S., and Rosen,V. (1995). Bone morphogenetic proteins inhibit adipocyte differentiation by bone marrow stromal cells. J. Cell Biochem. *58*, 393-402.

Gimble,J.M., Robinson,C.E., Wu,X., and Kelly,K.A. (1996). The function of adipocytes in the bone marrow stroma: an update. Bone *19*, 421-428.

Godfrey A.R. (2006). Lentiviral vectors as tools for gene manipulation. PhD Thesis, University of London

Green,H. and Kehinde,O. (1974). Sublines of mouse 3T3 cells that accumulate lipid. Cell *1*, 113-116.

Green,H. and Kehinde,O. (1975). An established preadipose cell line and its differentiation in culture. II. Factors affecting the adipose conversion. Cell *5*, 19-27.

Green,H. and Kehinde,O. (1976). Spontaneous heritable changes leading to increased adipose conversion in 3T3 cells. Cell *7*, 105-113.

Green,H. and Kehinde,O. (1979). Formation of normally differentiated subcutaneous fat pads by an established preadipose cell line. J. Cell Physiol *101*, 169-171.

Grimaldi,P.A. (2001). The roles of PPARs in adipocyte differentiation. Prog. Lipid Res. *40*, 269-281.

Gronthos,S., Zannettino,A.C., Hay,S.J., Shi,S., Graves,S.E., Kortesidis,A., and Simmons,P.J. (2003a). Molecular and cellular characterisation of highly purified stromal stem cells derived from human bone marrow. J. Cell Sci. *116*, 1827-1835.

Gronthos,S., Chen,S., Wang,C.Y., Robey,P.G., and Shi,S. (2003b). Telomerase accelerates osteogenesis of bone marrow stromal stem cells by upregulation of CBFA1, osterix, and osteocalcin. J. Bone Miner. Res. *18*, 716-722.

Hackl,H., Burkard,T.R., Sturn,A., Rubio,R., Schleiffer,A., Tian,S., Quackenbush,J., Eisenhaber,F., and Trajanoski,Z. (2005). Molecular processes during fat cell development revealed by gene expression profiling and functional annotation. Genome Biol. *6*, R108.

Hager,J., Dina,C., Francke,S., Dubois,S., Houari,M., Vatin,V., Vaillant,E., Lorentz,N., Basdevant,A., Clement,K., Guy-Grand,B., and Froguel,P. (1998). A genome-wide scan for human obesity genes reveals a major susceptibility locus on chromosome 10. Nat. Genet. *20*, 304-308.

Hahn,W.C., Counter,C.M., Lundberg,A.S., Beijersbergen,R.L., Brooks,M.W., and Weinberg,R.A. (1999). Creation of human tumour cells with defined genetic elements. Nature *400*, 464-468.

Hahn,W.C. and Weinberg,R.A. (2002). Rules for making human tumor cells. N. Engl. J. Med. *347*, 1593-1603.

Hahn,W.C., Dessain,S.K., Brooks,M.W., King,J.E., Elenbaas,B., Sabatini,D.M., DeCaprio,J.A., and Weinberg,R.A. (2002). Enumeration of the simian virus 40 early region elements necessary for human cell transformation. Mol. Cell Biol. *22*, 2111-2123.

Hall,P.A. and Watt,F.M. (1989). Stem cells: the generation and maintenance of cellular diversity. Development *106*, 619-633.

Hanada,K., Dennis,J.E., and Caplan,A.I. (1997). Stimulatory effects of basic fibroblast growth factor and bone morphogenetic protein-2 on osteogenic differentiation of rat bone marrow-derived mesenchymal stem cells. J. Bone Miner. Res. *12*, 1606-1614.

Hanahan,D. and Weinberg,R.A. (2000). The hallmarks of cancer. Cell *100*, 57-70.

Hannon,G.J. and Rossi,J.J. (2004). Unlocking the potential of the human genome with RNA interference. Nature *431*, 371-378.

Hansen,J.B., Petersen,R.K., Jorgensen,C., and Kristiansen,K. (2002). Deregulated MAPK activity prevents adipocyte differentiation of fibroblasts lacking the retinoblastoma protein. J. Biol. Chem. *277*, 26335-26339.

Harris,C.C. (1996). p53 tumor suppressor gene: from the basic research laboratory to the clinic - an abridged historical perspective. Carcinogenesis *17*, 1187-1198.

Harris,R.G., Herzog,E.L., Bruscia,E.M., Grove,J.E., Van Arnam,J.S., and Krause,D.S. (2004). Lack of a fusion requirement for development of bone marrow-derived epithelia. Science *305*, 90-93.

Heath,V.J., Gillespie,D.A., and Crouch,D.H. (2000). Inhibition of the terminal stages of adipocyte differentiation by cMyc. Exp. Cell Res. *254*, 91-98.

Helman,L.J. and Meltzer,P. (2003). Mechanisms of sarcoma development. Nat. Rev. Cancer *3*, 685-694.

Hermann,A., Gastl,R., Liebau,S., Popa,M.O., Fiedler,J., Boehm,B.O., Maisel,M., Lerche,H., Schwarz,J., Brenner,R., and Storch,A. (2004). Efficient generation of neural stem cell-like cells from adult human bone marrow stromal cells. J. Cell Sci. *117*, 4411-4422.

Higashiyama,S., Abraham,J.A., Miller,J., Fiddes,J.C., and Klagsbrun,M. (1991). A heparin-binding growth factor secreted by macrophage-like cells that is related to EGF. Science *251*, 936-939.

Hinney,A., Ziegler,A., Oeffner,F., Wedewardt,C., Vogel,M., Wulftange,H., Geller,F., Stubing,K., Siegfried,W., Goldschmidt,H.P., Remschmidt,H., and Hebebrand,J. (2000). Independent confirmation of a major locus for obesity on chromosome 10. J. Clin. Endocrinol. Metab *85*, 2962-2965.

Hirai,S., Yamanaka,M., Kawachi,H., Matsui,T., and Yano,H. (2005). Activin A inhibits differentiation of 3T3-L1 preadipocyte. Mol. Cell Endocrinol. *232*, 21-26.

Hoffman,L.M. and Carpenter,M.K. (2005). Characterization and culture of human embryonic stem cells. Nat. Biotechnol. *23*, 699-708.

Hong,J.H., Hwang,E.S., McManus,M.T., Amsterdam,A., Tian,Y., Kalmukova,R., Mueller,E., Benjamin,T., Spiegelman,B.M., Sharp,P.A., Hopkins,N., and Yaffe,M.B.

(2005). TAZ, a transcriptional modulator of mesenchymal stem cell differentiation. Science *309*, 1074-1078.

Horwitz,E.M., Prockop,D.J., Fitzpatrick,L.A., Koo,W.W., Gordon,P.L., Neel,M., Sussman,M., Orchard,P., Marx,J.C., Pyeritz,R.E., and Brenner,M.K. (1999). Transplantability and therapeutic effects of bone marrow-derived mesenchymal cells in children with osteogenesis imperfecta. Nat. Med. *5*, 309-313.

Horwitz,E.M., Gordon,P.L., Koo,W.K., Marx,J.C., Neel,M.D., McNall,R.Y., Muul,L., and Hofmann,T. (2002). Isolated allogeneic bone marrow-derived mesenchymal cells engraft and stimulate growth in children with osteogenesis imperfecta: Implications for cell therapy of bone. Natl. Acad. Sci. U. S.A *99*, 8932-8937.

Hosono,T., Mizuguchi,H., Katayama,K., Koizumi,N., Kawabata,K., Yamaguchi,T., Nakagawa,S., Watanabe,Y., Mayumi,T., and Hayakawa,T. (2005). RNA interference of PPARgamma using fiber-modified adenovirus vector efficiently suppresses preadipocyte-to-adipocyte differentiation in 3T3-L1 cells. Gene *348*, 157-165.

Hu,E., Kim,J.B., Sarraf,P., and Spiegelman,B.M. (1996). Inhibition of adipogenesis through MAP kinase-mediated phosphorylation of PPARgamma. Science *274*, 2100-2103.

Hung,S.C., Cheng,H., Pan,C.Y., Tsai,M.J., Kao,L.S., and Ma,H.L. (2002). In vitro differentiation of size-sieved stem cells into electrically active neural cells. Stem Cells *20*, 522-529.

Hung,S.C., Chang,C.F., Ma,H.L., Chen,T.H., and Low-Tone,H.L. (2004a). Gene expression profiles of early adipogenesis in human mesenchymal stem cells. Gene *340*, 141-150.

Hung,S.C., Yang,D.M., Chang,C.F., Lin,R.J., Wang,J.S., Low-Tone,H.L., and Yang,W.K. (2004b). Immortalization without neoplastic transformation of human mesenchymal stem cells by transduction with HPV16 E6/E7 genes. Int. J. Cancer *110*, 313-319.

Hutley,L. and Prins,J.B. (2005). Fat as an endocrine organ: relationship to the metabolic syndrome. Am. J. Med. Sci. *330*, 280-289.

Ignotz,R.A. and Massague,J. (1985). Type beta transforming growth factor controls the adipogenic differentiation of 3T3 fibroblasts. Natl. Acad. Sci. U. S.A *82*, 8530-8534.

Ikeda,R., Yoshida,K., Tsukahara,S., Sakamoto,Y., Tanaka,H., Furukawa,K., and Inoue,I. (2005). The promyelotic leukemia zinc finger promotes osteoblastic differentiation of human mesenchymal stem cells as an upstream regulator of CBFA1. J. Biol. Chem. *280*, 8523-8530.

Inoue,H., Furukawa,T., Giannakopoulos,S., Zhou,S., King,D.S., and Tanese,N. (2002). Largest subunits of the human SWI/SNF chromatin-remodeling complex promote transcriptional activation by steroid hormone receptors. J. Biol. Chem. *277*, 41674-41685.

Irizarry,R.A., Bolstad,B.M., Collin,F., Cope,L.M., Hobbs,B., and Speed,T.P. (2003a). Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res. *31*, e15.

Irizarry,R.A., Hobbs,B., Collin,F., Beazer-Barclay,Y.D., Antonellis,K.J., Scherf,U., and Speed,T.P. (2003b). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics. *4*, 249-264.

Ishii,M., Koike,C., Igarashi,A., Yamanaka,K., Pan,H., Higashi,Y., Kawaguchi,H., Sugiyama,M., Kamata,N., Iwata,T., Matsubara,T., Nakamura,K., Kurihara,H., Tsuji,K., and Kato,Y. (2005). Molecular markers distinguish bone marrow mesenchymal stem cells from fibroblasts. Biochem. Biophys. Res. Commun. *332*, 297-303.

Itani,O.A., Liu,K.Z., Cornish,K.L., Campbell,J.R., and Thomas,C.P. (2002). Glucocorticoids stimulate human sgk1 gene expression by activation of a GRE in its 5'-flanking region. Am. J. Physiol Endocrinol. Metab *283*, E971-E979.

Ito,M., Liu,Y., Yang,Z., Nguyen,J., Liang,F., Morris,R.J., and Cotsarelis,G. (2005). Stem cells in the hair follicle bulge contribute to wound repair but not to homeostasis of the epidermis. Nat. Med. *11*, 1351-1354.

Jackson,A.L., Bartz,S.R., Schelter,J., Kobayashi,S.V., Burchard,J., Mao,M., Li,B., Cavet,G., and Linsley,P.S. (2003). Expression profiling reveals off-target gene regulation by RNAi. Nat. Biotechnol. *21*, 635-637.

Jackson,K.A., Mi,T., and Goodell,M.A. (1999). Hematopoietic potential of stem cells isolated from murine skeletal muscle. Natl. Acad. Sci. U. S.A *96*, 14482-14486.

Jakkaraju,S., Zhe,X., Pan,D., Choudhury,R., and Schuger,L. (2005). TIPs are tension-responsive proteins involved in myogenic versus adipogenic differentiation. Dev. Cell *9*, 39-49.

Jang,Y.Y., Collector,M.I., Baylin,S.B., Diehl,A.M., and Sharkis,S.J. (2004). Hematopoietic stem cells convert into liver cells within days without fusion. Nat. Cell Biol. *6*, 532-539.

Jiang,Y., Jahagirdar,B.N., Reinhardt,R.L., Schwartz,R.E., Keene,C.D., Ortiz-Gonzalez,X.R., Reyes,M., Lenvik,T., Lund,T., Blackstad,M., Du,J., Aldrich,S., Lisberg,A., Low,W.C., Largaespada,D.A., and Verfaillie,C.M. (2002a). Pluripotency of mesenchymal stem cells derived from adult marrow. Nature *418*, 41-49.

Jiang,Y., Vaessen,B., Lenvik,T., Blackstad,M., Reyes,M., and Verfaillie,C.M. (2002b). Multipotent progenitor cells can be isolated from postnatal murine bone marrow, muscle, and brain. Exp. Hematol. *30*, 896-904.

Jones,E.A., Kinsey,S.E., English,A., Jones,R.A., Straszynski,L., Meredith,D.M., Markham,A.F., Jack,A., Emery,P., and McGonagle,D. (2002). Isolation and characterization of bone marrow multipotential mesenchymal progenitor cells. Arthritis Rheum. *46*, 3349-3360.

Joseph,N.M. and Morrison,S.J. (2005). Toward an understanding of the physiological function of Mammalian stem cells. Dev. Cell *9*, 173-183.

Kajita,M., McClinic,K.N., and Wade,P.A. (2004). Aberrant expression of the transcription factors snail and slug alters the response to genotoxic stress. Mol. Cell Biol. *24*, 7559-7566.

Karanam,S. and Moreno,C.S. (2004). CONFAC: automated application of comparative genomic promoter analysis to DNA microarray datasets. Nucleic Acids Res. *32*, W475-W484.

Kariko,K., Bhuyan,P., Capodici,J., and Weissman,D. (2004). Small interfering RNAs mediate sequence-independent gene suppression and induce immune activation by signaling through toll-like receptor 3. J. Immunol. *172*, 6545-6549.

Kemppainen,R.J., Cox,E., Behrend,E.N., Brogan,M.D., and Ammons,J.M. (2003). Identification of a glucocorticoid response element in the 3'-flanking region of the human Dexras1 gene. Biochim. Biophys. Acta *1627*, 85-89.

Kershaw,E.E. and Flier,J.S. (2004). Adipose tissue as an endocrine organ. J. Clin. Endocrinol. Metab *89*, 2548-2556.

Kim,J.B., Wright,H.M., Wright,M., and Spiegelman,B.M. (1998). ADD1/SREBP1 activates PPARgamma through the production of endogenous ligand. Proc. Natl. Acad. Sci. U.S.A *95*, 4333-4337.

Klaus S. (2001) Overview: Biological significance of fat and adipose tissues. In Adipose Tissue, S. Klaus, ed. (Georgetown, Texas: Landes Bioscience), pp 27-55

Knudsen,S. (2004). Introduction to DNA Microarray Technology. In Guide to Analysis of DNA Microarray Data, S.Knudsen, ed. (Hoboken, NJ: John Wiley and Sons Inc.), pp. 1-22.

Kolfschoten,I.G., van,L.B., Berns,K., Mullenders,J., Beijersbergen,R.L., Bernards,R., Voorhoeve,P.M., and Agami,R. (2005). A genetic screen identifies PITX1 as a suppressor of RAS activity and tumorigenicity. Cell *121*, 849-858.

Kondo,M., Weissman,I.L., and Akashi,K. (1997). Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell *91*, 661-672.

Koopman,R., Schaart,G., and Hesselink,M.K. (2001). Optimisation of oil red O staining permits combination with immunofluorescence and automated quantification of lipids. Histochem. Cell Biol. *116*, 63-68.

Kopen,G.C., Prockop,D.J., and Phinney,D.G. (1999). Marrow stromal cells migrate throughout forebrain and cerebellum, and they differentiate into astrocytes after injection into neonatal mouse brains. Proc. Natl. Acad. Sci. U.S.A *96*, 10711-10716.

Kralisch,S., Klein,J., Lossner,U., Bluher,M., Paschke,R., Stumvoll,M., and Fasshauer,M. (2005). Hormonal regulation of the novel adipocytokine visfatin in 3T3-L1 adipocytes. J. Endocrinol. *185*, R1-R8.

Krampera,M., Pasini,A., Rigo,A., Scupoli,M.T., Tecchio,C., Malpeli,G., Scarpa,A., Dazzi,F., Pizzolo,G., and Vinante,F. (2005). HB-EGF/HER-1 signaling in bone marrow mesenchymal stem cells: inducing cell expansion and reversibly preventing multilineage differentiation. Blood *106*, 59-66.

Krivtsov,A.V., Twomey,D., Feng,Z., Stubbs,M.C., Wang,Y., Faber,J., Levine,J.E., Wang,J., Hahn,W.C., Gilliland,G., Golub,T.R., and Armstrong,S.A. (2006). Transformation from committed progenitor to leukemia stem cell initiated by MLL-AF9. Nature *422*, 818-822.

Kuznetsov,S.A., Mankani,M.H., Gronthos,S., Satomura,K., Bianco,P., and Robey,P.G. (2001). Circulating skeletal stem cells. J. Cell Biol. *153*, 1133-1140.

Kwon,H.S., Huang,B., Unterman,T.G., and Harris,R.A. (2004). Protein kinase B-alpha inhibits human pyruvate dehydrogenase kinase-4 gene induction by dexamethasone through inactivation of FOXO transcription factors. Diabetes *53*, 899-910.

LaBarge,M.A. and Blau,H.M. (2002). Biological progression from adult bone marrow to mononucleate muscle stem cell to multinucleate muscle fiber in response to injury. Cell *111*, 589-601.

Laclef,C., Souil,E., Demignon,J., and Maire,P. (2003). Thymus, kidney and craniofacial abnormalities in Six 1 deficient mice. Mech. Dev. *120*, 669-679.

Lagasse,E., Connors,H., Al-Dhalimy,M., Reitsma,M., Dohse,M., Osborne,L., Wang,X., Finegold,M., Weissman,I.L., and Grompe,M. (2000). Purified hematopoietic stem cells can differentiate into hepatocytes in vivo. Nat. Med. *6*, 1229-1234.

Lagutin,O.V., Zhu,C.C., Kobayashi,D., Topczewski,J., Shimamura,K., Puelles,L., Russell,H.R., McKinnon,P.J., Solnica-Krezel,L., and Oliver,G. (2003). Six3 repression of Wnt signaling in the anterior neuroectoderm is essential for vertebrate forebrain development. Genes Dev. *17*, 368-379.

Lander,E.S., Linton,L.M., Birren,B., Nusbaum,C., Zody,M.C., Baldwin,J., Devon,K., Dewar,K., Doyle,M., FitzHugh,W., Funke,R., Gage,D., Harris,K., Heaford,A., Howland,J., Kann,L., Lehoczky,J., LeVine,R., McEwan,P., McKernan,K., Meldrim,J., Mesirov,J.P., Miranda,C., Morris,W., Naylor,J., Raymond,C., Rosetti,M., Santos,R., Sheridan,A., Sougnez,C., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Wyman,D., Rogers,J., Sulston,J., Ainscough,R., Beck,S., Bentley,D., Burton,J., Clee,C., Carter,N., Coulson,A., Deadman,R., Deloukas,P., Dunham,A., Dunham,I., Durbin,R., French,L., Grafham,D., Gregory,S., Hubbard,T., Humphray,S., Hunt,A., Jones,M., Lloyd,C., McMurray,A., Matthews,L., Mercer,S., Milne,S., Mullikin,J.C., Mungall,A., Plumb,R., Ross,M., Shownkeen,R., Sims,S., Waterston,R.H., Wilson,R.K., Hillier,L.W., McPherson,J.D., Marra,M.A., Mardis,E.R., Fulton,L.A., Chinwalla,A.T., Pepin,K.H., Gish,W.R., Chissoe,S.L., Wendl,M.C., Delehaunty,K.D., Miner,T.L., Delehaunty,A.,

Kramer,J.B., Cook,L.L., Fulton,R.S., Johnson,D.L., Minx,P.J., Clifton,S.W., Hawkins,T., Branscomb,E., Predki,P., Richardson,P., Wenning,S., Slezak,T., Doggett,N., Cheng,J.F., Olsen,A., Lucas,S., Elkin,C., Uberbacher,E., Frazier,M., Gibbs,R.A., Muzny,D.M., Scherer,S.E., Bouck,J.B., Sodergren,E.J., Worley,K.C., Rives,C.M., Gorrell,J.H., Metzker,M.L., Naylor,S.L., Kucherlapati,R.S., Nelson,D.L., Weinstock,G.M., Sakaki,Y., Fujiyama,A., Hattori,M., Yada,T., Toyoda,A., Itoh,T., Kawagoe,C., Watanabe,H., Totoki,Y., Taylor,T., Weissenbach,J., Heilig,R., Saurin,W., Artiguenave,F., Brottier,P., Bruls,T., Pelletier,E., Robert,C., Wincker,P., Smith,D.R., Doucette-Stamm,L., Rubenfield,M., Weinstock,K., Lee,H.M., Dubois,J., Rosenthal,A., Platzer,M., Nyakatura,G., Taudien,S., Rump,A., Yang,H., Yu,J., Wang,J., Huang,G., Gu,J., Hood,L., Rowen,L., Madan,A., Qin,S., Davis,R.W., Federspiel,N.A., Abola,A.P., Proctor,M.J., Myers,R.M., Schmutz,J., Dickson,M., Grimwood,J., Cox,D.R., Olson,M.V., Kaul,R., Raymond,C., Shimizu,N., Kawasaki,K., Minoshima,S., Evans,G.A., Athanasiou,M., Schultz,R., Roe,B.A., Chen,F., Pan,H., Ramser,J., Lehrach,H., Reinhardt,R., McCombie,W.R., de la,B.M., Dedhia,N., Blocker,H., Hornischer,K., Nordsiek,G., Agarwala,R., Aravind,L., Bailey,J.A., Bateman,A., Batzoglou,S., Birney,E., Bork,P., Brown,D.G., Burge,C.B., Cerutti,L., Chen,H.C., Church,D., Clamp,M., Copley,R.R., Doerks,T., Eddy,S.R., Eichler,E.E., Furey,T.S., Galagan,J., Gilbert,J.G., Harmon,C., Hayashizaki,Y., Haussler,D., Hermjakob,H., Hokamp,K., Jang,W., Johnson,L.S., Jones,T.A., Kasif,S., Kaspryzk,A., Kennedy,S., Kent,W.J., Kitts,P., Koonin,E.V., Korf,I., Kulp,D., Lancet,D., Lowe,T.M., McLysaght,A., Mikkelsen,T., Moran,J.V., Mulder,N., Pollara,V.J., Ponting,C.P., Schuler,G., Schultz,J., Slater,G., Smit,A.F., Stupka,E., Szustakowski,J., Thierry-Mieg,D., Thierry-Mieg,J., Wagner,L., Wallis,J., Wheeler,R., Williams,A., Wolf,Y.I., Wolfe,K.H., Yang,S.P., Yeh,R.F., Collins,F., Guyer,M.S., Peterson,J., Felsenfeld,A., Wetterstrand,K.A., Patrinos,A., Morgan,M.J., de,J.P., Catanese,J.J., Osoegawa,K., Shizuya,H., Choi,S., and Chen,Y.J. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Le Jeune,I., Shepherd,M., Van,H.G., Houslay,M.D., and Hall,I.P. (2002). Cyclic AMP-dependent transcriptional up-regulation of phosphodiesterase 4D5 in human airway smooth muscle cells. Identification and characterization of a novel PDE4D5 promoter. J. Biol. Chem. *277*, 35980-35989.

Lee,K., Villena,J.A., Moon,Y.S., Kim,K.H., Lee,S., Kang,C., and Sul,H.S. (2003). Inhibition of adipogenesis and development of glucose intolerance by soluble preadipocyte factor-1 (Pref-1). J. Clin. Invest *111*, 453-461.

Lee,O.K., Kuo,T.K., Chen,W.M., Lee,K.D., Hsieh,S.L., and Chen,T.H. (2004a). Isolation of multipotent mesenchymal stem cells from umbilical cord blood. Blood *103*, 1669-1675.

Lee,R.H., Kim,B., Choi,I., Kim,H., Choi,H.S., Suh,K., Bae,Y.C., and Jung,J.S. (2004b). Characterization and expression analysis of mesenchymal stem cells from human bone marrow and adipose tissue. Cell Physiol Biochem. *14*, 311-324.

Lehmann,J.M., Lenhard,J.M., Oliver,B.B., Ringold,G.M., and Kliewer,S.A. (1997). Peroxisome proliferator-activated receptors alpha and gamma are activated by indomethacin and other non-steroidal anti-inflammatory drugs. J. Biol. Chem. *272*, 3406-3410.

Leptin,M. (1991). twist and snail as positive and negative regulators during Drosophila mesoderm development. Genes Dev. *5*, 1568-1576.

Levy,V., Lindon,C., Harfe,B.D., and Morgan,B.A. (2005). Distinct stem cell populations regenerate the follicle and interfollicular epidermis. Dev. Cell *9*, 855-861.

Li,C. and Wong,W.H. (2001). Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc. Natl. Acad. Sci. U.S.A *98*, 31-36.

Li,D., Yea,S., Li,S., Chen,Z., Narla,G., Banck,M., Laborda,J., Tan,S., Friedman,J.M., Friedman,S.L., and Walsh,M.J. (2005). Kruppel-like factor-6 promotes preadipocyte differentiation through histone deacetylase 3-dependent repression of DLK1. J. Biol. Chem. *280*, 26941-26952.

Liechty,K.W., MacKenzie,T.C., Shaaban,A.F., Radu,A., Moseley,A.M., Deans,R., Marshak,D.R., and Flake,A.W. (2000). Human mesenchymal stem cells engraft and demonstrate site-specific differentiation after in utero transplantation in sheep. Nat. Med. *6*, 1282-1286.

Locklin,R.M., Oreffo,R.O., and Triffitt,J.T. (1999). Effects of TGFbeta and bFGF on the differentiation of human bone marrow stromal fibroblasts. Cell Biol. Int. *23*, 185-194.

Longo,K.A., Wright,W.S., Kang,S., Gerin,I., Chiang,S.H., Lucas,P.C., Opp,M.R., and MacDougald,O.A. (2004). Wnt10b inhibits development of white and brown adipose tissues. J. Biol. Chem. *279*, 35503-35509.

Lu,P., Blesch,A., and Tuszynski,M.H. (2004). Induction of bone marrow stromal cells to neurons: differentiation, transdifferentiation, or artifact? J. Neurosci. Res. *77*, 174-191.

Ludwig,T.E., Levenstein,M.E., Jones,J.M., Berggren,W.T., Mitchen,E.R., Frane,J.L., Crandall,L.J., Daigh,C.A., Conard,K.R., Piekarczyk,M.S., Llanas,R.A., and Thomson,J.A. (2006). Derivation of human embryonic stem cells in defined conditions. Nat. Biotechnol. *24*, 185-187.

Luo,Q., Kang,Q., Si,W., Jiang,W., Park,J.K., Peng,Y., Li,X., Luu,H.H., Luo,J., Montag,A.G., Haydon,R.C., and He,T.C. (2004). Connective tissue growth factor (CTGF) is regulated by Wnt and bone morphogenetic proteins signaling in osteoblast differentiation of mesenchymal stem cells. J. Biol. Chem. *279*, 55958-55968.

Magun,R., Burgering,B.M., Coffer,P.J., Pardasani,D., Lin,Y., Chabot,J., and Sorisky,A. (1996). Expression of a constitutively activated form of protein kinase B (c-Akt) in 3T3-L1 preadipose cells causes spontaneous differentiation. Endocrinology *137*, 3590-3593.

Majumdar,M.K., Keane-Moore,M., Buyaner,D., Hardy,W.B., Moorman,M.A., McIntosh,K.R., and Mosca,J.D. (2003). Characterization and functionality of cell surface molecules on human mesenchymal stem cells. J. Biomed. Sci. *10*, 228-241.

Makino,S., Fukuda,K., Miyoshi,S., Konishi,F., Kodama,H., Pan,J., Sano,M., Takahashi,T., Hori,S., Abe,H., Hata,J., Umezawa,A., and Ogawa,S. (1999).

Cardiomyocytes can be generated from marrow stromal cells in vitro. J. Clin. Invest *103*, 697-705.

Malkin,D., Li,F.P., Strong,L.C., Fraumeni,J.F., Jr., Nelson,C.E., Kim,D.H., Kassel,J., Gryka,M.A., Bischoff,F.Z., Tainsky,M.A., and . (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science *250*, 1233-1238.

Mankoo,B.S., Collins,N.S., Ashby,P., Grigorieva,E., Pevny,L.H., Candia,A., Wright,C.V., Rigby,P.W., and Pachnis,V. (1999). Mox2 is a component of the genetic hierarchy controlling limb muscle development. Nature *400*, 69-73.

Mankoo,B.S., Skuntz,S., Harrigan,I., Grigorieva,E., Candia,A., Wright,C.V., Arnheiter,H., and Pachnis,V. (2003). The concerted action of Meox homeobox genes is required upstream of genetic pathways essential for the formation, patterning and differentiation of somites. Development *130*, 4655-4664.

Marshman,E., Booth,C., and Potten,C.S. (2002). The intestinal epithelial stem cell. Bioessays *24*, 91-98.

Martin,G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Natl. Acad. Sci. U. S.A *78*, 7634-7638.

Marx,J. (2003). Cancer research. Mutant stem cells may seed cancer. Science *301*, 1308-1310.

Matin,M.M., Walsh,J.R., Gokhale,P.J., Draper,J.S., Bahrami,A.R., Morton,I., Moore,H.D., and Andrews,P.W. (2004). Specific knockdown of Oct4 and beta2-microglobulin expression by RNA interference in human embryonic stem cells and embryonic carcinoma cells. Stem Cells *22*, 659-668.

Matsumoto,S., Kishida,K., Shimomura,I., Maeda,N., Nagaretani,H., Matsuda,M., Nishizawa,H., Kihara,S., Funahashi,T., Matsuzawa,Y., Yamada,A., Yamashita,S., Tamura,S., and Kawata,S. (2002). Increased plasma HB-EGF associated with obesity and coronary artery disease. Biochem. Biophys. Res. Commun. *292*, 781-786.

McBeath,R., Pirone,D.M., Nelson,C.M., Bhadriraju,K., and Chen,C.S. (2004). Cell shape, cytoskeletal tension, and RhoA regulate stem cell lineage commitment. Dev. Cell *6*, 483-495.

Mezey,E., Chandross,K.J., Harta,G., Maki,R.A., and McKercher,S.R. (2000). Turning blood into brain: cells bearing neuronal antigens generated in vivo from bone marrow. Science *290*, 1779-1782.

Mi,H., Lazareva-Ulitsky,B., Loo,R., Kejariwal,A., Vandergriff,J., Rabkin,S., Guo,N., Muruganujan,A., Doremieux,O., Campbell,M.J., Kitano,H., and Thomas,P.D. (2005). The PANTHER database of protein families, subfamilies, functions and pathways. Nucleic Acids Res. *33*, D284-D288.

Minguell,J.J., Erices,A., and Conget,P. (2001). Mesenchymal stem cells. Exp. Biol. Med. (Maywood) 226, 507-520.

Miura,M., Gronthos,S., Zhao,M., Lu,B., Fisher,L.W., Robey,P.G., and Shi,S. (2003). SHED: stem cells from human exfoliated deciduous teeth. Proc. Natl. Acad. Sci. U.S.A 100, 5807-5812.

Miura,M., Miura,Y., Padilla-Nash,H.M., Molinolo,A.A., Fu,B., Patel,V., Seo,B.M., Sonoyama,W., Zheng,J.J., Baker,C.C., Chen,W., Ried,T., and Shi,S. (2006). Accumulated chromosomal instability in murine bone marrow mesenchymal stem cells leads to malignant transformation. Stem Cells 24, 1095-1103.

Miyamoto,T., Weissman,I.L., and Akashi,K. (2000). AML1/ETO-expressing nonleukemic stem cells in acute myelogenous leukemia with 8;21 chromosomal translocation. Proc. Natl. Acad. Sci. U.S.A 97, 7521-7526.

Mocellin,S., Rossi,C.R., Brandes,A., and Nitti,D. (2006). Adult soft tissue sarcomas: conventional therapies and molecularly targeted approaches. Cancer Treat. Rev. 32, 9-27.

Moldes,M., Lasnier,F., Feve,B., Pairault,J., and Djian,P. (1997). Id3 prevents differentiation of preadipose cells. Mol. Cell Biol. 17, 1796-1804.

Moldes,M., Boizard,M., Liepvre,X.L., Feve,B., Dugail,I., and Pairault,J. (1999). Functional antagonism between inhibitor of DNA binding (Id) and adipocyte determination and differentiation factor 1/sterol regulatory element-binding protein-1c (ADD1/SREBP-1c) trans-factors for the regulation of fatty acid synthase promoter in adipocytes. Biochem. J. 344 Pt 3, 873-880.

Montague,C.T., Farooqi,I.S., Whitehead,J.P., Soos,M.A., Rau,H., Wareham,N.J., Sewter,C.P., Digby,J.E., Mohammed,S.N., Hurst,J.A., Cheetham,C.H., Earley,A.R., Barnett,A.H., Prins,J.B., and O'rahilly,S. (1997). Congenital leptin deficiency is associated with severe early-onset obesity in humans. Nature 387, 903-908.

Moore,K.A., and Lemischka,I.R. (2006). Stem cells and their niches. Science 311, 1880-1885.

Mori,T., Sakaue,H., Iguchi,H., Gomi,H., Okada,Y., Takashima,Y., Nakamura,K., Nakamura,T., Yamauchi,T., Kubota,N., Kadowaki,T., Matsuki,Y., Ogawa,W., Hiramatsu,R., and Kasuga,M. (2005a). Role of Kruppel-like factor 15 (KLF15) in transcriptional regulation of adipogenesis. J. Biol. Chem. 280, 12867-12875.

Mori,T., Kiyono,T., Imabayashi,H., Takeda,Y., Tsuchiya,K., Miyoshi,S., Makino,H., Matsumoto,K., Saito,H., Ogawa,S., Sakamoto,M., Hata,J., and Umezawa,A. (2005b). Combination of hTERT and bmi-1, E6, or E7 induces prolongation of the life span of bone marrow stromal cells from an elderly donor without affecting their neurogenic potential. Mol. Cell Biol. 25, 5183-5195.

Morris,R.J., Liu,Y., Marles,L., Yang,Z., Trempus,C., Li,S., Lin,J.S., Sawicki,J.A., and Cotsarelis,G. (2004). Capturing and profiling adult hair follicle stem cells. Nat. Biotechnol. *22*, 411-417.

Morrison,R.F. and Farmer,S.R. (1999). Role of PPARgamma in regulating a cascade expression of cyclin-dependent kinase inhibitors, p18(INK4c) and p21(Waf1/Cip1), during adipogenesis. J. Biol. Chem. *274*, 17088-17097.

Morrison,S.J. and Weissman,I.L. (1994). The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. Immunity. *1*, 661-673.

Morrison,S.J., Shah,N.M., and Anderson,D.J. (1997). Regulatory mechanisms in stem cell biology. Cell *88*, 287-298.

Morrison,S.J., White,P.M., Zock,C., and Anderson,D.J. (1999). Prospective identification, isolation by flow cytometry, and in vivo self-renewal of multipotent mammalian neural crest stem cells. Cell *96*, 737-749.

Mueller,E., Drori,S., Aiyer,A., Yie,J., Sarraf,P., Chen,H., Hauser,S., Rosen,E.D., Ge,K., Roeder,R.G., and Spiegelman,B.M. (2002). Genetic analysis of adipogenesis through peroxisome proliferator-activated receptor gamma isoforms. J. Biol. Chem. *277*, 41925-41930.

Nakae,J., Kitamura,T., Kitamura,Y., Biggs,W.H., III, Arden,K.C., and Accili,D. (2003). The forkhead transcription factor Foxo1 regulates adipocyte differentiation. Dev. Cell *4*, 119-129.

Nakamura,T., Shiojima,S., Hirai,Y., Iwama,T., Tsuruzoe,N., Hirasawa,A., Katsuma,S., and Tsujimoto,G. (2003). Temporal gene expression changes during adipogenesis in human mesenchymal stem cells. Biochem. Biophys. Res. Commun. *303*, 306-312.

Napoli,C.A., Lemieux,C., and Jorgensen,R. (1990). Introduction of a chimeric chalcone synthetase gene in Petunia results in reversible cosuppression of homologous genes *in trans*. Plant Cell. *2*, 279-289.

Neubauer,M., Fischbach,C., Bauer-Kreisel,P., Lieb,E., Hacker,M., Tessmar,J., Schulz,M.B., Goepferich,A., and Blunk,T. (2004). Basic fibroblast growth factor enhances PPARgamma ligand-induced adipogenesis of mesenchymal stem cells. FEBS Lett. *577*, 277-283.

Neuhuber,B., Gallo,G., Howard,L., Kostura,L., Mackay,A., and Fischer,I. (2004). Reevaluation of in vitro differentiation protocols for bone marrow stromal cells: disruption of actin cytoskeleton induces rapid morphological changes and mimics neuronal phenotype. J. Neurosci. Res. *77*, 192-204.

Ngo,V.N., Davis,R.E., Lamy,L., Yu,X., Zhao,H., Lenz,G., Lam,L.T., Dave,S., Yang,L., Powell,J., and Staudt,L.M. (2006). A loss-of-function RNA interference screen for molecular targets in cancer. Nature *441*, 106-110.

Nichols,J., Zevnik,B., Anastassiadis,K., Niwa,H., Klewe-Nebenius,D., Chambers,I., Scholer,H., and Smith,A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell 95, 379-391.

Noth,U., Osyczka,A.M., Tuli,R., Hickok,N.J., Danielson,K.G., and Tuan,R.S. (2002). Multilineage mesenchymal differentiation potential of human trabecular bone-derived cells. J. Orthop. Res. 20, 1060-1069.

Ntambi,J.M. and Young-Cheul,K. (2000). Adipocyte differentiation and gene expression. J. Nutr. 130, 3122S-3126S.

Nuttall,M.E. and Gimble,J.M. (2000). Is there a therapeutic opportunity to either prevent or treat osteopenic disorders by inhibiting marrow adipogenesis? Bone 27, 177-184.

Ohyama,M., Terunuma,A., Tock,C.L., Radonovich,M.F., Pise-Masison,C.A., Hopping,S.B., Brady,J.N., Udey,M.C., and Vogel,J.C. (2006). Characterization and isolation of stem cell-enriched human hair follicle bulge cells. J. Clin. Invest 116, 249-260.

Okamoto,T., Aoyama,T., Nakayama,T., Nakamata,T., Hosaka,T., Nishijo,K., Nakamura,T., Kiyono,T., and Toguchida,J. (2002). Clonal heterogeneity in differentiation potential of immortalized human mesenchymal stem cells. Biochem. Biophys. Res. Commun. 295, 354-361.

Okuyama,H., Krishnamachary,B., Zhou,Y.F., Nagasawa,H., Bosch-Marce,M., and Semenza,G.L. (2006). Expression of Vascular Endothelial Growth Factor Receptor 1 in Bone Marrow-derived Mesenchymal Cells Is Dependent on Hypoxia-inducible Factor 1. J. Biol. Chem. 281, 15554-15563.

Oliveira,D.M. and Goodell,M.A. (2003). Transient RNA interference in hematopoietic progenitors with functional consequences. Genesis. 36, 203-208.

Orlic,D., Kajstura,J., Chimenti,S., Jakoniuk,I., Anderson,S.M., Li,B., Pickel,J., McKay,R., Nadal-Ginard,B., Bodine,D.M., Leri,A., and Anversa,P. (2001). Bone marrow cells regenerate infarcted myocardium. Nature 410, 701-705.

Osawa,M., Hanada,K., Hamada,H., and Nakauchi,H. (1996). Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. Science 273, 242-245.

Oshima,H., Rochat,A., Kedzia,C., Kobayashi,K., and Barrandon,Y. (2001). Morphogenesis and renewal of hair follicles from adult multipotent stem cells. Cell 104, 233-245.

Owen,M. (1978). Histogenesis of bone cells. Calcif. Tissue Res. 25, 205-207.

Owen,M. (1985). Lineage of osteogenic cells and their relationship to the stromal system. In Bone and Mineral Research Volume 3, W.A.Peck, ed. (Amsterdam: Elsevier), pp. 1-25.

Owen,M. (1988). Marrow stromal stem cells. J. Cell Sci. Suppl 10, 63-76.

Paddison,P.J., Caudy,A.A., and Hannon,G.J. (2002). Stable suppression of gene expression by RNAi in mammalian cells. Proc. Natl. Acad. Sci. U.S.A *99*, 1443-1448.

Paddison,P.J., Silva,J.M., Conklin,D.S., Schlabach,M., Li,M., Aruleba,S., Balija,V., O'Shaughnessy,A., Gnoj,L., Scobie,K., Chang,K., Westbrook,T., Cleary,M., Sachidanandam,R., McCombie,W.R., Elledge,S.J., and Hannon,G.J. (2004). A resource for large-scale RNA-interference-based screens in mammals. Nature *428*, 427-431.

Palmqvist,L., Glover,C.H., Hsu,L., Lu,M., Bossen,B., Piret,J.M., Humphries,R.K., and Helgason,C.D. (2005). Correlation of murine embryonic stem cell gene expression profiles with functional measures of pluripotency. Stem Cells *23*, 663-680.

Pardal,R., Clarke,M.F., and Morrison,S.J. (2003). Applying the principles of stem-cell biology to cancer. Nat. Rev. Cancer *3*, 895-902.

Passegue,E., Jamieson,C.H., Ailles,L.E., and Weissman,I.L. (2003). Normal and leukemic hematopoiesis: are leukemias a stem cell disorder or a reacquisition of stem cell characteristics? Proc. Natl. Acad. Sci. U.S.A *100 Suppl 1*, 11842-11849.

Pereira,R.C., Delany,A.M., and Canalis,E. (2004). CCAAT/enhancer binding protein homologous protein (DDIT3) induces osteoblastic cell differentiation. Endocrinology *145*, 1952-1960.

Pereira,R.F., Halford,K.W., O'Hara,M.D., Leeper,D.B., Sokolov,B.P., Pollard,M.D., Bagasra,O., and Prockop,D.J. (1995). Cultured adherent cells from marrow can serve as long-lasting precursor cells for bone, cartilage, and lung in irradiated mice. Proc. Natl. Acad. Sci. U.S.A *92*, 4857-4861.

Pereira,R.F., O'Hara,M.D., Laptev,A.V., Halford,K.W., Pollard,M.D., Class,R., Simon,D., Livezey,K., and Prockop,D.J. (1998). Marrow stromal cells as a source of progenitor cells for nonhematopoietic tissues in transgenic mice with a phenotype of osteogenesis imperfecta. Proc. Natl. Acad. Sci. U.S.A *95*, 1142-1147.

Perez-Mancera,P.A. and Sanchez-Garcia,I. (2005). Understanding mesenchymal cancer: the liposarcoma-associated FUS-DDIT3 fusion gene as a model. Semin. Cancer Biol. *15*, 206-214.

Persengiev,S.P., Zhu,X., and Green,M.R. (2004). Nonspecific, concentration-dependent stimulation and repression of mammalian gene expression by small interfering RNAs (siRNAs). RNA. *10*, 12-18.

Phinney,D.G., Kopen,G., Righter,W., Webster,S., Tremain,N., and Prockop,D.J. (1999). Donor variation in the growth properties and osteogenic potential of human marrow stromal cells. J. Cell Biochem. *75*, 424-436.

Pinhasov,A., Mandel,S., Torchinsky,A., Giladi,E., Pittel,Z., Goldsweig,A.M., Servoss,S.J., Brenneman,D.E., and Gozes,I. (2003). Activity-dependent neuroprotective protein: a novel gene essential for brain formation. Brain Res. Dev. Brain Res. *144*, 83-90.

Pittenger,M.F., Mackay,A.M., Beck,S.C., Jaiswal,R.K., Douglas,R., Mosca,J.D., Moorman,M.A., Simonetti,D.W., Craig,S., and Marshak,D.R. (1999). Multilineage potential of adult human mesenchymal stem cells. Science *284*, 143-147.

Pittenger,M.F. and Marshak,D.R. (2001) Mesenchymal stem cells of adult bone marrow. In Stem Cell Biology, D.R.Marshak, R.L.Gardner, and D.Gottlieb, eds. (New York: Cold Spring Harbor Laboratory Press), pp. 349-375.

Polyak,K. and Hahn,W.C. (2006). Roots and stems: stem cells in cancer. Nat. Med. *12*, 296-300.

Porras,A., Muszynski,K., Rapp,U.R., and Santos,E. (1994). Dissociation between activation of Raf-1 kinase and the 42-kDa mitogen-activated protein kinase/90-kDa S6 kinase (MAPK/RSK) cascade in the insulin/Ras pathway of adipocytic differentiation of 3T3 L1 cells. J. Biol. Chem. *269*, 12741-12748.

Postigo,A.A. (2003). Opposing functions of ZEB proteins in the regulation of the TGFbeta/BMP signaling pathway. EMBO J. *22*, 2443-2452.

Potten,C.S. and Loeffler,M. (1990). Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. Development *110*, 1001-1020.

Qi,H., Aguiar,D.J., Williams,S.M., La,P.A., Pan,W., and Verfaillie,C.M. (2003). Identification of genes responsible for osteoblast differentiation from human mesodermal progenitor cells. Proc. Natl. Acad. Sci. U.S.A *100*, 3305-3310.

Radtke,F. and Clevers,H. (2005). Self-renewal and cancer of the gut: two sides of a coin. Science *307*, 1904-1909.

Rangarajan,A., Hong,S.J., Gifford,A., and Weinberg,R.A. (2004). Species- and cell type-specific requirements for cellular transformation. Cancer Cell *6*, 171-183.

Reubinoff,B.E., Pera,M.F., Fong,C.Y., Trounson,A., and Bongso,A. (2000). Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. Nat. Biotechnol. *18*, 399-404.

Reusch,J.E., Colton,L.A., and Klemm,D.J. (2000). CREB activation induces adipogenesis in 3T3-L1 cells. Mol. Cell Biol. *20*, 1008-1020.

Reya,T., Morrison,S.J., Clarke,M.F., and Weissman,I.L. (2001). Stem cells, cancer, and cancer stem cells. Nature *414*, 105-111.

Reya,T. and Clevers,H. (2005). Wnt signalling in stem cells and cancer. Nature *434*, 843-850.

Reyes,M., Lund,T., Lenvik,T., Aguiar,D., Koodie,L., and Verfaillie,C.M. (2001). Purification and ex vivo expansion of postnatal human marrow mesodermal progenitor cells. Blood *98*, 2615-2625.

313

Reyes,M., Dudek,A., Jahagirdar,B., Koodie,L., Marker,P.H., and Verfaillie,C.M. (2002). Origin of endothelial progenitors in human postnatal bone marrow. J. Clin. Invest *109*, 337-346.

Reynolds,A., Leake,D., Boese,Q., Scaringe,S., Marshall,W.S., and Khvorova,A. (2004). Rational siRNA design for RNA interference. Nat. Biotechnol. *22*, 326-330.

Reznikoff,C.A., Brankow,D.W., and Heidelberger,C. (1973). Establishment and characterization of a cloned line of C3H mouse embryo cells sensitive to postconfluence inhibition of division. Cancer Res. *33*, 3231-3238.

Ricquier,D. and Bouillaud,F. (2000). The uncoupling protein homologues: UCP1, UCP2, UCP3, StUCP and AtUCP. Biochem. J. *345 Pt 2*, 161-179.

Rochford,J.J., Semple,R.K., Laudes,M., Boyle,K.B., Christodoulides,C., Mulligan,C., Lelliott,C.J., Schinner,S., Hadaschik,D., Mahadevan,M., Sethi,J.K., Vidal-Puig,A., and O'rahilly,S. (2004). ETO/MTG8 is an inhibitor of C/EBPbeta activity and a regulator of early adipogenesis. Mol. Cell Biol. *24*, 9863-9872.

Rodriguez,A.M., Elabd,C., Delteil,F., Astier,J., Vernochet,C., Saint-Marc,P., Guesnet,J., Guezennec,A., Amri,E.Z., Dani,C., and Ailhaud,G. (2004). Adipocyte differentiation of multipotent cells established from human adipose tissue. Biochem. Biophys. Res. Commun. *315*, 255-263.

Roelen,B.A. and Dijke,P. (2003). Controlling mesenchymal stem cell differentiation by TGFBeta family members. J. Orthop. Sci. *8*, 740-748.

Rosen,E.D., Sarraf,P., Troy,A.E., Bradwin,G., Moore,K., Milstone,D.S., Spiegelman,B.M., and Mortensen,R.M. (1999). PPAR gamma is required for the differentiation of adipose tissue in vivo and in vitro. Mol. Cell *4*, 611-617.

Rosen,E.D. and Spiegelman,B.M. (2000). Molecular regulation of adipogenesis. Annu. Rev. Cell Dev. Biol. *16*, 145-171.

Rosen,E.D., Hsu,C.H., Wang,X., Sakai,S., Freeman,M.W., Gonzalez,F.J., and Spiegelman,B.M. (2002). C/EBPalpha induces adipogenesis through PPARgamma: a unified pathway. Genes Dev. *16*, 22-26.

Ross,S.E., Hemati,N., Longo,K.A., Bennett,C.N., Lucas,P.C., Erickson,R.L., and MacDougald,O.A. (2000). Inhibition of adipogenesis by Wnt signaling. Science *289*, 950-953.

Ross,S.E., Erickson,R.L., Gerin,I., DeRose,P.M., Bajnok,L., Longo,K.A., Misek,D.E., Kuick,R., Hanash,S.M., Atkins,K.B., Andresen,S.M., Nebb,H.I., Madsen,L., Kristiansen,K., and MacDougald,O.A. (2002). Microarray analyses during adipogenesis: understanding the effects of Wnt signaling on adipogenesis and the roles of liver X receptor alpha in adipocyte metabolism. Mol. Cell Biol. *22*, 5989-5999.

314

Roufosse,C.A., Direkze,N.C., Otto,W.R., and Wright,N.A. (2004). Circulating mesenchymal stem cells. Int. J. Biochem. Cell Biol. *36*, 585-597.

Rubin,C.S., Lai,E., and Rosen,O.M. (1977). Acquisition of increased hormone sensitivity during in vitro adipocyte development. J. Biol. Chem. *252*, 3554-3557.

Rubin,C.S., Hirsch,A., Fung,C., and Rosen,O.M. (1978). Development of hormone receptors and hormonal responsiveness in vitro. Insulin receptors and insulin sensitivity in the preadipocyte and adipocyte forms of 3T3-L1 cells. J. Biol. Chem. *253*, 7570-7578.

Rubinson,D.A., Dillon,C.P., Kwiatkowski,A.V., Sievers,C., Yang,L., Kopinja,J., Rooney,D.L., Ihrig,M.M., McManus,M.T., Gertler,F.B., Scott,M.L., and Van,P.L. (2003). A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. Nat. Genet. *33*, 401-406.

Rubio,D., Garcia-Castro,J., Martin,M.C., de la,F.R., Cigudosa,J.C., Lloyd,A.C., and Bernad,A. (2005). Spontaneous human adult stem cell transformation. Cancer Res. *65*, 3035-3039.

Russell,T.R. and Ho,R. (1976). Conversion of 3T3 fibroblasts into adipose cells: triggering of differentiation by prostaglandin F2alpha and 1-methyl-3-isobutyl xanthine. Proc. Natl. Acad. Sci. U.S.A *73*, 4516-4520.

Saad,S., Stevens,V.A., Wassef,L., Poronnik,P., Kelly,D.J., Gilbert,R.E., and Pollock,C.A. (2005). High glucose transactivates the EGF receptor and up-regulates serum glucocorticoid kinase in the proximal tubule. Kidney Int. *68*, 985-997.

Safford,K.M., Hicok,K.C., Safford,S.D., Halvorsen,Y.D., Wilkison,W.O., Gimble,J.M., and Rice,H.E. (2002). Neurogenic differentiation of murine and human adipose-derived stromal cells. Biochem. Biophys. Res. Commun. *294*, 371-379.

Sanchez-Ramos,J., Song,S., Cardozo-Pelaez,F., Hazzi,C., Stedeford,T., Willing,A., Freeman,T.B., Saporta,S., Janssen,W., Patel,N., Cooper,D.R., and Sanberg,P.R. (2000). Adult bone marrow stromal cells differentiate into neural cells in vitro 10. Exp. Neurol. *164*, 247-256.

Schena,M., Shalon,D., Davis,R.W., and Brown,P.O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science *270*, 467-470.

Scherr,M., Battmer,K., Ganser,A., and Eder,M. (2003). Modulation of gene expression by lentiviral-mediated delivery of small interfering RNA. Cell Cycle *2*, 251-257.

Schwartz,R.E., Reyes,M., Koodie,L., Jiang,Y., Blackstad,M., Lund,T., Lenvik,T., Johnson,S., Hu,W.S., and Verfaillie,C.M. (2002). Multipotent adult progenitor cells from bone marrow differentiate into functional hepatocyte-like cells. J. Clin. Invest *109*, 1291-1302.

315

Sekiya,I., Vuoristo,J.T., Larson,B.L., and Prockop,D.J. (2002). In vitro cartilage formation by human adult stem cells from bone marrow stroma defines the sequence of cellular and molecular events during chondrogenesis. Proc. Natl. Acad. Sci. U.S.A 99, 4397-4402.

Sekiya,I., Larson,B.L., Vuoristo,J.T., Cui,J.G., and Prockop,D.J. (2004). Adipogenic differentiation of human adult stem cells from bone marrow stroma (MSCs). J. Bone Miner. Res. 19, 256-264.

Sell S. (2004) Stem Cells. In Stem Cells Handbook, S.Sell, ed. (Totowa, NJ: Humana Press), pp. 1-17.

Seo,M.J., Suh,S.Y., Bae,Y.C., and Jung,J.S. (2005). Differentiation of human adipose stromal cells into hepatic lineage in vitro and in vivo. Biochem. Biophys. Res. Commun. 328, 258-264.

Serakinci,N., Guldberg,P., Burns,J.S., Abdallah,B., Schrodder,H., Jensen,T., and Kassem,M. (2004). Adult human mesenchymal stem cell as a target for neoplastic transformation. Oncogene 23, 5095-5098.

Shay,J.W. and Bacchetti,S. (1997). A survey of telomerase activity in human cancer. Eur. J. Cancer 33, 787-791.

Shen,C., Buck,A.K., Liu,X., Winkler,M., and Reske,S.N. (2003). Gene silencing by adenovirus-delivered siRNA. FEBS Lett. 539, 111-114.

Shi,S., Gronthos,S., Chen,S., Reddi,A., Counter,C.M., Robey,P.G., and Wang,C.Y. (2002). Bone formation by human postnatal bone marrow stromal stem cells is enhanced by telomerase expression. Nat. Biotechnol. 20, 587-591.

Shi,X., Shi,W., Li,Q., Song,B., Wan,M., Bai,S., and Cao,X. (2003). A glucocorticoid-induced leucine-zipper protein, GILZ, inhibits adipogenesis of mesenchymal cells. EMBO Rep. 4, 374-380.

Shi,Y. (2003). Mammalian RNAi for the masses. Trends Genet. 19, 9-12.

Siminovitch,L., McCulloch,E.A., and Till,J.E. (1963). The distribution of colony-forming cells among spleen colonies. J. Cell Physiol 62, 327-336.

Simmons,P.J. and Torok-Storb,B. (1991). Identification of stromal cell precursors in human bone marrow by a novel monoclonal antibody, STRO-1. Blood 78, 55-62.

Simonsen,J.L., Rosada,C., Serakinci,N., Justesen,J., Stenderup,K., Rattan,S.I., Jensen,T.G., and Kassem,M. (2002). Telomerase expression extends the proliferative life-span and maintains the osteogenic potential of human bone marrow stromal cells. Nat. Biotechnol. 20, 592-596.

Singh,S.K., Clarke,I.D., Terasaki,M., Bonn,V.E., Hawkins,C., Squire,J., and Dirks,P.B. (2003). Identification of a cancer stem cell in human brain tumors. Cancer Res. 63, 5821-5828.

Singh,S.K., Hawkins,C., Clarke,I.D., Squire,J.A., Bayani,J., Hide,T., Henkelman,R.M., Cusimano,M.D., and Dirks,P.B. (2004). Identification of human brain tumour initiating cells. Nature *432*, 396-401.

Sledz,C.A., Holko,M., de Veer,M.J., Silverman,R.H., and Williams,B.R. (2003). Activation of the interferon system by short-interfering RNAs. Nat. Cell Biol. *5*, 834-839.

Sledz,C.A. and Williams,B.R. (2005). RNA interference in biology and disease. Blood *106*, 787-794.

Smith,A.G., Heath,J.K., Donaldson,D.D., Wong,G.G., Moreau,J., Stahl,M., and Rogers,D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature *336*, 688-690.

Smyth,G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. *3*, Article3.

Solchaga,L.A., Penick,K., Porter,J.D., Goldberg,V.M., Caplan,A.I., and Welter,J.F. (2005). FGF-2 enhances the mitotic and chondrogenic potentials of human adult bone marrow-derived mesenchymal stem cells. J. Cell Physiol *203*, 398-409.

Song,L., Webb,N.E., Song,Y., and Tuan,R.S. (2006). Identification and functional analysis of candidate genes regulating mesenchymal stem cell self-renewal and multipotency. Stem Cells.

Sordella,R., Jiang,W., Chen,G.C., Curto,M., and Settleman,J. (2003). Modulation of Rho GTPase signaling regulates a switch between adipogenesis and myogenesis. Cell *113*, 147-158.

Soukas,A., Socci,N.D., Saatkamp,B.D., Novelli,S., and Friedman,J.M. (2001). Distinct transcriptional profiles of adipogenesis in vivo and in vitro. J. Biol. Chem. *276*, 34167-34174.

Spangrude,G.J., Heimfeld,S., and Weissman,I.L. (1988). Purification and characterization of mouse hematopoietic stem cells. Science *241*, 58-62.

Spiegelman,B.M. and Flier,J.S. (1996). Adipogenesis and obesity: rounding out the big picture. Cell *87*, 377-389.

Stears,R.L., Martinsky,T., and Schena,M. (2003). Trends in microarray analysis. Nat. Med. *9*, 140-145.

Takagi,T., Moribe,H., Kondoh,H., and Higashi,Y. (1998). DeltaEF1, a zinc finger and homeodomain transcription factor, is required for skeleton patterning in multiple lineages. Development *125*, 21-31.

Tan,S.H., Reverter,A., Wang,Y., Byrne,K.A., McWilliam,S.M., and Lehnert,S.A. (2006). Gene expression profiling of bovine in vitro adipogenesis using a cDNA microarray. Funct. Integr. Genomics 1-15.

317

Tanaka,T., Yoshida,N., Kishimoto,T., and Akira,S. (1997). Defective adipocyte differentiation in mice lacking the C/EBPbeta and/or C/EBPdelta gene. EMBO J. *16*, 7432-7443.

Tang,Q.Q. and Lane,M.D. (1999). Activation and centromeric localization of CCAAT/enhancer-binding proteins during the mitotic clonal expansion of adipocyte differentiation. Genes Dev. *13*, 2231-2241.

Tang,Q.Q. and Lane,M.D. (2000). Role of C/EBP homologous protein (CHOP-10) in the programmed activation of CCAAT/enhancer-binding protein-beta during adipogenesis. Proc. Natl. Acad. Sci. U.S.A *97*, 12446-12450.

Tang,Q.Q., Otto,T.C., and Lane,M.D. (2004). Commitment of C3H10T1/2 pluripotent stem cells to the adipocyte lineage. Proc. Natl. Acad. Sci. U.S.A *101*, 9607-9611.

Tavassoli,M. and Crosby,W.H. (1968). Transplantation of marrow to extramedullary sites. Science *161*, 54-56.

Taylor,G., Lehrer,M.S., Jensen,P.J., Sun,T.T., and Lavker,R.M. (2000). Involvement of follicular stem cells in forming not only the follicle but also the epidermis. Cell *102*, 451-461.

Taylor,S.M. and Jones,P.A. (1979). Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine. Cell *17*, 771-779.

Thomas,D. and Kansara,M. (2006). Epigenetic modifications in osteogenic differentiation and transformation. J. Cell Biochem.

Thomson,J.A., Itskovitz-Eldor,J., Shapiro,S.S., Waknitz,M.A., Swiergiel,J.J., Marshall,V.S., and Jones,J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Till,J.E. and McCulloch,E.A. (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. Radiat. Res. *14*, 213-222.

Toma,C., Pittenger,M.F., Cahill,K.S., Byrne,B.J., and Kessler,P.D. (2002). Human mesenchymal stem cells differentiate to a cardiomyocyte phenotype in the adult murine heart. Circulation *105*, 93-98.

Tong,Q., Dalgin,G., Xu,H., Ting,C.N., Leiden,J.M., and Hotamisligil,G.S. (2000). Function of GATA transcription factors in preadipocyte-adipocyte transition. Science *290*, 134-138.

Tong,Q., Tsai,J., Tan,G., Dalgin,G., and Hotamisligil,G.S. (2005). Interaction between GATA and the C/EBP family of transcription factors is critical in GATA-mediated suppression of adipocyte differentiation. Mol. Cell Biol. *25*, 706-715.

Tontonoz,P., Hu,E., Graves,R.A., Budavari,A.I., and Spiegelman,B.M. (1994a). mPPAR gamma 2: tissue-specific regulator of an adipocyte enhancer. Genes Dev. *8*, 1224-1234.

318

Tontonoz,P., Hu,E., and Spiegelman,B.M. (1994b). Stimulation of adipogenesis in fibroblasts by PPAR gamma 2, a lipid-activated transcription factor. Cell 79, 1147-1156.

Trayhurn,P. and Beattie,J.H. (2001). Physiological role of adipose tissue: white adipose tissue as an endocrine and secretory organ. Proc. Nutr. Soc. 60, 329-339.

Tsutsumi,S., Shimazu,A., Miyazaki,K., Pan,H., Koike,C., Yoshida,E., Takagishi,K., and Kato,Y. (2001). Retention of multilineage differentiation potential of mesenchymal cells during proliferation in response to FGF. Biochem. Biophys. Res. Commun. 288, 413-419.

Tumbar,T., Guasch,G., Greco,V., Blanpain,C., Lowry,W.E., Rendl,M., and Fuchs,E. (2004). Defining the epithelial stem cell niche in skin. Science 303, 359-363.

Uchida,N., Buck,D.W., He,D., Reitsma,M.J., Masek,M., Phan,T.V., Tsukamoto,A.S., Gage,F.H., and Weissman,I.L. (2000). Direct isolation of human central nervous system stem cells. Proc. Natl. Acad. Sci. U.S.A 97, 14720-14725.

Urist,M.R. and McLean,F.C. (1952). Osteogenetic potency and new-bone formation by induction in transplants to the anterior chamber of the eye. J. Bone Joint Surg. Am. 34-A, 443-476.

Urs,S., Smith,C., Campbell,B., Saxton,A.M., Taylor,J., Zhang,B., Snoddy,J., Jones,V.B., and Moustaid-Moussa,N. (2004). Gene expression profiling in human preadipocytes and adipocytes by microarray analysis. J. Nutr. 134, 762-770.

Vaidyanathan,G., Cismowski,M.J., Wang,G., Vincent,T.S., Brown,K.D., and Lanier,S.M. (2004). The Ras-related protein AGS1/RASD1 suppresses cell growth. Oncogene 23, 5858-5863.

van der Krol,A.R., Mur,L.A., de Lange,P., Mol,J.N., and Stuitje,A.R. (1990). Inhibition of flower pigmentation by antisense CHS genes: promoter and minimal sequence requirements for the antisense effect. Plant Mol. Biol. 14, 457-466.

Venter,J.C., Adams,M.D., Myers,E.W., Li,P.W., Mural,R.J., Sutton,G.G., Smith,H.O., Yandell,M., Evans,C.A., Holt,R.A., Gocayne,J.D., Amanatides,P., Ballew,R.M., Huson,D.H., Wortman,J.R., Zhang,Q., Kodira,C.D., Zheng,X.H., Chen,L., Skupski,M., Subramanian,G., Thomas,P.D., Zhang,J., Gabor Miklos,G.L., Nelson,C., Broder,S., Clark,A.G., Nadeau,J., McKusick,V.A., Zinder,N., Levine,A.J., Roberts,R.J., Simon,M., Slayman,C., Hunkapiller,M., Bolanos,R., Delcher,A., Dew,I., Fasulo,D., Flanigan,M., Florea,L., Halpern,A., Hannenhalli,S., Kravitz,S., Levy,S., Mobarry,C., Reinert,K., Remington,K., bu-Threideh,J., Beasley,E., Biddick,K., Bonazzi,V., Brandon,R., Cargill,M., Chandramouliswaran,I., Charlab,R., Chaturvedi,K., Deng,Z., Di,F., V, Dunn,P., Eilbeck,K., Evangelista,C., Gabrielian,A.E., Gan,W., Ge,W., Gong,F., Gu,Z., Guan,P., Heiman,T.J., Higgins,M.E., Ji,R.R., Ke,Z., Ketchum,K.A., Lai,Z., Lei,Y., Li,Z., Li,J., Liang,Y., Lin,X., Lu,F., Merkulov,G.V., Milshina,N., Moore,H.M., Naik,A.K., Narayan,V.A., Neelam,B., Nusskern,D., Rusch,D.B., Salzberg,S., Shao,W., Shue,B., Sun,J., Wang,Z., Wang,A., Wang,X., Wang,J., Wei,M., Wides,R., Xiao,C., Yan,C., Yao,A., Ye,J., Zhan,M., Zhang,W., Zhang,H., Zhao,Q., Zheng,L., Zhong,F., Zhong,W., Zhu,S., Zhao,S., Gilbert,D., Baumhueter,S., Spier,G., Carter,C., Cravchik,A.,

Woodage,T., Ali,F., An,H., Awe,A., Baldwin,D., Baden,H., Barnstead,M., Barrow,I., Beeson,K., Busam,D., Carver,A., Center,A., Cheng,M.L., Curry,L., Danaher,S., Davenport,L., Desilets,R., Dietz,S., Dodson,K., Doup,L., Ferriera,S., Garg,N., Gluecksmann,A., Hart,B., Haynes,J., Haynes,C., Heiner,C., Hladun,S., Hostin,D., Houck,J., Howland,T., Ibegwam,C., Johnson,J., Kalush,F., Kline,L., Koduru,S., Love,A., Mann,F., May,D., McCawley,S., McIntosh,T., McMullen,I., Moy,M., Moy,L., Murphy,B., Nelson,K., Pfannkoch,C., Pratts,E., Puri,V., Qureshi,H., Reardon,M., Rodriguez,R., Rogers,Y.H., Romblad,D., Ruhfel,B., Scott,R., Sitter,C., Smallwood,M., Stewart,E., Strong,R., Suh,E., Thomas,R., Tint,N.N., Tse,S., Vech,C., Wang,G., Wetter,J., Williams,S., Williams,M., Windsor,S., Winn-Deen,E., Wolfe,K., Zaveri,J., Zaveri,K., Abril,J.F., Guigo,R., Campbell,M.J., Sjolander,K.V., Karlak,B., Kejariwal,A., Mi,H., Lazareva,B., Hatton,T., Narechania,A., Diemer,K., Muruganujan,A., Guo,N., Sato,S., Bafna,V., Istrail,S., Lippert,R., Schwartz,R., Walenz,B., Yooseph,S., Allen,D., Basu,A., Baxendale,J., Blick,L., Caminha,M., Carnes-Stine,J., Caulk,P., Chiang,Y.H., Coyne,M., Dahlke,C., Mays,A., Dombroski,M., Donnelly,M., Ely,D., Esparham,S., Fosler,C., Gire,H., Glanowski,S., Glasser,K., Glodek,A., Gorokhov,M., Graham,K., Gropman,B., Harris,M., Heil,J., Henderson,S., Hoover,J., Jennings,D., Jordan,C., Jordan,J., Kasha,J., Kagan,L., Kraft,C., Levitsky,A., Lewis,M., Liu,X., Lopez,J., Ma,D., Majoros,W., McDaniel,J., Murphy,S., Newman,M., Nguyen,T., Nguyen,N., and Nodell,M. (2001). The sequence of the human genome. Science *291*, 1304-1351.

Viguerie,N., Vidal,H., Arner,P., Holst,C., Verdich,C., Avizou,S., Astrup,A., Saris,W.H., Macdonald,I.A., Klimcakova,E., Clement,K., Martinez,A., Hoffstedt,J., Sorensen,T.I., and Langin,D. (2005). Adipose tissue gene expression in obese subjects during low-fat and high-fat hypocaloric diets. Diabetologia *48*, 123-131.

Vogelstein,B. and Kinzler,K.W. (2004). Cancer genes and the pathways they control. Nat. Med. *10*, 789-799.

Wabitsch,M., Bruderlein,S., Melzner,I., Braun,M., Mechtersheimer,G., and Moller,P. (2000). LiSa-2, a novel human liposarcoma cell line with a high capacity for terminal adipose differentiation. Int. J. Cancer *88*, 889-894.

Wagers,A.J. and Weissman,I.L. (2004). Plasticity of adult stem cells. Cell *116*, 639-648.

Wagner,W., Wein,F., Seckinger,A., Frankhauser,M., Wirkner,U., Krause,U., Blake,J., Schwager,C., Eckstein,V., Ansorge,W., and Ho,A.D. (2005). Comparative characteristics of mesenchymal stem cells from human bone marrow, adipose tissue, and umbilical cord blood. Exp. Hematol. *33*, 1402-1416.

Wakitani,S., Goto,T., Pineda,S.J., Young,R.G., Mansour,J.M., Caplan,A.I., and Goldberg,V.M. (1994). Mesenchymal cell-based repair of large, full-thickness defects of articular cartilage. J. Bone Joint Surg. Am. *76*, 579-592.

Wang,E.A., Israel,D.I., Kelly,S., and Luxenberg,D.P. (1993). Bone morphogenetic protein-2 causes commitment and differentiation in C3H10T1/2 and 3T3 cells. Growth Factors *9*, 57-71.

Wang,P.P., Wang,J.H., Yan,Z.P., Hu,M.Y., Lau,G.K., Fan,S.T., and Luk,J.M. (2004). Expression of hepatocyte-like phenotypes in bone marrow stromal cells after HGF induction. Biochem. Biophys. Res. Commun. *320*, 712-716.

Weissman,I.L., Anderson,D.J., and Gage,F. (2001). Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations. Annu. Rev. Cell Dev. Biol. *17*, 387-403.

Westbrook,T.F., Martin,E.S., Schlabach,M.R., Leng,Y., Liang,A.C., Feng,B., Zhao,J.J., Roberts,T.M., Mandel,G., Hannon,G.J., Depinho,R.A., Chin,L., and Elledge,S.J. (2005). A genetic screen for candidate tumor suppressors identifies REST. Cell *121*, 837-848.

Williams,B.R. and Haque,S.J. (1997). Interacting pathways of interferon signaling. Semin. Oncol. *24*, S9.

Williams,I.H. and Polakis,S.E. (1977). Differentiation of 3T3-L1 fibroblasts to adipocytes. The effect of indomethacin, prostaglandin E1 and cyclic AMP on the process of differentiation. Biochem. Biophys. Res. Commun. *77*, 175-186.

Williams,R.L., Hilton,D.J., Pease,S., Willson,T.A., Stewart,C.L., Gearing,D.P., Wagner,E.F., Metcalf,D., Nicola,N.A., and Gough,N.M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature *336*, 684-687.

Woodbury,D., Schwarz,E.J., Prockop,D.J., and Black,I.B. (2000). Adult rat and human bone marrow stromal cells differentiate into neurons. J. Neurosci. Res. *61*, 364-370.

Wotton,D., Lo,R.S., Lee,S., and Massague,J. (1999). A Smad transcriptional corepressor. Cell *97*, 29-39.

Wu,G.D., Nolta,J.A., Jin,Y.S., Barr,M.L., Yu,H., Starnes,V.A., and Cramer,D.V. (2003). Migration of mesenchymal stem cells to heart allografts during chronic rejection. Transplantation *75*, 679-685.

Wu,J., Srinivasan,S.V., Neumann,J.C., and Lingrel,J.B. (2005). The KLF2 transcription factor does not affect the formation of preadipocytes but inhibits their differentiation into adipocytes. Biochemistry *44*, 11098-11105.

Wu,Z., Xie,Y., Bucher,N.L., and Farmer,S.R. (1995). Conditional ectopic expression of C/EBP beta in NIH-3T3 cells induces PPAR gamma and stimulates adipogenesis. Genes Dev. *9*, 2350-2363.

Wu,Z., Bucher,N.L., and Farmer,S.R. (1996). Induction of peroxisome proliferator-activated receptor gamma during the conversion of 3T3 fibroblasts into adipocytes is mediated by C/EBPbeta, C/EBPdelta, and glucocorticoids. Mol. Cell Biol. *16*, 4128-4136.

Wu,Z., Rosen,E.D., Brun,R., Hauser,S., Adelmant,G., Troy,A.E., McKeon,C., Darlington,G.J., and Spiegelman,B.M. (1999). Cross-regulation of C/EBP alpha and

PPAR gamma controls the transcriptional pathway of adipogenesis and insulin sensitivity. Mol. Cell *3*, 151-158.

Wu,Z. and Irizarry,R.A. (2004). Preprocessing of oligonucleotide array data. Nat. Biotechnol. *22*, 656-658.

Xu,R.H., Peck,R.M., Li,D.S., Feng,X., Ludwig,T., and Thomson,J.A. (2005). Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. Nat. Methods *2*, 185-190.

Xu,W., Zhang,X., Qian,H., Zhu,W., Sun,X., Hu,J., Zhou,H., and Chen,Y. (2004). Mesenchymal stem cells from adult human bone marrow differentiate into a cardiomyocyte phenotype in vitro. Exp. Biol. Med. (Maywood.) *229*, 623-631.

Xu,Y., Mirmalek-Sani,S.H., Zhang,J., and Oreffo,R.O. (2006). The use of small interfering RNAs to inhibit adipocyte differentiation in human preadipocytes and fetal-femur-derived mesenchymal cells. Exp. Cell Res.

Yau,T.O., Chan,C.Y., Chan,K.L., Lee,M.F., Wong,C.M., Fan,S.T., and Ng,I.O. (2005). HDPR1, a novel inhibitor of the WNT/beta-catenin signaling, is frequently downregulated in hepatocellular carcinoma: involvement of methylation-mediated gene silencing. Oncogene *24*, 1607-1614.

Yeh,W.C., Cao,Z., Classon,M., and McKnight,S.L. (1995). Cascade regulation of terminal adipocyte differentiation by three members of the C/EBP family of leucine zipper proteins. Genes Dev. *9*, 168-181.

Yin,Y., Yuan,H., Wang,C., Pattabiraman,N., Rao,M., Pestell,R.G., and Glazer,R.I. (2006). 3-phosphoinositide-dependent protein kinase-1 activates the peroxisome proliferator-activated receptor-gamma and promotes adipocyte differentiation. Mol. Endocrinol. *20*, 268-278.

Ying,Q.L., Nichols,J., Chambers,I., and Smith,A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell *115*, 281-292.

Ylostalo,J., Smith,J.R., Pochampally,R.R., Matz,R., Sekiya,I., Larson,B.L., Vuoristo,J.T., and Prockop,D.J. (2006). Use of differentiating adult stem cells (marrow stromal cells) to identify new downstream target genes for transcription factors. Stem Cells *24*, 642-652.

Yoo,E.J., Chung,J.J., Choe,S.S., Kim,K.H., and Kim,J.B. (2006). Down-regulation of histone deacetylases stimulates adipocyte differentiation. J. Biol. Chem. *281*, 6608-6615.

Yoshida,Y., Tanaka,S., Umemori,H., Minowa,O., Usui,M., Ikematsu,N., Hosoda,E., Imamura,T., Kuno,J., Yamashita,T., Miyazono,K., Noda,M., Noda,T., and Yamamoto,T. (2000). Negative regulation of BMP/Smad signaling by Tob in osteoblasts. Cell *103*, 1085-1097.

322

Young,H.E., Steele,T.A., Bray,R.A., Hudson,J., Floyd,J.A., Hawkins,K., Thomas,K., Austin,T., Edwards,C., Cuzzourt,J., Duenzl,M., Lucas,P.A., and Black,A.C., Jr. (2001). Human reserve pluripotent mesenchymal stem cells are present in the connective tissues of skeletal muscle and dermis derived from fetal, adult, and geriatric donors. Anat. Rec. *264*, 51-62.

Young,R.G., Butler,D.L., Weber,W., Caplan,A.I., Gordon,S.L., and Fink,D.J. (1998). Use of mesenchymal stem cells in a collagen matrix for Achilles tendon repair. J. Orthop. Res. *16*, 406-413.

Zeeberg,B.R., Feng,W., Wang,G., Wang,M.D., Fojo,A.T., Sunshine,M., Narasimhan,S., Kane,D.W., Reinhold,W.C., Lababidi,S., Bussey,K.J., Riss,J., Barrett,J.C., and Weinstein,J.N. (2003). GoMiner: a resource for biological interpretation of genomic and proteomic data. Genome Biol. *4*, R28.

Zhang,B., Kirov,S., and Snoddy,J. (2005). WebGestalt: an integrated system for exploring gene sets in various biological contexts. Nucleic Acids Res. *33*, W741-W748.

Zhang,Y., Proenca,R., Maffei,M., Barone,M., Leopold,L., and Friedman,J.M. (1994). Positional cloning of the mouse obese gene and its human homologue. Nature *372*, 425-432.

Zhao,J.J., Roberts,T.M., and Hahn,W.C. (2004). Functional genetics and experimental models of human cancer. Trends Mol. Med. *10*, 344-350.

Zhao,L., Gregoire,F., and Sul,H.S. (2000). Transient induction of ENC-1, a Kelch-related actin-binding protein, is required for adipocyte differentiation. J. Biol. Chem. *275*, 16845-16850.

Zhou,L.B., Halvorsen,Y.D., Cryan,E.V., Pelton,P.D., Burris,T.P., and Demarest,K.T. (1999). Analysis of the pattern of gene expression during human adipogenesis by DNA microarray. Biotechnology Techniques *13*, 513-517.

Zinszner,H., Albalat,R., and Ron,D. (1994). A novel effector domain from the RNA-binding protein TLS or EWS is required for oncogenic transformation by CHOP. Genes Dev. *8*, 2513-2526.

Zou,G.M. and Yoder,M.C. (2005). Application of RNA interference to study stem cell function: current status and future perspectives. Biol. Cell *97*, 211-219.

Zufferey,R., Dull,T., Mandel,R.J., Bukovsky,A., Quiroz,D., Naldini,L., and Trono,D. (1998). Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. J. Virol. *72*, 9873–9880.

Zuk,P.A., Zhu,M., Mizuno,H., Huang,J., Futrell,J.W., Katz,A.J., Benhaim,P., Lorenz,H.P., and Hedrick,M.H. (2001). Multilineage cells from human adipose tissue: implications for cell-based therapies. Tissue Eng *7*, 211-228.

Zuk,P.A., Zhu,M., Ashjian,P., De Ugarte,D.A., Huang,J.I., Mizuno,H., Alfonso,Z.C., Fraser,J.K., Benhaim,P., and Hedrick,M.H. (2002). Human adipose tissue is a source of multipotent stem cells. Mol. Biol. Cell *13*, 4279-4295.

Zurita,M., Vaquero,J., Oya,S., and Miguel,M. (2005). Schwann cells induce neuronal differentiation of bone marrow stromal cells. Neuroreport *16*, 505-508.

Zvaifler,N.J., Marinova-Mutafchieva,L., Adams,G., Edwards,C.J., Moss,J., Burger,J.A., and Maini,R.N. (2000). Mesenchymal precursor cells in the blood of normal individuals. Arthritis Res. *2*, 477-488.

Websites used in this thesis:
Oligoengine: www.oligoengine.com
Genecards: www.genecards.org
NetAffix: www.affymetrix.com/analysis/index.affx
OMIM, EntrezGene and PubMed: all at www.ncbi.nlm.nih.gov
Ingenuity® Systems, www.ingenuity.com

APPENDIX

The following legends describe Tables, Folders and Figures that can be found on the CD-ROM attached into the inside back cover of this thesis.

Appendix Folder 1. .cel files for each timepoint replicate, containing the raw (unprocessed) signal data produced from Affymetrix microarray scanning.

Appendix Folder 2. Alternative viewing option (see Appendix Figure 1) of hierarchical clustering of significantly changed genes, represented as cluster dendrogram and heatmap. The complete heatmap and cluster dendrogram, searchable by gene name, can be viewed by downloading and installing the Treeview Software (Eisen *et al* 1998) from http://rana.lbl.gov/EisenSoftware.htm (free for non-commercial, academic and not-for-profit users). To view the heatmap, open the "hc.cdt" file in Appendix Folder 2 using the Treeview software.

Appendix Figure 1. Hierarchical clustering of significantly changed genes, represented as cluster dendrogram and heatmap. The heatmap has been segmented due to it's large size, and is displayed as the 17 nodes of the cluster dendrogram that form the 14 expression clusters used for analysis in this thesis. For each node, it's position in the cluster dendrogram is indicated by a purple colour.

Appendix Table 1. Full list of genes identified as having at least one significant expression change between two consecutive timepoints in the microarray time-course study. Significant expression changes were identified by comparing the expression of each gene between consecutive timepoints (i.e. 0h-1h, 1h-3h, 3h-9h, 9h-7d). The significance cut-off used was $q<0.05$, and q values were calculated as described in section 2.7.3. Only the 5000 most significant changes in the 9h-7d phase are included in this list as further changes were not included in the hierarchical clustering analysis. Negative fold-change values indicate down-regulated gene expression between two consecutive timepoints (due to the $\log_2$ scale of RMA expression units). NA indicates probesets with no gene annotation (these may represent ESTs, for example).

Appendix Table 2. Expression values for significantly changed gene on each array replicate. Expression values are represented in RMA units (log$_2$). NA indicates probesets with no gene annotation.

Appendix Table 3. Gene lists for each expression cluster. These gene lists were used to retrieve Gene Ontology annotations and perform GO category enrichment calculations.

Appendix Table 4. Gene Ontology Biological Process, Molecular Function and Cellular Component categories that were significantly enriched in each expression cluster (p<0.01 for clusters 3 and 10, p<0.05, for all other clusters, and containing at least 10% of the genes in that cluster). GO category annotations were retrieved and significant enrichment calculated as described in section 2.7.4.1. Numbers preceding GO terms relate to the level in the GO hierarchical tree at which the GO term was found. Numbers proceeding expression clusters are the total number of genes in the node/number of genes with annotated functions. Rows in red indicate GO terms passing both significance criteria. Observed: number of genes observed in an expression cluster with a specific GO term; Expected: number of genes annotated with that GO term expected to be observed, given the size of the cluster and the number of genes with the specific GO term in the reference list (all genes on the U133 Plus 2.0 GeneChip). Enrichment ratio: observed/expected; p-value, calculated using hypergeometric test.