Induced pluripotent stem cells — opportunities for disease modelling and drug discovery

Marica Grskovic*¦¦, Ashkan Javaherian*¦¦, Berta Strulovici‡ and George Q. Daley§||¶#**¦¦

Abstract | The ability to generate induced pluripotent stem cells (iPSCs) from patients, and an increasingly refined capacity to differentiate these iPSCs into disease-relevant cell types, promises a new paradigm in drug development — one that positions human disease pathophysiology at the core of preclinical drug discovery. Disease models derived from iPSCs that manifest cellular disease phenotypes have been established for several monogenic diseases, but iPSCs can likewise be used for phenotype-based drug screens in complex diseases for which the underlying genetic mechanism is unknown. Here, we highlight recent advances as well as limitations in the use of iPSC technology for modelling a 'disease in a dish' and for testing compounds against human disease phenotypes *in vitro*. We discuss how iPSCs are being exploited to illuminate disease pathophysiology, identify novel drug targets and enhance the probability of clinical success of new drugs.

*iPierian, 951 Gateway Blvd, South San Francisco, California 94080, USA.
‡Weizmann Institute of Science, Rehovot 76100, Israel.
§Stem Cell Transplantation Program, Division of Pediatric Haematology/Oncology, Manton Center for Orphan Disease Research, Children's Hospital Boston and Dana Farber Cancer Institute; Boston, Massachusetts 02115, USA.
||Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA.
¶Division of Hematology, Brigham and Women's Hospital; Boston, Massachusetts 02115, USA.
#Department of Biological Chemistry and Molecular Pharmacology, Harvard Stem Cell Institute, Harvard Medical School, Boston, Massachusetts 02115, USA.
**Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA.
¦¦These authors contributed equally to this work.
Correspondence to G.Q.D. e-mail: George.Daley@childrens.harvard.edu
doi:10.1038/nrd3577
Published online 11 November 2011

Current approaches to target-driven drug discovery typically involve screening a large compound library against a single enzyme or receptor, followed by prioritization of hits based on chemical tractability and optimization through medicinal chemistry to achieve potency and selectivity. Typically, at a later stage in preclinical development, candidate compounds are tested in a relevant primary cell system or animal model, often with disappointing results. However, some notable successes include drugs such as the kinase inhibitor sorafenib (Nexavar; Bayer)[1], the HIV integrase inhibitor raltegravir (Isentress; Merck)[2] and several drugs targeting G protein-coupled receptors, such as olanzapine (Zyprexa; Lilly), desloratadine (Clarinex; Schering-Plough) and ranitidine (Zantac; GlaxoSmithKline)[3]. In contrast to target-directed biochemical screens, the use of cell-based phenotypic assays has increased, driven by advances in high-content image processing and the appreciation that cell-based screens represent a more physiological system for high-throughput screening and lead optimization.

Cell-based assays have advantages over single-target biochemical assays as they simultaneously confirm cell permeability and tolerable toxicity, but they have many associated limitations. The lack of a known target in cell-based screens hinders the ability to determine the chemical groups responsible for biological activity through structure–activity relationship studies, and requires purely empirical attempts to optimize drug-like properties through trial and error. Although there is a compelling rationale for the use of human primary — preferably disease-bearing — cells for drug screening, lead discovery and optimization, most primary cells are difficult to access and have a finite lifespan in culture. Adaptation of human primary cells to immortal growth in culture — a requisite for their use in cell-based screens — typically entails selection for genetic alterations that may influence the cell's response to drugs, thus compromising the fidelity of drug screens and counterscreens.

The availability of pluripotent embryonic stem cells (ESCs)[4], along with their capacity for unlimited proliferation in culture and their potential to differentiate into any human cell type — neurons, cardiomyocytes or hepatocytes[5] — has provided a potentially invaluable drug discovery tool, particularly for *in vitro* compound toxicity testing during preclinical development[6]. However, most ESCs represent generic cells that are not reflective of any specific clinical condition. The revolutionary discovery of induced pluripotent stem cells (iPSCs)[7–11], whereby a patient's somatic cells can be reprogrammed into an embryonic pluripotent state by the forced expression of a defined set of transcription factors, has the transformative potential to enable *in vitro* disease modelling, inform drug discovery and provide a resource for cell replacement therapy. Recent reports have demonstrated that recapitulating a disease phenotype *in vitro* is feasible

© 2011 Macmillan Publishers Limited. All rights reserved

for numerous monogenic diseases[12–20]. The availability of such disease-relevant pathological cells from actual patients — in limitless amounts —could substantially benefit drug discovery (FIG. 1).

In this Review we discuss the therapeutic potential of iPSCs for drug discovery, with special emphasis on their potential to improve the probability that drugs discovered through human cell-based phenotypic assays will show efficacy in the clinic. In addition, we address current challenges associated with the generation and use of iPSCs, and discuss future perspectives on iPSC technology in drug discovery and development.

ESCs

In 1998 Thomson and colleagues isolated human ESCs from human blastocysts[21], providing the medical research community with a formidable master stock of cells for deriving large quantities of a range of differentiated adult cell types. Using defined growth factors and matrices, it is possible to differentiate ESCs into a desired cell type, rendering them a powerful model for human developmental biology, drug discovery and cell replacement therapy[5,22]. For example, human cardiomyocytes, hepatocytes or neurons can be derived from ESCs and used in drug toxicity studies, thereby providing improved physiological models that are potentially more reliable for the prediction of clinical outcomes[4].

Of particular interest are human ESC-derived neurons, a previously unobtainable cell type that can now be used in primary screens for the discovery of neurodegenerative disease-modulating or psychotropic drugs. Notably, McNeish *et al.*[23] have carried out a high-throughput drug screen of over two million compounds in ESC-derived neurons. Although the screen was carried out in mouse ESC-derived neurons, validation of chemical leads and secondary assays was performed using human ESC-derived neurons. This study demonstrates the feasibility of using human pluripotent stem cells (PSCs) and paves the way for their application in drug screening and optimization. In addition, ESCs derived from known genotypes can be used to assess the variability in response to drugs or examine the molecular mechanisms underlying genetic diseases. Over the past decade, both academic and industry groups have assembled banks of human ESCs from healthy and genetically defective embryos identified through pre-implantation genetic testing, thereby capturing a multitude of diseases (TABLE 1).

Although ESCs offer a considerable theoretical advantage over commonly used immortalized or primary cells, their more widespread use has faced notable obstacles. First, human ESCs generate considerable debate, prompting only limited and cautious investments into this technology by large pharmaceutical companies. Second, only rare single-gene disorders are represented among the disease-specific ESC lines, as multifactorial diseases cannot be discerned by pre-implantation diagnostics. Last, it remains challenging to expand and differentiate ESCs on a large scale, and ESCs have been shown to accumulate karyotypic aberrations[24–26]; such problems are common to all PSCs and are discussed in detail below.

Derivation of iPSCs

In 2006 Shinya Yamanaka[7] demonstrated that the retroviral-mediated introduction of four transcription factors into mouse fibroblasts could convert them into cells closely resembling pluripotent ESCs — a revolutionary breakthrough that has sparked immense interest (FIG. 1). In an elegantly designed set of experiments, Takahashi and Yamanaka tested 24 genes that have previously been implicated in the biology of ESCs for their ability to reprogramme somatic cells into ESC-like cells. They found that the retroviral-mediated introduction of a combination of four transcription factors —octamer binding protein 4 (also known as POU5F1), SOX2, Krüppel-like factor 4 and MYC — was sufficient to induce expression of the pluripotency programme in somatic cells and yield colonies with similar morphology and growth characteristics to ESCs. These induced cells — iPSCs — were able to form embryoid bodies *in vitro* and teratomas *in vivo*, and contributed to the formation of diverse tissues in chimeric embryos when they were injected into mouse blastocysts. However, these early iPSCs were not faithfully reprogrammed; they failed to fully activate some key pluripotency genes and, when injected into mouse blastocysts, did not yield live chimaeras[27,28]. Subsequent reports refined the original derivation methods to produce bona fide PSCs that chimerized mouse embryos and thus fulfilled the classical functional criteria of ESCs[27–31]. Indeed, when iPSCs are injected into tetraploid blastocysts they are capable of generating entire mice, indicating that they, similarly to ESCs, possess full developmental potential[32–34].

Reports of human iPSCs followed within a year[7–11,35], and human iPSCs have now been generated from several human tissues using a variety of approaches (TABLE 2). Most commonly, human iPSCs are generated from dermal fibroblasts owing to their accessibility and relatively high efficiency of reprogramming, with peripheral blood also emerging as an attractive source of donor cells[36,37]. Recently, iPSCs have been generated from Epstein–Barr virus (EBV)-immortalized B cell lines, providing the opportunity to obtain samples from disease cohort repositories such as the Coriell Institute for Medical Research or the UK Biobank[38,39]. Two major hurdles stand in the way of reliable, consistent derivation of iPSCs on an industrial scale: the reliance on viral vectors as the most efficient means to deliver reprogramming factors, and the overall inefficiency of the reprogramming process itself. The most widely practiced method, transduction of reprogramming factors via a retrovirus or lentivirus, results in random integration of foreign genetic elements into the genome, which may cause insertional mutagenesis and inadvertently affect the differentiation of iPSCs into relevant cell types. Random integrations render each iPSC line genetically distinct. Although the integrated viruses are transcriptionally silenced following reprogramming, re-expression of any of the reprogramming factors may interfere with differentiation and subsequent cell behaviour[40].

Alternative approaches, such as the use of single[41] or multiple transient transfections[42], non-integrating

Embryonic stem cells
(ESCs). Pluripotent cells derived from a pre-implantation-stage embryo. These cells are capable of dividing without differentiating for a prolonged period in culture.

Induced pluripotent stem cells
(iPSCs). Pluripotent cells derived from differentiated somatic cells through treatment with exogenous factors.

Disease phenotype
A molecular, cellular or functional manifestation of a disease in patient-derived cells.

Pluripotent stem cells
(PSCs). Undifferentiated cells that have the ability to self-renew and the potential to differentiate into cells of the three primary germ layers: endoderm, mesoderm or ectoderm.

Embryoid bodies
Aggregates of cells derived from pluripotent cells, formed by growing pluripotent cells in suspension in the absence of self-renewal-promoting factors. Following their aggregation, these cells differentiate into various differentiated cell types partly recapitulating early embryonic development.

Reprogramming
The process by which a differentiated somatic cell acquires the features of a pluripotent stem cell or a differentiated cell of a different cell type.

© 2011 Macmillan Publishers Limited. All rights reserved

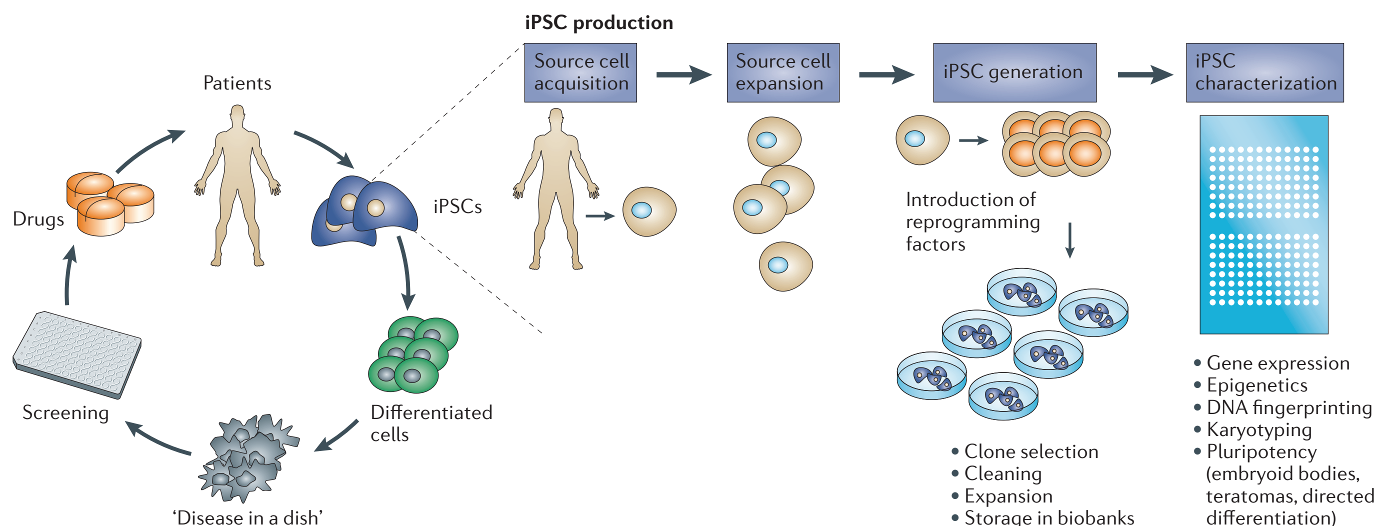

Figure 1 | **An integrated model for drug discovery and development based on the iPSC technology platform.** In this model, the drug discovery process starts with the patient samples used for the derivation of induced pluripotent stem cells (iPSCs), followed by directed differentiation of these cells into cells that have a crucial role in the disease. The hallmark of the technology that makes it valuable for drug discovery is the ability to recapitulate crucial aspects of the disease and create a 'disease in a dish' model for drug screening. A schematic diagram of the iPSC production process is shown. Starting with source cell acquisition from reliable patient sources with informed consent, fibroblasts are expanded and iPSCs are derived; the process is followed by a thorough characterization of the iPSCs, and their expansion and storage in a biobank.

vectors[43–45], excisable vectors[46–48], direct protein transduction[49–51], RNA-based Sendai viruses[52–54], mRNA-based transcription factor delivery[55,56] and microRNA transfections[57], have been reported to solve some of the concerns related to viral integration (TABLE 2). However, many of these approaches either suffer from poor efficiency or they are costly and time-consuming. Chemical compounds can increase reprogramming efficiency and replace certain reprogramming factors[58,59]. In addition, iPSCs have recently been generated by viral delivery of a single reprogramming factor, octamer binding protein 4, in combination with bone morphogenetic proteins[60] or a cocktail of small-molecule compounds[61]. Complete chemical reprogramming has not yet been proven to be successful but remains a much sought-after goal. Despite considerable efforts aimed at optimizing the reprogramming methodology, current approaches remain cumbersome, favouring a highly disciplined and industrialized approach for deriving large panels of iPSC lines from multiple sources.

Implementation of iPSCs in drug discovery

Implementing iPSC technology as a platform for drug discovery requires the seamless integration of multiple steps. First, recruitment of a patient cohort, along with appropriate healthy controls; second, derivation of high-quality, thoroughly characterized iPSCs by a scalable process, and storage of these iPSCs in a well-annotated biobank that allows the tracking and retrieval of samples and accompanying medical information from a relational database; third, differentiation of patient-derived iPSCs into the key cell types that are affected in disease; fourth, discovery of a disease phenotype; and fifth, configuration of an assay (based on the disease phenotype) that is robust, scalable and amenable to automation in a medium- to high-throughput manner for screening. The ability to grow billions of high-quality cells in a reproducible manner is essential.

Patient recruitment. Diseases must be amenable to modelling *in vitro*, and this is not feasible for all diseases. Ideally, the disease must manifest as a disorder at the cellular level, with a relevant phenotype demonstrable through cell culture. Diseases that are associated with insufficient production of known proteins are particularly suitable for *in vitro* modelling, particularly if the levels of such proteins can be detected by immunofluorescence or immunohistochemistry. Examples of such diseases include spinal muscular atrophy (SMA)[20], familial dysautonomia[13] and muscular dystrophy[10]. Moreover, iPSCs derived from the most aggressive or early-onset forms of a genetic disease are likely to reveal a phenotype most quickly and robustly. Recruiting patients for skin punch biopsy or peripheral blood samples can be a challenging and time-consuming process, requiring coordination with patient advocacy groups, academic clinicians and clinics that treat such patients. Large academic institutions with affiliated hospitals, which have a stem cell core and an institutional review board (IRB) process in place for collecting patient samples, have a strategic advantage in collecting tissues and producing iPSCs.

In many cases, heritable diseases or familial forms of more common diseases are rare, and only a handful of specialty clinics throughout the world have access to such patients. Online websites and social networks are

Spinal muscular atrophy
A monogenic neuro-developmental disorder in which a reduced level of survival of motor neuron (SMN) protein leads to the degeneration of motor neurons during childhood.

© 2011 Macmillan Publishers Limited. All rights reserved

Table 1 | **Diseases in which either ESCs or iPSCs have been derived from embryos or patients**

Disease	Molecular defect	Phenotype demonstrated	Refs
Human ESCs			
Alport syndrome	Mutation in *COL4A5*	Not determined	124
Androgen insensitivity syndrome	Deletion of androgen receptor gene	Not determined	124
Fabry syndrome	Mutation in *GLA*	Not determined	125
Fanconi anaemia (carrier)	Mutations in *FANCA*	Not determined	126
Marfan syndrome	Mutation in *FBN1*	Not determined	126
Multiple endocrine neoplasia type 2A	Mutation in *RET*	Not determined	125
Myotonic dystrophy	Trinucleotide expansion in *DMPK* or tetranucleotide expansion in *CNBP*	Decreased expression of two members of the SLITRK family; altered neurite outgrowth, neuritogenesis and synaptogenesis in motor neuron and muscle cell co-cultures	124, 126–128
Neurofibromatosis type 1	Point mutation in *NF1*	Not determined	126
Saethre–Chotzen syndrome	Mutation in *TWIST*	Not determined	124
Spinocerebellar ataxia type 2	Trinucleotide expansion in *ATXN2*	Not determined	125
X-linked myotubular myopathy	Mutation in *MTM1*	Not determined	125
Human ESCs and iPSCs			
Becker muscular dystrophy	Mutation in dystrophin gene	Not determined	10,126
Cystic fibrosis	Mutations in *CFTR*	Not determined	125,127, 129,130
Duchenne muscular dystrophy	Mutation in dystrophin gene	Loss of dystrophin expression in muscle tissue derived from diseased iPSCs; restored by human artificial chromosome-mediated dystrophin expression	124,126, 131
Fragile X syndrome	Trinucleotide (CGG) expansion, silencing of *FMR1*	Not determined	12,125, 126
Gaucher's disease	Point mutation in β-glucocerebrosidase	Not determined	124
Huntington's disease	Trinucleotide expansion in huntingtin gene	Enhanced caspase activity following growth factor withdrawal in iPSC-derived neurons from patients	10, 125–127, 132,133
X-linked adrenoleukodystrophy	Mutation in *ABCD1*	VLCFA levels increased in iPSC-derived oligodendrocytes; reduced after treatment with lovastatin or 4-phenylbutyrate	126,134
iPSCs			
ADA–SCID	Mutations in *ADA*	Not determined	10
Atypical Werner syndrome	Mutation in *LMNA*	Nuclear membrane abnormalities, increased senescence and susceptibility to apoptosis observed in iPSC-derived fibroblasts	135
β-thalassaemia	Deletion in β-globin gene	Not determined	91
Crigler–Najjar syndrome	Mutation in *UGT1A1*	Not determined	136
Type 1 diabetes	Multifactorial; unknown	Not determined	10,57
Down syndrome	Trisomy 21	Not determined	10
Dyskeratosis congenita	Mutations in *DKC1*, *TERT* or *TCAB1*	Progressive telomere shortening and loss of self-renewal of iPSCs	137,138
Dystrophic epidermolysis bullosa	Mutations in *COL7A1*	Lack of expression of type VII collagen, restored following gene correction; no difference between diseased and control formation of three-dimensional skin equivalents	139,140
Familial amyotrophic lateral sclerosis	Mutation in *SOD1* or *VAPB*	Reduced levels of VAPB in fibroblasts, iPSCs and motor neurons derived from patients with VAPB mutation	35,141
Familial dysautonomia	Mutation in *IKBKAP*	Decreased expression of genes involved in neurogenesis and neuronal differentiation; defects in neural crest migration	13,35
Familial hypercholesterolaemia	Mutation in gene encoding LDL receptor	iPSC-derived hepatocytes have an impaired ability to incorporate LDL	19

© 2011 Macmillan Publishers Limited. All rights reserved

Table 1 (cont.) | **Diseases in which either ESCs or iPSCs have been derived from embryos or patients**

Disease	Molecular defect	Phenotype demonstrated	Refs
iPSCs (cont.)			
Glycogen storage disease type 1A	Deficiency in glucose-6-phosphate	Hyperaccumulation of glycogen	19,136
Gyrate atrophy	Mutation in *OAT*	Not determined	142
Hereditary tyrosinaemia type 1	Mutation in fumarylacetoacetate hydrolase	Not determined	136
Hutchinson–Gilford progeria syndrome	Mutations in *LMNA*	Accelerated cell senescence, progerin accumulation, DNA damage, nuclear abnormalities, inclusions in VSMCs; phenotype corrected by HDAd-based gene repair	17,135, 143,144
Inherited dilated cardiomyopathy	Mutation in *LMNA* causing *LMNA* haploinsufficiency	Nuclear membrane abnormalities, increased senescence and susceptibility to apoptosis in iPSC-derived fibroblasts	135
Lesch–Nyhan syndrome (carrier)	Heterozygosity of *HPRT1*	Not determined	10,145
Long QT syndrome	Mutation in genes encoding KCNQ1 or KCNH2	Arrhythmogenicity in cardiac cells; treatment with ranolazine rescues arrhythmia	15,97
MPS type I (Hurler syndrome)	*IDUA* deficiency	Not determined	146
MPS type IIIB	Defective α-*N*-acetylglucosaminidase	iPSCs and differentiated neurons derived from patients show defects in storage vesicles and Golgi apparatus	147
Parkinson's disease	Unknown or mutations in *LRRK2* or *PINK1*	Impaired mitochondrial function in PINK1-mutated dopaminergic neurons, corrected by lentiviral expression of PINK1; sensitivity to oxidative stress in LRRK2-mutant neurons	10,87, 98,107, 148,149
Polycythaemia vera	Heterozygous point mutation in *JAK2*	Enhanced erythropoiesis	91
Progressive familial hereditary cholestasis	Unknown	Not determined	136
Retinitis pigmentosa	Mutations in *RP1*, *RP9*, *PRPH2* or *RHO*	Decreased numbers of differentiated rod cells and expression of cellular stress markers	150
Rett syndrome	Mutation in *MECP2*	Decreased synapse number, reduced number of spines and elevated LINE1 retrotransposon mobility	15,16, 103
Schizophrenia	Unknown	iPSC-derived neurons from patients show diminished neuronal connectivity and decreased neurite number, PSD95 and glutamate receptor expression; neuronal connectivity is improved following treatment with loxapine	151
Scleroderma	Unknown	Not determined	130
Shwachman–Bodian–Diamond syndrome	Mutation in *SBDS*	Not determined	10
Sickle cell anaemia	Mutation in *HBB*	Not determined	130
Spinal muscular atrophy	Mutation in *SMN1*	Reduced SMN levels in iPSCs, reduced size and number of motor neurons; valproic acid and tobramycin increases the number of SMN-rich structures (gems) in iPSCs derived from patients	20
Wilson's disease	Mutations in *ATP7B*	Mislocalization of mutated ATP7B and defective copper transport in iPSC-derived hepatocyte-like cells; rescued by lentiviral gene correction or treatment with the chaperone drug curcumin	152
X-linked chronic granulomatous disease	*CYBB* deficiency	Lack of ROS production in neutrophils, corrected by insertion of *CYBB* minigene	153
α1-antitrypsin deficiency	Mutation in α1-antitrypsin	α1-antitrypsin polymerization	19,130

ABCD1, ATP-binding cassette subfamily D member 1; ADA, adenosine deaminase; ATP7B, copper-transporting ATPase β polypeptide; ATXN2, ataxin 2; CFTR, cystic fibrosis transmembrane conductance regulator; CNBP, cellular nucleic acid binding protein; COL4A5, collagen type IV α-5 chain; CYBB, cytochrome b-245 β polypeptide; DKC1, dyskeratosis congenita 1; DMPK, dystrophia myotonica protein kinase; ESC, embryonic stem cell; FANCA, Fanconi anaemia group A; FBN1, fibrillin 1; FMR1, fragile X mental retardation 1; GLA, galactosidase α; HBB, haemoglobin subunit β; HDAd; helper-dependent adenovirus; HPRT1, hypoxanthine phosphoribosyltransferase 1; IDUA, α-L-iduronidase; IKBKAP, IκB kinase complex-associated protein; iPSC, induced pluripotent stem cell; JAK2, Janus kinase 2; KCNH2, potassium voltage-gated channel subfamily H member 2; KCNQ1, potassium voltage-gated channel subfamily KQT member 1; LDL, low-density lipoprotein; LINE1, long interspersed nucleotide element 1; LMNA, lamin A; LRRK2, leucine-rich repeat kinase 2; MECP2, methyl-CpG-binding protein 2; MPS, mucopolysaccharidosis; MTM1, myotubularin 1; NF1, neurofibromin 1; OAT, ornithine-δ-aminotransferase; PINK1, PTEN-induced putative kinase 1; PRPH2, peripherin 2; PSD95, postsynaptic density protein 95; ROS, reactive oxygen species; RP1, retinitis pigmentosa 1 protein; SBDS, Shwachman–Bodian–Diamond syndrome protein; SCID, severe combined immunodeficiency; SLITRK, SLIT and neurotrophic tyrosine kinase receptor (NTRK)-like protein; SMN1, survival of motor neuron protein 1; SOD1, superoxide dismutase 1; TCAB1, telomerase Cajal body protein 1; TERT, telomerase reverse transcriptase; UGT1A1, UDP glucuronosyltransferase 1 family polypeptide A1; VAPB, vesicle-associated membrane protein-associated protein B; VLCFA, very long chain fatty acid; VSMC, vascular smooth muscle cell.

© 2011 Macmillan Publishers Limited. All rights reserved

increasingly being used as a platform to meet and provide advice and support for patients with the rarest diseases. Although most patients are eager to contribute to research, strict IRB protocols need to be established and followed before patient recruitment, and proper consent must be obtained for the use of iPSCs for drug discovery and commercialization[62]. The Guidelines for Clinical Translation of Stem Cells from the International Society for Stem Cell Research provide templates of consent documents that encompass all the required elements.

Derivation and expansion of patient iPSCs. Dermal fibroblasts are currently the starting material of choice for reprogramming owing to their accessibility, easy storage and handling, and efficient yield of iPSCs. From a single 2–3 mm skin punch biopsy performed under local anaesthesia without a suture, fibroblasts can routinely be expanded and stored for reprogramming. Techniques for reprogramming fresh peripheral blood[36,37,53,63,64] and EBV-immortalized B cell[38,39] samples have been published by several groups, but derivation of iPSCs from cryo-preserved blood samples is not yet widely practised, and the efficiency and yield of this technique is unknown. The choice of source cells may also depend on the particular differentiated cell type that most accurately reflects the disease phenotype: some mouse and human iPSCs have been shown to bear 'epigenetic memory' of their somatic tissue of origin, favouring their differentiation along lineages related to the donor cell while restricting alternative cell fates[65–69].

Skilful identification of properly reprogrammed clones requires familiarity with the morphology of PSCs and proper experience in manual passaging of the cells. Although iPSC lines are heterogeneous with respect to their growth properties and differentiation ability[70], it is currently not known whether this heterogeneity reflects differences among starting cell populations (inter-patient variability) or inherent variability among different iPSC clones derived from the same starting cell population (intra-patient variability). Optimized reprogramming strategies that use small molecules, modified culture conditions[71] or extended passaging may produce more uniform and stable behaviour in iPSC lines. In addition, identification of molecular markers that are predictive of iPSC differentiation potential will substantially accelerate the selection of high-quality cell lines. By correlating genome-wide expression data with differentiation capacities of 32 human ESC and iPSC lines, Meissner *et al.*[72] have recently generated an assay for quickly comparing and characterizing PSCs. This assay yields a scorecard that can be used to predict the differentiation efficiency of individual cell lines. It should be noted, however, that even though iPSC lines may appear to have distinct differentiation potentials under particular conditions, all cell lines may differentiate with similar efficiencies when an optimal differentiation condition is applied[73,74]. The expression of genes in the delta homolog 1 (*Dlk1*)–type III iodothyronine deiodinase (*Dio3*) locus was found to correlate with high pluripotent differentiation potential in mouse iPSCs[75,76], but its relevance in human iPSCs remains unexplored.

Comprehensive genomic analyses of iPSCs have revealed a striking number of chromosomal aberrations and point mutations that may pre-exist in the donor cell, occur during reprogramming or be acquired through prolonged cell culture[77–81]. As for ESCs, these aberrations may confer a growth advantage to iPSCs such that the mutated cells eventually dominate the population[25], altering the differentiation capacity of the iPSCs and interfering with the disease phenotype[82,83]. To minimize the risk that acquired genetic (or epigenetic) aberrancies might alter a disease phenotype or confound drug discovery, several iPSC clones from the same patient should be assessed for consistent functional readouts. As long-term cultures result in a greater accumulation of chromosomal aberrations in ESCs[25], early passages of iPSCs may be more suited for disease modelling and drug screening. However, Hussein *et al.*[81] have recently shown that the iPSC genome is dynamic in culture; the reprogramming process initially generates a variety of copy number variations that appear to be selected against during prolonged culture, leading to the generation of iPSCs with a more normal genetic complement. Therefore, careful analysis of the genomic integrity of iPSCs is essential for both early and late passage lines.

Given these limitations, the use of either human ESCs or iPSCs for screening presents particular challenges that are rarely encountered with other cell types. The growth and expansion of iPSCs is technically more demanding and requires greater expertise than most commonly used cell types. Automation of PSC culture is being developed, but the susceptibility of some of the iPSC and human ESC lines to spontaneous differentiation necessitates continuous monitoring and sometimes painstaking manual purging of differentiated cells. Recent advances in media formulation, the use of growth substrates and the introduction of RHO-associated protein kinase inhibitors to human ESC and iPSC cultures have facilitated the maintenance and propagation of PSCs through multiple passages[84]. The establishment of high-quality cell lines and rigorous cell characterization, which should include testing for batch-to-batch variability, are essential for the success of large-scale screening.

Directed differentiation. Differentiation of human PSCs is a multistep process that can take several weeks or months for highly specialized target cell types. Such an extremely long preparation time adds to the complexity of screening operations and contributes to the inconsistency of results. Scale-up methods are prerequisites for achieving robust screening with iPSC-derived cellular models; these methods include the expansion of cells at intermediate stages of differentiation, such as the expansion of embryoid bodies[85] or differentiated yet proliferative progenitor cells, in combination with a reduction in the number of manipulations during differentiation and the use of chemically defined media. The development of large-scale cell banks of iPSCs, or ideally of differentiated products such as neural progenitors, can further reduce cell preparation time and complexity, and provide consistent cell preparations for multiple screening campaigns. At present, the implementation of such

Epigenetic
A heritable change in gene expression that is not caused by the DNA sequence.

Copy number variations
Duplications or deletions in the genome that lead to variability in the number of genes.

© 2011 Macmillan Publishers Limited. All rights reserved

Table 2 | **Summary of iPSC derivation methods***

Cell types	Reprogramming factors and methods			
	Integrating	**Excisable**	**Non-integrating**	**Chemicals**
Fibroblasts	• Retroviral KSOM or KSO[7,154] • Lentiviral KSOM • Inducible lentiviral KSOM[10] • Lentiviral OSNL[8] • Lentiviral mir-302 cluster[155]	• Floxed lentiviral KSOM or KSO[87] • Transposon KSOM[46]	• Adenoviral[156] • Plasmids[44,157] • Protein[49] • mRNA[56] • RNA virus[52] • miRNA mimics[158]	• KSO + valproic acid[110] • KSOM + vitamin C[159] • KSOM + SB431542, PD0325901 and thiazovivin[160] • KSOM + sodium butyrate (NaB)[161]
Bone marrow (mobilized) or peripheral blood cells	• Retroviral KSOM[36,37] • Inducible lentiviral KSOM[36]	Not applicable	• RNA virus[53] • Plasmids[63,64]	Not applicable
Cord blood cells	• Lentiviral OSNL[162] • KSOM[163] • Retroviral OCT4 and SOX2 (REF. 164)	Not applicable	• Plasmids[63,64]	Not applicable
EBV-immortalized blood cells	-	Not applicable	• Plasmids[38,39]	Not applicable
HUVECs	• Retroviral KSOM[165]	Not applicable	Not applicable	• OCT4 + PS48, NaB, A-83-01 and PD0325901 (REF. 161)
Adipose-derived stem cells	• Lentiviral KSOM[166] • Retroviral KSO[167]	Not applicable	• Plasmid minicircles[157] • miRNA mimics[158]	Not applicable
Keratinocytes	• Retroviral KSOM and KSO[168] • Inducible lentiviral KSOM cassette[169]	Not applicable	Not applicable	• OCT4 and KLF4 + tranylcypromine and CHIR99021 (REF. 170) • OCT4 + PS48, NaB, A-83-01 and PD0325901 (REF. 61)
Neural stem cells	• Retroviral OCT4 (REF. 171)	Not applicable	Not applicable	Not applicable
Astrocytes	• Retroviral KSOM[172]	Not applicable	Not applicable	Not applicable
Hepatocytes	• Retroviral KSOM[173]	Not applicable	Not applicable	Not applicable
Amniotic cells	• Retroviral KSOM[174] • Lentiviral OSN[175]	Not applicable	Not applicable	• OCT4 + PS48, NaB, A-83-01 and PD0325901 (REF. 61)

A-83-01, 3-(6-Methyl-2-pyridinyl)-*N*-phenyl-4-(4-quinolinyl)-1*H*-pyrazole-1-carbothioamide; CHIR99021, 6-((2-((4-(2,4-dichlorophenyl)-5-(4-methyl-1*H*-imidazol-2-yl)pyrimidin-2-yl)amino)ethyl)amino)nicotinonitrile; EBV, Epstein–Barr virus; HUVEC, human umbilical vein endothelial cell; iPSC, induced pluripotent stem cell; KLF4, Krüppel-like factor 4; KSO, culture medium containing the transcription factors KLF4, SOX2 and OCT4; KSOM, culture medium containing the transcription factors KLF4, SOX2, OCT4 and MYC; miRNA, microRNA; OCT4, octamer binding protein 4; OSN, a lentivirus containing transcription factors OCT4, SOX2 and Nanog; OSNL, a lentivirus containing transcription factors OCT4, SOX2, Nanog and LIN28; PD0325901, *N*-[(2*R*)-2,3-dihydroxypropoxy]-3,4-difluoro-2-[(2-fluoro-4-iodophenyl)amino]-benzamide; PS48, (2*Z*)-5-(4-chlorophenyl)-3-phenyl-2-pentenoic acid; SB431542, 4-[4-(1,3-benzodioxol-5-yl)-5-(2-pyridinyl)-1*H*-imidazol-2-yl]benzamide. *For simplicity, only the first reports of a given method on a given cell type are listed.

approaches for drug screening has been reported only for mouse ESCs and iPSCs[23]. The adaptation and extension of these approaches to human iPSCs will aid in the standardization of their routine use in drug screening.

In theory, human iPSCs can give rise to all cell types of an adult organism; in practice, however, *in vitro* differentiation protocols have been developed for only a small subset of specific cell types: chiefly neurons, haematopoietic progenitors, hepatocytes, cardiomyocytes and keratinocytes, among others[20,22,35,86–92]. In many cases the differentiation process is inefficient and produces only heterogeneous cell populations (consisting of cells at different stages of maturation) and/or overtly mixed cell populations. The establishment of efficient protocols for directed differentiation has been hampered by our poor understanding of human embryonic development and the long timelines associated with the differentiation of human PSCs.

Existing methods for the differentiation of mouse ESCs or iPSCs are not always applicable to human cells, and the differentiation of human PSCs can take much longer (up to several weeks or months) owing to the longer time required for human gestation and development. Furthermore, differentiation typically requires expensive recombinant growth factors and media, and is therefore costly. Small molecules that mimic growth factors and directly activate developmental pathways reduce costs, but such small molecules do not exist for all growth conditions. It may be possible to drive iPSCs towards a particular lineage or force progenitors to mature more quickly by overexpressing lineage-specific transcription factors during differentiation[93–95]. Although shortened differentiation protocols would be beneficial for high-throughput screening, ectopic expression of differentiation factors may inadvertently mask disease phenotypes, especially in developmental diseases.

© 2011 Macmillan Publishers Limited. All rights reserved

Fluorescence-activated cell sorting or magnetic bead purification can be applied to enrich for pure cultures of target cells, but in many cases appropriate surface antigens and corresponding antibodies have not yet been defined. Regardless, such steps are cumbersome and impractical for use in high-throughput screening. Alternatively, one can selectively measure effects of compounds on predefined subpopulations of cells through image-based high-content assays. However, such image-based assays may only be applicable to a limited number of cell types for which differentiation and/or maturation markers have been well established. For some cell types, such as neurons, it may be beneficial to use a complex mixed culture or co-culture with astrocytes, thereby creating a more physiologically relevant screening platform. The single greatest barrier to the widespread application of iPSCs in drug development remains the difficulty of achieving faithfully directed differentiation of disease-relevant cell types on a large scale.

Disease modelling. The key to modelling any disease with iPSCs is the identification of a disease-relevant cellular pathology. The most compelling demonstrations to date have pertained to early-onset diseases that have a strong genetic component and affect a highly defined cell or tissue type[13–16,20,91,96]. For diseases with a known molecular mechanism, such as mRNA mis-splicing in SMA or familial dysautonomia[13,20], small-molecule screening campaigns using these human models are bound to identify efficacious compounds because the assay is directly relevant to the pathological mechanisms of the disease. Although theoretically appealing, modelling diseases that are more genetically complex — such as sporadic Alzheimer's disease or Parkinson's disease — represents a daunting challenge that requires the identification of celular phenotypes that relate to known aspects of disease pathology. In these cases the identification of pathological phenotypes that are common to both sporadic and genetically inherited forms of the disease may illuminate common mechanisms of the disease being modelled.

Robust disease phenotypes have been demonstrated in several monogenic diseases, including SMA[20], fragile X syndrome[12], familial dysautonomia[13], LEOPARD syndrome[14], Rett syndrome[15,16], Hutchinson–Gilford progeria syndrome (HGPS)[17], Long QT syndrome (LQTS)[18,97], various models of human liver disease (such as α1-antitrypsin deficiency, familial hypercholesterolaemia and glycogen storage disease type 1A)[19], as well as familial Parkinson's disease[98,85]. The current list of disease phenotypes (TABLE 1) will rapidly become obsolete, and thus we only highlight examples in which patient-derived cells have been used to recapitulate *in vitro* models of various diseases.

Future collaborations between academic groups and industry may provide the quickest route for the development of novel therapeutics for patients. Leading academic laboratories that have created these disease models could work closely with industry partners that have access to large chemical libraries, high-throughput screening capabilities and the expertise in clinical trials to bring novel therapeutics to the clinical setting.

LEOPARD syndrome
An autosomal dominant multisystem disease caused by a mutation in the gene encoding protein tyrosine phosphatase non-receptor type 11. The disease affects the skin as well as the skeletal and cardiovascular systems.

SMA. SMA is an autosomal recessive childhood neuromuscular disease that is characterized by the loss of lower motor neurons, and is the most common cause of death by a heritable disease in infants. There are currently no effective treatments available. SMA is caused by a decrease in levels of the survival of motor neuron (SMN) protein due to mutations in the *SMN1* gene[99]. Although the SMN protein is expressed ubiquitously, its loss leads to motor neuron degeneration, consistent with its specific role in these cells[100]. In the past, researchers have screened for compounds that elevate SMN levels in various engineered cell lines and patient fibroblasts, but these compounds have failed in the clinic — possibly because the mechanisms that regulate SMN protein levels in fibroblasts or engineered cells are substantially different from those in human motor neurons *in vivo*. Ebert and colleagues[20] created two iPSC lines — one from a patient with SMA and the other from an unaffected relative — and differentiated them into motor neurons. In these cultures motor neuron numbers at later time points were shown to be reduced, specifically in the cells derived from patients with SMA, demonstrating for the first time that the process of reprogramming and directed differentiation faithfully captured and recapitulated the disease phenotype. Two compounds, valproic acid and tobramycin, increased the number of SMN-rich structures (called gems) in the patient-derived iPSCs. Although it remains to be seen whether the compounds that elevate SMN levels in patient-derived iPSCs can have the same effect in motor neurons and thus rescue motor neuron loss in patients, these early experiments provide an important validation that neurons derived from iPSCs of patients with SMA can be used to model the disease.

Familial dysautonomia. Familial dysautonomia (also known as Riley–Day syndrome), a disorder of the autonomic nervous system that results in abnormal migration of neural crest cells, affects the survival of sensory, sympathetic and parasympathetic neurons. There are currently no effective treatments or appropriate disease models for familial dysautonomia. Most cases of familial dysautonomia are caused by mutations in the gene encoding IκB kinase complex-associated protein (IKBKAP) that result in the skipping of exon 20 (REF. 101). Lee, Studer and colleagues[13] derived iPSCs from three young patients with familial dysautonomia and induced the differentiation of these iPSCs into neural crest cells. Three defects were found in these neural crest cells: a splicing defect in *IKBKAP*, a cellular migration defect and a defect in neurogenesis. By testing several drugs known to affect the splicing of *IKBKAP* in iPSC-derived neural crest cells of patients with familial dysautonomia, they documented that kinetin markedly reduced the splicing defect and modestly affected neurogenesis, but had no effect on cell migration. Expression profiling of diseased neural crest cells identified the differential expression of a series of genes that are important for neurogenesis and cell survival. This study confirmed the utility of iPSC-based cellular models in phenotypic screens for drug discovery, and for the identification of novel molecular targets and disease mechanisms.

© 2011 Macmillan Publishers Limited. All rights reserved

Rett syndrome. A study by Marchetto *et al.*[15] focused on another developmental neurological disease, Rett syndrome, which is part of the larger group of autism spectrum disorders[12]. Rett syndrome is caused by mutations in methyl-CpG-binding protein 2, a protein involved in DNA methylation that regulates an array of different genes[102]. iPSCs derived from healthy controls and patients with Rett syndrome were differentiated into glutamatergic and GABA (γ-aminobutyric acid)-ergic neurons, and the authors assessed neurogenesis, synapse number and neuronal morphology. Although the authors observed no changes in neurogenesis, they were able to measure a substantial reduction in synapse number as well as a reduction in the number of spines — the small protrusions in neuronal processes where glutamatergic synapses are formed. A reduced number of spines has previously been observed in the post-mortem brains of patients with Rett syndrome. Probing the mechanisms of abnormal synaptogenesis using electrophysiology and calcium imaging showed that calcium oscillations and the frequency of spontaneous postsynaptic currents were decreased in the neurons of patients with Rett syndrome. In another elegant study, Muotri *et al.*[16] showed that iPSC-derived neural cells from patients with Rett syndrome manifest increased L1 transposon motility, providing yet another mechanism that may underlie Rett syndrome and become a target for therapy[15,16,103].

HGPS. HGPS is a rare congenital disease caused by a mutation in the lamin A (*LMNA*) gene that leads to a truncated and farnesylated form of LMNA called progerin[104]. Patients carrying autosomal dominant mutations in *LMNA* show signs of early ageing and often die in their early teens as a result of myocardial infarction or stroke. Many tissues are ravaged by this disease, particularly those of mesenchymal lineage, including vascular smooth muscle cells (VSMCs). To gain an insight into disease biology, Zhang *et al.*[17] derived iPSCs from two patients with HGPS carrying different mutations in *LMNA* (iPSCs derived from their parents were used as controls). The authors then differentiated these cells into five lineages: fibroblasts, endothelial cells, neural progenitor cells, VSMCs and mesenchymal stem cells. They found three major pathological defects in cells differentiated from the two patients: DNA damage, mislocalization of lamina-associated polypeptide 2 (also known as TMPO) and nuclear dysmorphology. Interestingly, the most severe defects were in the VSMCs, which express the highest levels of progerin. Unexpectedly, calponin 1 inclusion bodies were found in patient-derived VSMCs, suggesting that patient-derived cells have defects in handling protein load. Calponin 1 is an actin-binding protein that is important for contractility, and its aggregation may lead to poor contractile properties of VSMCs. Consistent with this, patient-derived VSMCs were considerably more susceptible to external stressors such as hypoxia and electrical stimulation, which perhaps explains why these cells are pivotal in the manifestation of the disease. In addition, patient-derived mesenchymal stem cells were more vulnerable to serum starvation, indicating a selective vulnerability of these stem cells to stress conditions. When patient-derived iPSCs were differentiated into several lineages, VSMCs emerged as the most vulnerable cell type associated with the various disease pathologies.

LQTS. Congenital LQTS is an inherited condition that predisposes patients to life-threatening cardiac arrhythmias characterized by delayed repolarization of the cardiomyocyte action potential and a prolonged QT interval in electrocardiograms. Some drugs exacerbate the repolarization defect in individuals with LQTS, inducing drug-associated adverse events such as sudden cardiac death; these untoward but rare side effects have led to the withdrawal of medications from the market or severe restrictions on their use[105,106]. A lack of human cardiomyocytes from patients with the genetic mutations associated with LQTS has hindered attempts to develop protective drugs for this condition, as well as attempts to screen preclinical drug candidates to eliminate those drugs that promote arrhythmia.

Two groups have recently derived iPSCs from patients with LQTS, differentiated them into cardiomyocytes and documented phenotypes that are indicative of LQTS. Moretti *et al.*[18] generated iPSCs from members of a family affected by type 1 LQTS, carrying a mutation in the gene encoding potassium voltage-gated channel subfamily KQT member 1. Cardiomyocytes derived from these iPSCs exhibited prolonged action potentials and defective potassium channel properties. Consistent with the known complications of LQTS, the cardiomyocytes were susceptible to catecholamine-induced tachyarrhythmia, which can be attenuated with β-adrenergic receptor blockers. Itzhaki *et al.*[97] derived iPSCs from an individual with type 2 LQTS, carrying a mutation in the gene encoding potassium voltage-gated channel subfamily H member 2. The patient-derived iPSCs were differentiated into beating cardiomyocytes, which likewise showed the characteristic prolongation of action-potential duration and arrhythmogenicity. This cell-based model was validated by testing the effect of several ion channel blockers as anti-arrhythmic agents. As drug-induced exacerbation of LQTS is one of the most common forms of drug toxicity, these studies reinforce the concept that genetically diverse iPSC-derived cardiomyocytes can be exploited to assess adverse drug effects.

Parkinson's disease. Parkinson's disease, which is caused by the progressive loss of midbrain dopaminergic neurons, is the second most common neurodegenerative disorder. There is no known cure for Parkinson's disease, and neurodegeneration progresses despite initial symptomatic control, inevitably leading to the worsening of symptoms and a loss of therapeutic efficacy. Recently, patient-derived neurons from familial forms of Parkinson's disease have been used to model aspects of the disease *in vitro*[98,107]. In one study Seibler *et al.*[107] derived dopaminergic neurons from patients with mutations in the gene encoding PTEN-induced putative kinase 1, an outer mitochondrial membrane protein that is believed to regulate the mitochondrial translocation of the E3 ubiquitin protein ligase parkin, another protein

QT interval
A measure of the time between the start of the Q wave and the end of the T wave in the electrical cycle of the heart.

© 2011 Macmillan Publishers Limited. All rights reserved

associated with familial Parkinson's disease[85]. Patient-derived dopaminergic neurons exhibited impaired recruitment of parkin to mitochondria, increased mitochondria copy number and increased expression of the mitochondrial regulator peroxisome proliferator-activated receptor-γ co-activator 1α. These phenotypes were rescued by exogenous expression of wild-type PTEN-induced putative kinase 1.

In another study Nguyen and colleagues[98] derived iPSCs from a patient with a mutation in the gene encoding leucine-rich repeat kinase 2; mutations in this gene are the most common cause of familial Parkinson's disease[77]. Patient-derived dopaminergic neurons showed increased vulnerability to stress by hydrogen peroxide, 6-hydroxydopamine and the proteasome inhibitor MG-132, consistent with the notion that both genetic and environmental factors contribute to the development of Parkinson's disease. It remains to be seen whether dopaminergic neurons derived from patients with sporadic Parkinson's disease show similar vulnerability to stress agents, and whether the disease phenotypes and mechanisms identified in neurons derived from patients with familial Parkinson's disease extend to those derived from patients with sporadic Parkinson's disease, thus providing common targets for therapeutic intervention.

Coaxing iPSCs out of their embryonic state into reliable populations of mature adult cells represents another challenge for the recapitulation of disease phenotypes. iPSC-derived cells that are differentiated and grown in culture for only a few weeks or months typically have characteristics more akin to early embryonic or fetal cells rather than adult cells. It may therefore be necessary to develop culture conditions and differentiation methods to fully mature such cells before disease phenotypes can be revealed. This may be particularly important when modelling adult-onset diseases such as Parkinson's disease or Alzheimer's disease, in which disease manifestation may occur only in mature neurons with synaptic properties that are typically present in an adult brain. It remains to be seen whether cells derived from patients with adult-onset diseases will faithfully recapitulate disease phenotypes. The experiments by Nguyen and colleagues suggest that at least some aspects of familial Parkinson's disease can be modelled using immature neurons, and it may be possible to further refine culture conditions to reveal differences between healthy and diseased cells.

Drug screening and optimization

iPSC-based disease models can be useful in preclinical studies at virtually all stages of drug development (FIG. 2). In the conventional target-centric drug discovery model, compounds are typically not tested in a relevant patient population until Phase II clinical trials have been carried out, thus contributing to a high drug attrition rate[108]. By contrast, iPSC technology provides the means to obtain drug efficacy and toxicity data in a disease-relevant context at the earliest stages of drug development and indeed throughout the drug discovery process, which ensures lower attrition at the expensive, late stages of the clinical pipeline.

iPierian has recently applied such an integrated iPSC-based drug discovery strategy to target SMA by deriving iPSCs from several patients with type 1 SMA, differentiating these iPSCs into motor neurons, and screening them against a library of 200,000 compounds (FIG. 3). Akin to an earlier report[20], the reduced levels of SMN protein and gems in patient-derived cells were reproduced *in vitro*. In addition, a high-content screening assay was created that quantifies the amount and subcellular localization of this protein in motor neurons — a cell type uniquely sensitive to lower levels of SMN[99]. The screening campaign using lower SMN levels as the assay readout was specifically geared towards the identification of small molecules with the capacity to increase SMN levels to those found in non-diseased motor neurons. The custom algorithms identified neurons that express the motor neuron markers Islet and HB9 (also known as MNX1) in which SMN levels had been elevated by small molecules. As industrial and academic investigators become more accustomed to this paradigm, further developments are expected in this field.

Although pharmaceutical companies have proprietary compound libraries with an excess of one million compounds, academic laboratories are largely restricted to purchasing commercially available compound libraries that contain molecules in the public domain as well as US Food and Drug Administration (FDA)-approved drugs. These chemical libraries can be readily exploited because repurposing approved compounds for a new disease accelerates the delivery of therapeutics to patients. Chemical libraries such as the Johns Hopkins Clinical Compound Collections, MicroSource Spectrum Collection, Prestwick Collection and the Sigma-Aldrich Library of Pharmacologically Active Compounds contain most of the nearly 3,400 FDA-approved drugs currently on the market[109].

Target identification using iPSCs. The most promising drug targets have an unequivocal link to human disease because of phenotypic similarities to genetically engineered mouse strains and, where possible, human genetics. In complex diseases for which molecular mechanisms are unknown, a cellular pathology captured in iPSC-derived cells could be used for target-agnostic drug screening. Such phenotypic screens could enable the discovery of therapies even before developing a deep molecular understanding of the disease process. RNA interference (RNAi) technology has been applied on a genome-wide scale for several biological processes, and potential new targets have been proposed[110,111]. However, even when a target is identified and validated via RNAi knockdown in surrogate cells it remains uncertain whether modulating the target chemically rather than genetically — and in the diseased cells of the patients — will have a therapeutic effect.

Access to physiologically relevant cell types from patient-derived iPSCs in combination with RNAi technologies presents the opportunity to not only identify disease-relevant biological signatures that can shed light on mechanisms of disease but to also discover novel drug targets. For example, systematic RNAi knockdown

© 2011 Macmillan Publishers Limited. All rights reserved

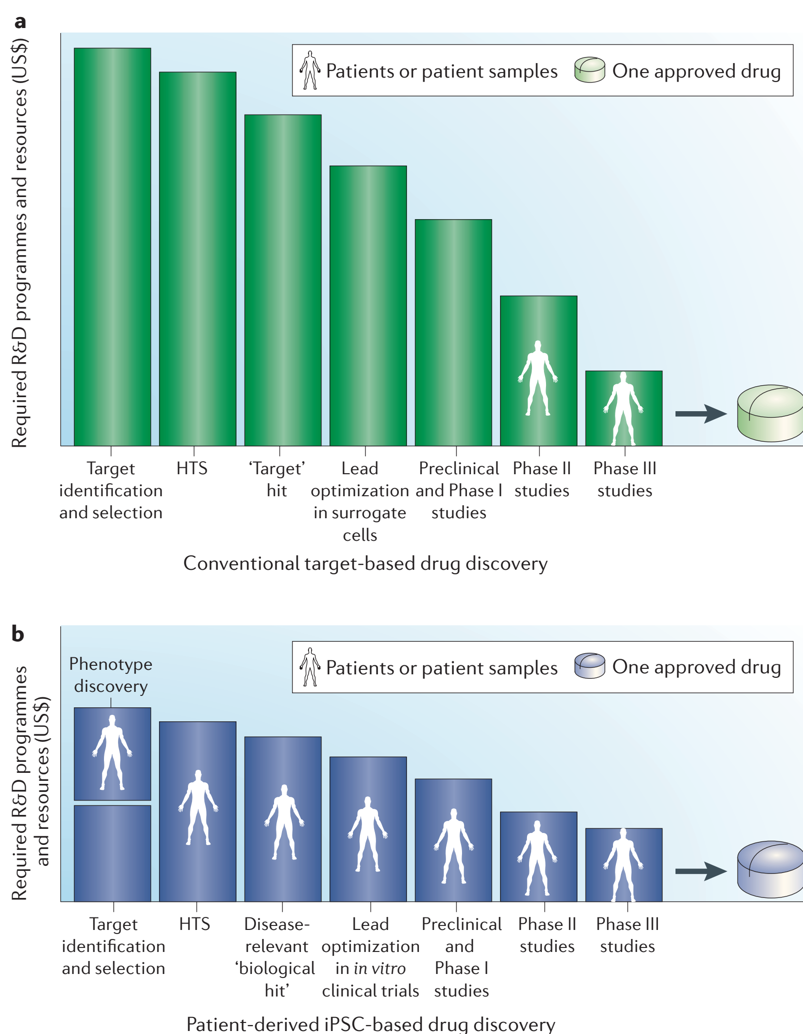

Figure 2 | **Value creation with the iPSC-based drug discovery paradigm: comparison between conventional target-centric drug discovery and patient-derived iPSC-enabled drug discovery.** In the conventional model (part **a**) the high attrition rate of drug development behoves drug companies to invest in a large number of research and development (R&D) programmes without knowledge of the probability of success of such programmes until Phase II clinical trials have been carried out. In the patient-derived induced pluripotent stem cell (iPSC)-based drug discovery model (part **b**), patient samples are introduced into the drug discovery process at the very beginning, complementing current drug discovery technologies. iPSC technology preserves human disease biology in disease-relevant cell types, providing the opportunity to break with surrogate models such as engineered cell lines and mostly unpredictable animal models, and provides the means to obtain efficacy and toxicity results very early and indeed throughout the drug discovery process. In this model fewer programmes are needed, and the programmes have a higher likelihood of succeeding. The overall investment in R&D is therefore lower, as there is likely to be less attrition at later stages of the clinical pipeline. The vertical axes in both graphs represent arbitrary units scaled to reflect the anticipated lower cost of iPSC-based discovery programmes.HTS, high-throughput screening.

of gene candidates identified via expression profiling of healthy and diseased cells can identify potential therapeutic targets in cases in which gene knockdown can rescue a disease phenotype. Compounds that mimic the effect of gene knockdown can then be tested for efficacy in the more disease-relevant iPSC models, and be further chemically optimized. In instances in which targets are cell-surface proteins, therapeutic monoclonal antibodies could be designed and developed. Ultimately, the most effective drug development strategy will combine disease-relevant cell-based screens with target-directed compound optimization or other biologically active molecules through rational drug design. Although RNAi libraries enable the probing of the entire genome, the high costs of maintaining such libraries remain a challenge, and off-target effects need to be eliminated before such data are interpretable.

iPSCs in toxicology. Adverse drug reactions represent a major challenge for pharmaceutical companies, healthcare providers and drug regulators, and are major contributors to the high cost of drug development[112]. The development of predictive human cellular systems that complement current toxicity tests in animals and primary cells is therefore vital. Given that differentiation protocols can generate human cells that are most susceptible to drug toxicity[6], patient-derived stem cell-based assays of drug toxicity offer great potential for assessing adverse drug effects. The ability to derive and test cells from individuals with known drug sensitivities — either therapeutic or toxic — would help to identify the molecular basis of variable human drug responses. Moreover, panels of iPSCs offer the unprecedented opportunity to capture the heterogeneity of patient populations resulting from variables such as gender and ethnicity. As outlined above, cardiomyocytes differentiated from iPSCs taken from patients with LQTS have been used to reflect the tendency of drugs to exacerbate cardiac conduction defects, and thus could be used to screen compounds for cardiotoxicity[16,53]. However, this is just the first proof of principle for modelling the multitude of genetic traits that dictate drug sensitivity, metabolism or toxicity. The use of iPSCs to personalize drug development may prove to be a powerful means of reducing drug toxicity, stratifying patient response and reducing late-stage clinical failures.

The '*in vitro* clinical trial'

Most clinical trials fail because of safety issues or a lack of efficacy[108,113,114]. Even compounds that appear to be safe and efficacious in animal models or *in vitro* cell systems can fail spectacularly in clinical trials in patients[115–119]. It is perhaps not surprising that positive results obtained in preclinical drug testing are not replicated in the diverse human population if they are based on imperfect model systems — for example, genetically inbred mice or highly contrived human cell lines. According to estimates from the International Haplotype Mapping project, there are approximately 10 million single nucleotide polymorphisms (point mutations with a frequency of >1%) in the human population, with even rarer mutations existing in certain individuals[120]. Copy number variations, deletions, insertions and inversions add to this genetic variation in the human population, as does marked variation at the epigenetic level in cells and tissues. Taken together, the challenges associated with developing an accurate model system are staggering.

© 2011 Macmillan Publishers Limited. All rights reserved

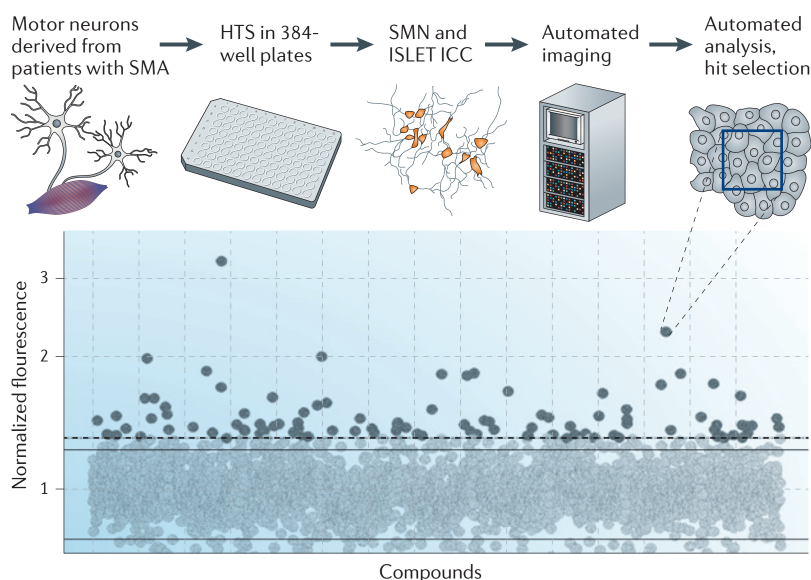

Figure 3 | **Schematic diagram of the iPSC-driven lead discovery platform based on iPierian's SMA programme.** After the identification of a disease phenotype in motor neurons differentiated from induced pluripotent stem cells (iPSCs) that are derived from patients with spinal muscular atrophy (SMA), an image-based high-throughput screening (HTS) assay is executed on patient-derived motor neurons in 384-well plates. The assay is based on endogenous levels of survival of motor neuron (SMN) protein, and the use of non-engineered cells provides the right physiological context for drug candidate identification. Compound-treated cultures are immunostained with antibodies specific to SMN and islet cells, imaged using Molecular Device's ImageExpress Ultra and analysed using algorithms developed in-house to identify compound hits that increase SMN expression levels in islet-positive cells. Dark grey dots represent compound hits that increase SMN expression levels by immunostaining; light grey dots represent compounds that do not alter SMN levels sufficiently to be considered hits. Cells within one well of a 384-well plate are shown (inset), where a 'hit' was identified by automated imaging. ICC, immunocytochemistry.

An exciting idea for capitalizing on the iPSC technology is to test candidate compounds in cell-based assays from a large panel of patient-derived iPSCs before these compounds are tested in the clinic. Such an '*in vitro* clinical trial' might reveal the range of drug responsiveness that is due to the heterogeneity of the human population, and might indicate whether stratified subpopulations respond to a drug, thereby allowing another level of compound triage and refinement in clinical trial design. This additional information might inform the magnitude of effect and thus the size of the patient population (also known as the 'power') needed to achieve meaningful therapeutic end points in clinical testing in patients.

In a traditional drug discovery and development paradigm, there is typically little opportunity to test candidate compounds in patient cells before clinical trials are initiated. However, pharmaceutical companies have recently started to implement small-scale clinical proof-of-concept trials for the purpose of demonstrating that clinical candidate compounds show a reassuring signal of clinical efficacy in patients before launching into a full-fledged clinical trial. Although such studies involve a smaller number of patients than a typical Phase II trial, they remain costly and time-consuming, and the number of patients used to yield conclusive results is too low. iPSC technology has the potential to increase the success of clinical trials because it could enable the testing of candidate compounds *in vitro* for efficacy, toxicity and dose response in the actual diseased cells, and across a large number of patients who are reflective of the desired drug target product profile[121,122].

An added benefit of an *in vitro* clinical trial is that the known issues with patient compliance or risks to patients are eliminated. In addition, the repository of iPSCs and progenitor cell lines created from patient-derived cells is amenable to almost limitless expansion, thus enabling tests of many more candidate compounds than possible in actual clinical trials involving patients. Moreover, cells from a single patient can serve both as control and treatment arms at multiple time points. As expected, the number of patient-derived cell lines that need to be created to properly represent a patient population also depends on the complexity of the disease. For monogenic diseases such as Huntington's disease or SMA, which have a strong and highly consistent phenotype, a smaller number of cell lines would typically be needed than for complex diseases such as type 2 diabetes or Alzheimer's disease.

In addition to representing the 'disease in a dish', iPSCs afford the opportunity to study multiple cell types from the same patient — some for drug effect, others for drug toxicity. From this perspective iPSCs can represent a 'patient in a dish', thereby prospectively reflecting a clinical trial more effectively than by testing just one cell type. However, simultaneous differentiation of multiple iPSC clones and parallel screening of several patient-derived cell lines would be very expensive. The differentiation of iPSC lines remains highly variable, and reducing this variability by optimized differentiation methods will be an important factor in enabling these studies. An alternative strategy would be to carry out a primary screen in a single patient-derived cell line, followed by secondary confirmatory screens in several additional patient-derived cell lines to identify lead compounds that are effective across many patients or certain subpopulations of patients. Patient-stratification strategies have proven to be particularly useful in the field of oncology. There are numerous examples of cancer therapies that are efficacious in only a patient subpopulation with a particular genotype. For example, in one study Engelman and colleagues[123] showed that patients with *MET* gene amplification were resistant to the epidermal growth factor receptor inhibitor gefitinib (Iressa; AstraZeneca/Teva).

The creation of an iPSC biobank for the US population could provide a reservoir of cellular models for many diseases, whereby drugs can be tested against many different genetic backgrounds for efficacy and toxicity. An iPSC biobank that is shared among academic and industry researchers could facilitate the development of therapeutics and regenerative medicine.

Future perspectives

We have outlined the key concerns and the theoretical potential of iPSCs in drug discovery. The crucial experiments performed in monogenic diseases exemplify the

© 2011 Macmillan Publishers Limited. All rights reserved

power of iPSC technology in illuminating human disease at the cellular and molecular level, and highlight the potential to de-risk drug discovery programmes at the preclinical stage. Although the direct translation of findings using the 'disease in a dish' model to predict clinical outcomes in patients has yet to be proven, the unacceptable rate of drug attrition in late phases of drug discovery calls for a new drug discovery paradigm — one in which compelling evidence for drug efficacy is obtained in relevant, patient-derived cell systems at early stages of preclinical development (FIG. 2). Furthermore, the ability to differentiate iPSCs from the same patient into cell types that enable testing for cardiac, hepatic and neuronal toxicity at preclinical stages has the potential to de-risk compounds at an earlier stage of drug development. This approach can practically reduce the cost burden of failed drug development programmes, and enable the enrichment of programmes with the highest probability of success. To meet its therapeutic potential, the iPSC-based drug discovery platform will have to achieve more efficient reprogramming, more faithfully directed differentiation and more robust disease phenotypes, especially for complex and common diseases. The validation of iPSCs as a drug discovery technology may encourage the pharmaceutical industry to embrace a new paradigm, reducing obligatory preclinical testing in non-predictive animal models, to proceed directly from *in vitro* clinical trials to actual clinical trials in patients. Ultimately, the predictive value of the use of our proposed *in vitro* clinical trials in predicting efficacy in patients, as well as identifying responder and non-responder populations, will be revealed gradually as we continue to use the iPSC platform more broadly and deeper within the drug discovery pipeline.

1. Wilhelm, S. *et al.* Discovery and development of sorafenib: a multikinase inhibitor for treating cancer. *Nature Rev. Drug Discov.* **5**, 835–844 (2006).
2. Hazuda, D. J. *et al.* Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. *Science* **287**, 646–650 (2000).
3. Jacoby, E., Bouhelal, R., Gerspacher, M. & Seuwen, K. The 7 TM G-protein-coupled receptor target family. *ChemMedChem* **1**, 761–782 (2006).
4. Pouton, C. W. & Haynes, J. M. Embryonic stem cells as a source of models for drug discovery. *Nature Rev. Drug Discov.* **6**, 605–616 (2007).
5. Murry, C. E. & Keller, G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132**, 661–680 (2008).
6. Davila, J. C. *et al.* Use and application of stem cells in toxicology. *Toxicol. Sci.* **79**, 214–223 (2004).
7. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
This was the first demonstration that mouse somatic cells can be reprogrammed into PSCs following the introduction of a defined set of transcription factors.
8. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
9. Lowry, W. E. *et al.* Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl Acad. Sci. USA* **105**, 2883–2888 (2008).
10. Park, I. H. *et al.* Disease-specific induced pluripotent stem cells. *Cell* **134**, 877–886 (2008).
This paper describes human iPSCs created from several diseases, including complex adult-onset diseases.
11. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
References 8, 9 and 11 were the first reports of the generation of iPSCs from human somatic cells by defined transcription factors.
12. Urbach, A., Bar-Nur, O., Daley, G. Q. & Benvenisty, N. Differential modeling of fragile X syndrome by human embryonic stem cells and induced pluripotent stem cells. *Cell Stem Cell* **6**, 407–411 (2010).
13. Lee, G. *et al.* Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature* **461**, 402–406 (2009).
This work demonstrates that reprogrammed cells from patients with monogenic developmental diseases can recapitulate a disease-relevant phenotype, and that these disease phenotypes can be modified by drugs.
14. Carvajal-Vergara, X. *et al.* Patient-specific induced pluripotent stem-cell-derived models of LEOPARD syndrome. *Nature* **465**, 808–812 (2010).
15. Marchetto, M. C. *et al.* A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* **143**, 527–539 (2010).
16. Muotri, A. R. *et al.* L1 retrotransposition in neurons is modulated by MeCP2. *Nature* **468**, 443–446 (2010).
17. Zhang, J. *et al.* A human iPSC model of Hutchinson Gilford progeria reveals vascular smooth muscle and mesenchymal stem cell defects. *Cell Stem Cell* **8**, 31–45 (2011).
This paper demonstrates that several disease-relevant cell types differentiated from reprogrammed cells (taken from patients) show disease phentoypes.
18. Moretti, A. *et al.* Patient-specific induced pluripotent stem-cell models for long-QT syndrome. *N. Engl. J. Med.* **363**, 1397–1409 (2010).
19. Rashid, S. T. *et al.* Modeling inherited metabolic disorders of the liver using human induced pluripotent stem cells. *J. Clin. Invest.* **120**, 3127–3136 (2010).
20. Ebert, A. D. *et al.* Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature* **457**, 277–280 (2009).
This work also demonstrates that reprogrammed cells from patients with monogenic developmental diseases can recapitulate a disease-relevant phenotype, and that these disease phenotypes can be modified by drugs.
21. Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145–1147 (1998).
22. Keller, G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev.* **19**, 1129–1155 (2005).
23. McNeish, J. *et al.* High-throughput screening in embryonic stem cell-derived neurons identifies potentiators of α-amino-3-hydroxyl-5-methyl-4-isoxazolepropionate-type glutamate receptors. *J. Biol. Chem.* **285**, 17209–17217 (2010).
This was the first demonstration that cells differentiated from iPSCs can be used in high-throughput screens to discover new chemical entities.
24. Maitra, A. *et al.* Genomic alterations in cultured human embryonic stem cells. *Nature Genet.* **37**, 1099–1103 (2005).
25. Baker, D. E. *et al.* Adaptation to culture of human embryonic stem cells and oncogenesis *in vivo*. *Nature Biotech.* **25**, 207–215 (2007).
26. Narva, E. *et al.* High-resolution DNA analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity. *Nature Biotech.* **28**, 371–377 (2010).
This paper shows that chromosomal abnormalities accumulate during the passaging of human embryonic stem cells.
27. Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313–317 (2007).
28. Maherali, N. *et al.* Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* **1**, 55–70 (2007).
29. Blelloch, R., Venere, M., Yen, J. & Ramalho-Santos, M. Generation of induced pluripotent stem cells in the absence of drug selection. *Cell Stem Cell* **1**, 245–247 (2007).
30. Meissner, A., Wernig, M. & Jaenisch, R. Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. *Nature Biotech.* **25**, 1177–1181 (2007).
31. Wernig, M., Meissner, A., Cassady, J. P. & Jaenisch, R. c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell* **2**, 10–12 (2008).
32. Boland, M. J. *et al.* Adult mice generated from induced pluripotent stem cells. *Nature* **461**, 91–94 (2009).
33. Kang, L., Wang, J., Zhang, Y., Kou, Z. & Gao, S. iPS cells can support full-term development of tetraploid blastocyst-complemented embryos. *Cell Stem Cell* **5**, 135–138 (2009).
34. Zhao, X. Y. *et al.* iPS cells produce viable mice through tetraploid complementation. *Nature* **461**, 86–90 (2009).
References 32–34 demonstrate the full developmental potential of mouse iPSCs through tetraploid complementation — the most stringent test of developmental potency.
35. Dimos, J. T. *et al.* Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science* **321**, 1218–1221 (2008).
This was the first demonstration that skin cells from patients can be reprogrammed and differentiated into disease-relevant cell types.
36. Loh, Y. H. *et al.* Reprogramming of T cells from human peripheral blood. *Cell Stem Cell* **7**, 15–19 (2010).
37. Loh, Y. H. *et al.* Generation of induced pluripotent stem cells from human blood. *Blood* **113**, 5476–5479 (2009).
38. Rajesh, D. *et al.* Human lymphoblastoid B-cell lines reprogrammed to EBV-free induced pluripotent stem cells. *Blood* **118**, 1797–1800 (2011).
39. Choi, S. M. *et al.* Reprogramming of EBV-immortalized B-lymphocyte cell lines into induced pluripotent stem cells. *Blood* **118**, 1801–1805 (2011).
References 38 and 39 first reported the generation of iPSCs from immortalized blood cell lines, providing the opportunity to obtain patient cells from biobanks.
40. Sommer, C. A. *et al.* Excision of reprogramming transgenes improves the differentiation potential of iPS cells generated with a single excisable vector. *Stem Cells* **28**, 64–74 (2010).
41. Si-Tayeb, K. *et al.* Generation of human induced pluripotent stem cells by simple transient transfection of plasmid DNA encoding reprogramming factors. *BMC Dev. Biol.* **10**, 81 (2010).

© 2011 Macmillan Publishers Limited. All rights reserved

42. Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T. & Yamanaka, S. Generation of mouse induced pluripotent stem cells without viral vectors. *Science* **322**, 949–953 (2008).
43. Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G. & Hochedlinger, K. Induced pluripotent stem cells generated without viral integration. *Science* **322**, 945–949 (2008).
References 42 and 43 were the first reports of reprogramming using vectors that are not integrated into the genome.
44. Yu, J. *et al.* Human induced pluripotent stem cells free of vector and transgene sequences. *Science* **324**, 797–801 (2009).
45. Okita, K. *et al.* A more efficient method to generate integration-free human iPS cells. *Nature Methods* **8**, 409–412 (2011).
46. Kaji, K. *et al.* Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* **458**, 771–775 (2009).
47. Lacoste, A., Berenshteyn, F. & Brivanlou, A. H. An efficient and reversible transposable system for gene delivery and lineage-specific differentiation in human embryonic stem cells. *Cell Stem Cell* **5**, 332–342 (2009).
48. Woltjen, K. *et al. piggyBac* transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* **458**, 766–770 (2009).
49. Kim, D. *et al.* Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell* **4**, 472–476 (2009).
50. Zhou, H. *et al.* Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell* **4**, 381–384 (2009).
51. Cho, H. J. *et al.* Induction of pluripotent stem cells from adult somatic cells by protein-based reprogramming without genetic manipulation. *Blood* **116**, 386–395 (2010).
52. Fusaki, N., Ban, H., Nishiyama, A., Saeki, K. & Hasegawa, M. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc. Jpn Acad. Ser. B Phys. Biol. Sci.* **85**, 348–362 (2009).
53. Seki, T. *et al.* Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. *Cell Stem Cell* **7**, 11–14 (2010).
54. Nishimura, K. *et al.* Development of defective and persistent sendai virus vector: a unique gene delivery/expression system ideal for cell reprogramming. *J. Biol. Chem.* **286**, 4760–4771 (2010).
55. Yakubov, E., Rechavi, G., Rozenblatt, S. & Givol, D. Reprogramming of human fibroblasts to pluripotent stem cells using mRNA of four transcription factors. *Biochem. Biophys. Res. Commun.* **394**, 189–193 (2010).
56. Warren, L. *et al.* Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* **7**, 618–630 (2010).
57. Maehr, R. *et al.* Generation of pluripotent stem cells from patients with type 1 diabetes. *Proc. Natl Acad. Sci. USA* **106**, 15768–15773 (2009).
58. Desponts, C. & Ding, S. Using small molecules to improve generation of induced pluripotent stem cells from somatic cells. *Methods Mol. Biol.* **636**, 207–218 (2010).
59. Li, W. & Ding, S. Small molecules that modulate embryonic stem cell fate and somatic cell reprogramming. *Trends Pharmacol. Sci.* **31**, 36–45 (2010).
60. Chen, J. *et al.* BMPs functionally replace Klf4 and support efficient reprogramming of mouse fibroblasts by Oct4 alone. *Cell Res.* **21**, 205–212 (2010).
61. Zhu, S. *et al.* Reprogramming of human primary somatic cells by OCT4 and chemical compounds. *Cell Stem Cell* **7**, 651–655 (2010).
62. Aalto-Setala, K., Conklin, B. R. & Lo, B. Obtaining consent for future research with induced pluripotent cells: opportunities and challenges. *PLoS Biol.* **7**, e42 (2009).
63. Chou, B. K. *et al.* Efficient human iPS cell derivation by a non-integrating plasmid from blood cells with unique epigenetic and gene expression signatures. *Cell Res.* **21**, 518–529 (2011).
64. Hu, K. *et al.* Efficient generation of transgene-free induced pluripotent stem cells from normal and neoplastic bone marrow and cord blood mononuclear cells. *Blood* **117**, e109–e119 (2011).
65. Kim, K. *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature* **467**, 285–290 (2010).
66. Polo, J. M. *et al.* Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nature Biotech.* **28**, 848–855 (2010).
67. Bar-Nur, O., Russ, H. A., Efrat, S. & Benvenisty, N. Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet β cells. *Cell Stem Cell* **9**, 17–23 (2011).
68. Ohi, Y. *et al.* Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nature Cell Biol.* **13**, 541–549 (2011).
69. Ghosh, Z. *et al.* Persistent donor cell gene expression among human induced pluripotent stem cells contributes to differences with human embryonic stem cells. *PLoS ONE* **5**, e8975 (2010).
70. Hu, B. Y. *et al.* Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proc. Natl Acad. Sci. USA* **107**, 4335–4340 (2010).
71. Hanna, J. *et al.* Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc. Natl Acad. Sci. USA* **107**, 9222–9227 (2010).
72. Bock, C. *et al.* Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* **144**, 439–452 (2011).
This paper describes the generation of an assay to predict the differentiation efficiency of distinct iPSC clones, providing a tool for quality control of such iPSC clones.
73. Boulting, G. L. *et al.* A functionally characterized test set of human induced pluripotent stem cells. *Nature Biotech.* **29**, 279–286 (2011).
74. Kattman, S. J. *et al.* Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. *Cell Stem Cell* **8**, 228–240 (2011).
75. Liu, L. *et al.* Activation of the imprinted *Dlk1-Dio3* region correlates with pluripotency levels of mouse stem cells. *J. Biol. Chem.* **285**, 19483–19490 (2010).
76. Stadtfeld, M. *et al.* Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* **465**, 175–181 (2010).
77. Mayshar, Y. *et al.* Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell* **7**, 521–531 (2010).
78. Pasi, C. E. *et al.* Genomic instability in induced stem cells. *Cell Death. Differ.* **18**, 745–753 (2011).
79. Gore, A. *et al.* Somatic coding mutations in human induced pluripotent stem cells. *Nature* **471**, 63–67 (2011).
80. Laurent, L. C. *et al.* Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. *Cell Stem Cell* **8**, 106–118 (2011).
81. Hussein, S. M. *et al.* Copy number variation and selection during reprogramming to pluripotency. *Nature* **471**, 58–62 (2011).
This study demonstrates that reprogramming leads to an increase in copy number variation that is selected against and disappears after prolonged passaging of iPSCs.
82. Blum, B. & Benvenisty, N. The tumorigenicity of diploid and aneuploid human pluripotent stem cells. *Cell Cycle* **8**, 3822–3830 (2009).
83. Enver, T. *et al.* Cellular differentiation hierarchies in normal and culture-adapted human embryonic stem cells. *Hum. Mol. Genet.* **14**, 3129–3140 (2005).
84. Watanabe, K. *et al.* A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nature Biotech.* **25**, 681–686 (2007).
85. Svendsen, C. N. *et al.* A new method for the rapid and long term growth of human neural precursor cells. *J. Neurosci. Methods* **85**, 141–152 (1998).
86. Chambers, S. M. *et al.* Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature Biotech.* **27**, 275–280 (2009).
87. Soldner, F. *et al.* Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. *Cell* **136**, 964–977 (2009).
88. Osakada, F. *et al. In vitro* differentiation of retinal cells from human pluripotent stem cells by small-molecule induction. *J. Cell Sci.* **122**, 3169–3179 (2009).
89. Sullivan, G. J. *et al.* Generation of functional human hepatic endoderm from human induced pluripotent stem cells. *Hepatology* **51**, 329–335 (2010).
90. Choi, K. D. *et al.* Hematopoietic and endothelial differentiation of human induced pluripotent stem cells. *Stem Cells* **27**, 559–567 (2009).
91. Ye, Z. *et al.* Human-induced pluripotent stem cells from blood cells of healthy donors and patients with acquired blood disorders. *Blood* **114**, 5473–5480 (2009).
92. Taura, D. *et al.* Adipogenic differentiation of human induced pluripotent stem cells: comparison with that of human embryonic stem cells. *FEBS Lett.* **583**, 1029–1033 (2009).
93. Friling, S. *et al.* Efficient production of mesencephalic dopamine neurons by Lmx1a expression in embryonic stem cells. *Proc. Natl Acad. Sci. USA* **106**, 7613–7618 (2009).
94. Hester, M. E. *et al.* Rapid and efficient generation of functional motor neurons from human pluripotent stem cells using gene delivered transcription factor codes. *Mol. Ther.* **19**, 1905–1912 (2011).
95. Panman, L. *et al.* Transcription factor-induced lineage selection of stem-cell-derived neural progenitor cells. *Cell Stem Cell* **8**, 663–675 (2011).
References 93–95 describe improvements in the differentiation of iPSCs using lineage-specific transcription factors.
96. Raya, A. *et al.* Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature* **460**, 53–59 (2009).
97. Itzhaki, I. *et al.* Modelling the long QT syndrome with induced pluripotent stem cells. *Nature* **471**, 225–229 (2011).
98. Nguyen, H. N. *et al.* LRRK2 mutant iPSC-derived DA neurons demonstrate increased susceptibility to oxidative stress. *Cell Stem Cell* **8**, 267–280 (2011).
Together with references 107 and 151, this study identifies disease-relevant phenotypes in reprogrammed cells from patients with adult-onset diseases.
99. Burghes, A. H. & Beattie, C. E. Spinal muscular atrophy: why do low levels of survival motor neuron protein make motor neurons sick? *Nature Rev. Neurosci.* **10**, 597–609 (2009).
100. Crawford, T. O. & Pardo, C. A. The neurobiology of childhood spinal muscular atrophy. *Neurobiol. Dis.* **3**, 97–110 (1996).
101. Slaugenhaupt, S. A. *et al.* Tissue-specific expression of a splicing mutation in the *IKBKAP* gene causes familial dysautonomia. *Am. J. Hum. Genet.* **68**, 598–605 (2001).
102. Chahrour, M. *et al.* MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* **320**, 1224–1229 (2008).
103. Hotta, A. *et al.* Isolation of human iPS cells using EOS lentiviral vectors to select for pluripotency. *Nature Methods* **6**, 370–376 (2009).
104. Hennekam, R. C. Hutchinson–Gilford progeria syndrome: review of the phenotype. *Am. J. Med. Genet. A* **140**, 2603–2624 (2006).
105. Kannankeril, P., Roden, D. M. & Darbar, D. Drug-induced long QT syndrome. *Pharmacol. Rev.* **62**, 760–781 (2010).
106. Roden, D. M. Drug-induced prolongation of the QT interval. *N. Engl. J. Med.* **350**, 1013–1022 (2004).
107. Seibler, P. *et al.* Mitochondrial parkin recruitment is impaired in neurons derived from mutant PINK1 induced pluripotent stem cells. *J. Neurosci.* **31**, 5970–5976 (2011).
Together with references 98 and 151, this study identifies disease-relevant phenotypes in reprogrammed cells from patients with adult-onset diseases.
108. Kola, I. & Landis, J. Can the pharmaceutical industry reduce attrition rates? *Nature Rev. Drug Discov.* **3**, 711–715 (2004).
109. Chong, C. R., Chen, X., Shi, L., Liu, J. O. & Sullivan, D. J. Jr. A clinical drug library screen identifies astemizole as an antimalarial agent. *Nature Chem. Biol.* **2**, 415–416 (2006).
110. Majercak, J. *et al. LRRTM3* promotes processing of amyloid-precursor protein by BACE1 and is a positional candidate gene for late-onset Alzheimer's disease. *Proc. Natl Acad. Sci. USA* **103**, 17967–17972 (2006).
111. Zhou, H. *et al.* Genome-scale RNAi screen for host factors required for HIV replication. *Cell Host Microbe* **4**, 495–504 (2008).
112. Wysowski, D. K. & Swartz, L. Adverse drug event surveillance and drug withdrawals in the United States, 1969–2002: the importance of reporting suspected reactions. *Arch. Intern. Med.* **165**, 1363–1369 (2005).

© 2011 Macmillan Publishers Limited. All rights reserved

113. Pearson, H. The bitterest pill. *Nature* **444**, 532–533 (2006).
114. Kola, I. The state of innovation in drug development. *Clin. Pharmacol. Ther.* **83**, 227–230 (2008).
115. Barter, P. J. *et al.* Effects of torcetrapib in patients at high risk for coronary events. *N. Engl. J. Med.* **357**, 2109–2122 (2007).
116. Tall, A. R., Yvan-Charvet, L. & Wang, N. The failure of torcetrapib: was it the molecule or the mechanism? *Arterioscler. Thromb. Vasc. Biol.* **27**, 257–260 (2007).
117. Joy, T. R. & Hegele, R. A. The failure of torcetrapib: what have we learned? *Br. J. Pharmacol.* **154**, 1379–1381 (2008).
118. Diener, H. C. *et al.* NXY-059 for the treatment of acute stroke: pooled analysis of the SAINT I and II trials. *Stroke* **39**, 1751–1758 (2008).
119. Erondu, N. *et al.* Neuropeptide Y5 receptor antagonism does not induce clinically meaningful weight loss in overweight and obese adults. *Cell Metab.* **4**, 275–282 (2006).
120. International HapMap Consortium. A haplotype map of the human genome. *Nature* **437**, 1299–1320 (2005).
121. Tebbey, P. W. & Rink, C. Target product profile: a renaissance for its definition and use. *J. Med. Market.* **9**, 301–307 (2009).
122. US Food and Drug Administration. Guidance For Industry And Review Staff: Target Product Profile — A Strategic Development Process Tool. *FDA website* [online], http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm080593.pdf (2007).
123. Engelman, J. A. *et al. MET* amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* **316**, 1039–1043 (2007).
124. Frumkin, T. *et al.* Human embryonic stem cells carrying mutations for severe genetic disorders. *In Vitro Cell. Dev. Biol. Anim.* **46**, 327–336 (2010).
125. Tropel, P. *et al.* High-efficiency derivation of human embryonic stem cell lines following pre-implantation genetic diagnosis. *In Vitro Cell. Dev. Biol. Anim.* **46**, 376–385 (2010).
126. Verlinsky, Y. *et al.* Human embryonic stem cell lines with genetic disorders. *Reprod. Biomed. Online* **10**, 105–110 (2005).
127. Mateizel, I. *et al.* Derivation of human embryonic stem cell lines from embryos obtained after IVF and after PGD for monogenic disorders. *Hum. Reprod.* **21**, 503–511 (2006).
128. Marteyn, A. *et al.* Mutant human embryonic stem cells reveal neurite and synapse formation defects in type 1 myotonic dystrophy. *Cell Stem Cell* **8**, 434–444 (2011).
129. Deleu, S. *et al.* Human cystic fibrosis embryonic stem cell lines derived on placental mesenchymal stromal cells. *Reprod. Biomed. Online* **18**, 704–716 (2009).
130. Somers, A. *et al.* Generation of transgene-free lung disease-specific human induced pluripotent stem cells using a single excisable lentiviral stem cell cassette. *Stem Cells* **28**, 1728–1740 (2010).
131. Kazuki, Y. *et al.* Complete genetic correction of iPS cells from Duchenne muscular dystrophy. *Mol. Ther.* **18**, 386–393 (2010).
132. Niclis, J. C. *et al.* Human embryonic stem cell models of Huntington disease. *Reprod. Biomed. Online* **19**, 106–113 (2009).
133. Zhang, N., An, M. C., Montoro, D. & Ellerby, L. M. Characterization of human Huntington's disease cell model from induced pluripotent stem cells. *PLoS Curr.* **2**, RRN1193 (2010).
134. Jang, J. *et al.* Induced pluripotent stem cell models from X-linked adrenoleukodystrophy patients. *Ann. Neurol.* **70**, 402–409 (2011).
135. Ho, J. C. *et al.* Generation of induced pluripotent stem cell lines from 3 distinct laminopathies bearing heterogeneous mutations in lamin A/C. *Aging* **3**, 380–390 (2011).
136. Ghodsizadeh, A. *et al.* Generation of liver disease-specific induced pluripotent stem cells along with efficient differentiation to functional hepatocyte-like cells. *Stem Cell Rev.* **6**, 622–632 (2010).
137. Agarwal, S. *et al.* Telomere elongation in induced pluripotent stem cells from dyskeratosis congenita patients. *Nature* **464**, 292–296 (2010).
138. Batista, L. F. *et al.* Telomere shortening and loss of self-renewal in dyskeratosis congenita induced pluripotent stem cells. *Nature* **474**, 399–402 (2011).
139. Tolar, J. *et al.* Induced pluripotent stem cells from individuals with recessive dystrophic epidermolysis bullosa. *J. Invest. Dermatol.* **131**, 848–856 (2011).
140. Itoh, M., Kiuru, M., Cairo, M. S. & Christiano, A. M. Generation of keratinocytes from normal and recessive dystrophic epidermolysis bullosa-induced pluripotent stem cells. *Proc. Natl Acad. Sci. USA* **108**, 8797–8802 (2011).
141. Mitne-Neto, M. *et al.* Downregulation of VAPB expression in motor neurons derived from induced pluripotent stem cells of ALS8 patients. *Hum. Mol. Genet.* **20**, 3642–3652 (2011).
142. Howden, S. E. *et al.* Genetic correction and analysis of induced pluripotent stem cells from a patient with gyrate atrophy. *Proc. Natl Acad. Sci. USA* **108**, 6537–6542 (2011).
143. Liu, G. H. *et al.* Recapitulation of premature ageing with iPSCs from Hutchinson–Gilford progeria syndrome. *Nature* **472**, 221–225 (2011).
144. Liu, G. H. *et al.* Targeted gene correction of laminopathy-associated *LMNA* mutations in patient-specific iPSCs. *Cell Stem Cell* **8**, 688–694 (2011).
145. Khan, I. F. *et al.* Engineering of human pluripotent stem cells by AAV-mediated gene targeting. *Mol. Ther.* **18**, 1192–1199 (2010).
146. Tolar, J. *et al.* Hematopoietic differentiation of induced pluripotent stem cells from patients with mucopolysaccharidosis type I (Hurler syndrome). *Blood* **117**, 839–847 (2011).
147. Lemonnier, T. *et al.* Modeling neuronal defects associated with a lysosomal disorder using patient-derived induced pluripotent stem cells. *Hum. Mol. Genet.* **20**, 3653–3666 (2011).
148. Hargus, G. *et al.* Differentiated Parkinson patient-derived induced pluripotent stem cells grow in the adult rodent brain and reduce motor asymmetry in Parkinsonian rats. *Proc. Natl Acad. Sci. USA* **107**, 15921–15926 (2010).
149. Swistowski, A. *et al.* Efficient generation of functional dopaminergic neurons from human induced pluripotent stem cells under defined conditions. *Stem Cells* **28**, 1893–1904 (2010).
150. Jin, Z. B. *et al.* Modeling retinal degeneration using patient-specific induced pluripotent stem cells. *PLoS ONE* **6**, e17084 (2011).
151. Brennand, K. J. *et al.* Modelling schizophrenia using human induced pluripotent stem cells. *Nature* **473**, 221–225 (2011).
Together with references 98 and 107, this study identifies disease-relevant phenotypes in reprogrammed cells from patients with adult-onset diseases.
152. Zhang, S. *et al.* Rescue of ATP7B function in hepatocyte-like cells from Wilson's disease induced pluripotent stem cells using gene therapy or the chaperone drug curcumin. *Hum. Mol. Genet.* **20**, 3176–3187 (2011).
153. Zou, J. *et al.* Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPS cells: functional correction by zinc finger nuclease-mediated safe harbor targeting. *Blood* **117**, 5561–5572 (2011).
154. Nakagawa, M. *et al.* Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nature Biotech.* **26**, 101–106 (2008).
155. Anokye-Danso, F. *et al.* Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell* **8**, 376–388 (2011).
156. Zhou, W. & Freed, C. R. Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. *Stem Cells* **27**, 2667–2674 (2009).
157. Jia, F. *et al.* A nonviral minicircle vector for deriving human iPS cells. *Nature Methods* **7**, 197–199 (2010).
158. Miyoshi, N. *et al.* Reprogramming of mouse and human cells to pluripotency using mature microRNAs. *Cell Stem Cell* **8**, 633–638 (2011).
159. Esteban, M. A. *et al.* Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell* **6**, 71–79 (2010).
160. Lin, T. *et al.* A chemical platform for improved induction of human iPSCs. *Nature Methods* **6**, 805–808 (2009).
161. Mali, P. *et al.* Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes. *Stem Cells* **28**, 713–720 (2010).
162. Haase, A. *et al.* Generation of induced pluripotent stem cells from human cord blood. *Cell Stem Cell* **5**, 434–441 (2009).
163. Eminli, S. *et al.* Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nature Genet.* **41**, 968–976 (2009).
164. Giorgetti, A. *et al.* Generation of induced pluripotent stem cells from human cord blood using *OCT4* and *SOX2*. *Cell Stem Cell* **5**, 353–357 (2009).
165. Lagarkova, M. A. *et al.* Induction of pluripotency in human endothelial cells resets epigenetic profile on genome scale. *Cell Cycle* **9**, 937–946 (2010).
166. Sun, N. *et al.* Feeder-free derivation of induced pluripotent stem cells from adult human adipose stem cells. *Proc. Natl Acad. Sci. USA* **106**, 15720–15725 (2009).
167. Aoki, T. *et al.* Generation of induced pluripotent stem cells from human adipose-derived stem cells without c-MYC. *Tissue Eng. Part A* **16**, 2197–2206 (2010).
168. Aasen, T. *et al.* Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nature Biotech.* **26**, 1276–1284 (2008).
169. Carey, B. W. *et al.* Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proc. Natl Acad. Sci. USA* **106**, 157–162 (2009).
170. Li, W. *et al.* Generation of human-induced pluripotent stem cells in the absence of exogenous *Sox2*. *Stem Cells* **27**, 2992–3000 (2009).
171. Kim, J. B. *et al.* Direct reprogramming of human neural stem cells by *OCT4*. *Nature* 28 Aug 2009 (doi:10.1038/nature08436).
172. Ruiz, S. *et al.* High-efficient generation of induced pluripotent stem cells from human astrocytes. *PLoS ONE* **5**, e15526 (2010).
173. Liu, H., Ye, Z., Kim, Y., Sharkis, S. & Jang, Y. Y. Generation of endoderm-derived human induced pluripotent stem cells from primary hepatocytes. *Hepatology* **51**, 1810–1819 (2010).
174. Li, C. *et al.* Pluripotency can be rapidly and efficiently induced in human amniotic fluid-derived cells. *Hum. Mol. Genet.* **18**, 4340–4349 (2009).
175. Zhao, H. X. *et al.* Rapid and efficient reprogramming of human amnion-derived cells into pluripotency by three factors *OCT4/SOX2/NANOG*. *Differentiation* **80**, 123–129 (2010).

Acknowledgements
The authors would like to especially thank S. Irion, E. Vaisberg, I. Griswold-Prenner and C. Johnson for their crucial input and help with writing the manuscript. Special thanks to A. Rosenthal for his help in editing and integrating the manuscript, and M. Smith and B. Keon for their help with the graphic design.

Competing interests statement
The authors declare competing financial interests: see web version for details.

FURTHER INFORMATION
George Q. Daley's homepage: http://daley.med.harvard.edu/index.htm
Coriell Institute for Medical Research: http://www.coriell.org
Guidelines for Clinical Translation of Stem Cells: http://www.isscr.org/GuidelinesforClinicalTranslation/2480.htm
International Haplotype Mapping project: http://hapmap.ncbi.nlm.nih.gov
iPierian website: http://www.ipierian.com
Johns Hopkins Clinical Compound Collections: http://www.hopkinschemcore.org/facility/com_coll.html
MicroSource Spectrum Collection: http://www.msdiscovery.com/spectrum.html
Prestwick Collection: http://www.prestwickchemical.com/index.php?pa=26
Sigma-Aldrich Library of Pharmacologically Active Compounds: http://www.sigmaaldrich.com/chemistry/drug-discovery/validation-libraries/lopac1280-navigator.html
UK Biobank: http://www.ukbiobank.ac.uk

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

© 2011 Macmillan Publishers Limited. All rights reserved