We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,800 Open access books available
116,000 International authors and editors
120M Downloads
154 Countries delivered to
TOP 1% Our authors are among the most cited scientists
12.2% Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
1. Introduction

Human pluripotent stem cells (hPSCs, which include both human embryonic and induced pluripotent stem cells [hESCs and hiPSCs]) provide a unique model system to study early human development, derive functionally mature cell populations, and hold great promise to advance medical treatments for currently incurable diseases. Modulating gene function in these cells is of critical importance to gain insight into the molecular underpinnings of the pluripotent state and the process of differentiation. Furthermore, efficient transgenesis in hPSC is essential in generating cells carrying specific disease-associated alleles, thereby enabling the study of human diseases at the cellular level.

In contrast to murine ESCs, hESCs have proven quite recalcitrant to gene modification using many of the traditional methods. The reason(s) for this difference between mESCs and hESCs are poorly understood, but it has been postulated that hESCs represent a slightly different stage with distinct developmental potential than mESCs and that this difference accounts for the differences observed in transgenesis and homologous recombination (Buecker et al., 2010).

DNA can be introduced into hPSCs using viral and non-viral methods, yielding both transient and stable transduction. Stable integration into the genome can occur at random sites, thus potentially disrupting endogenous gene expression or function. In addition, transduced cell lines can be selected to carry targeted gene insertions through site-specific recombination or homologous recombination. In this chapter, we describe several methods to transduce genetic elements into hPSCs and discuss their strengths and weaknesses.

2. Random transgenesis

Several methods have been used successfully to randomly integrate DNA into the genome of hPSCs and thereby generating marker lines that express reporter genes in a cell- or tissue-specific manner. The most commonly used methods to transduce DNA into cells, either transiently or stably, involve chemical-based transfection reagents, electroporation, or viral infection. The advantages and limitations of each transgenic method will be discussed in the following section. In general, transgenic approaches to create reporter lines are limited by the fact that regulatory elements controlling cell-specific gene expression are often either poorly defined or located at great distances from the gene. To faithfully reproduce a gene
expression pattern may require DNA fragment lengths that exceed conventional recombinant DNA methods. In addition, stably transduced reporter genes may be under control of endogenous regulatory elements near the site of integration. For example, genetic integration into certain compact regions of the genome can lead to silencing or lower levels of expression due to reduced accessibility of transcription factors. Another caveat of using random insertional transgenesis is that in rare instances gene insertion may interrupt expression of other genes that influence cell phenotype. As a result of these positional effects, screening and analysis of multiple clones for reporter gene expression patterns over time and during differentiation must be performed. Unfortunately, generation of single-cell clones is often difficult given the low cloning efficiency of hESCs (Amit et al., 2000). Recently, this challenge has been partially mitigated by the use of neurotrophins (such as NT3 and NT4) (Pyle et al., 2006) or selective inhibitors of Rho-associated kinase ROCK (e.g., Y-27632) (Damoiseaux et al., 2009; Watanabe et al., 2007), both of which increase survival of dissociated hESCs.

The methods described in this section have been used to introduce transgenes encoding fluorescent proteins or drug resistance markers under the transcriptional control of cell-specific promoters into hPSCs. Specific sub-populations of cells expressing the reporter gene can be enriched and isolated using drug selection or flow cytometry-based cell sorting.

2.1 Electroporation

Electroporation involves the use of electric pulses of varying voltage, length, and number to transiently increase the permeability of the cell membrane and permit entry of DNA. Several groups have reported the use of electroporation for the transfection of hESCs (Eiges et al., 2001; Lakshmipathy et al., 2004; Mohr et al., 2006; Zwaka and Thomson, 2003). Using a plasmid containing a gene encoding green fluorescent protein (GFP) under control of a cytomegalovirus immediate-early gene 1 (CMV) promoter/enhancer, Mohr et al. investigated the electroporation parameters of pulse, voltage, duration, and number on hESC viability and transfection efficiency (Mohr et al., 2006). A single 300V, 0.5 millisecond pulse achieved an optimum survival rate of 50% and a transient transfection rate of 10%. These values were consistent with other reported transfection efficiencies using electroporation (Lakshmipathy et al., 2004). It is important to note that this study focused on the transient transfection of plasmid DNA. In contrast, stable transfection efficiencies are much lower because of the low rate of DNA integration into the genome. Typically, transient transgene expression in hESCs is not retained for longer than 5 passages (Vallier et al., 2004). By comparison, the yield of stably transfected clones using electroporation has been reported to be on the order of 1 in $10^{-5}$ (Adewumi et al., 2007; Zwaka and Thomson, 2003). Moreover, less than 1% of electroporated hESCs in single-cell suspension survive (Zwaka and Thomson, 2003). However, electroporation of hESCs in clumps suspended in standard cell culture medium greatly improved cell survival (Zwaka and Thomson, 2003).

2.2 Nucleofection

Nucleofection is a recently developed and commercialized technology (Amaxa® Nucleofector® Technology, Lonza) that expands on the principles of electroporation. Specifically, nucleofection is a customized procedure in which buffer solutions and electric pulses are optimized for specific cell types (Siemen et al., 2005; Siemen et al., 2008). In
contrast to electroporation, in which efficiency is dependent on cell cycle because DNA only enters the nucleus when the nuclear envelope is fragmented during cell division, nucleofection leads to a direct transfer of the DNA into the cell nucleus (Hamm et al., 2002). Nucleofection of a GFP reporter construct into hESCs yielded survival rates of >70% and a transient transgene expression of 66% in surviving cells. Furthermore, transfected cells maintained expression of markers of pluripotency, such as TRA-1-60 and OCT4, and could be expanded to stably expressing clones. Additionally, the small amount of cells and DNA required for nucleofection has allowed it to be a valuable tool for high-throughput genetic manipulation of hESCs (Moore et al., 2010).

2.3 Chemical-based transfection methods

As an alternative to electroporation, chemical transfection reagents (lipid and polymer-based reagents) have been used in attempts to achieve substantial levels of transgene expression in hESCs. The commonly used calcium-phosphate (Ca/PO\textsubscript{4}) mediated transfection methods have had limited success in hESCs, due to the cytotoxicity of the Ca/PO\textsubscript{4} precipitate.

An early report compared the efficiency of three different lipofection-based reagents, Lipofectamine (Invitrogen), Fugene (Boehringer Mannheim) and ExGen 500 (Fermentas), in hESCs by using an expression construct in which GFP was under control of the elongation factor I (EF1\text{α}) promoter (Eiges et al., 2001). Transient transfection with ExGen 500 resulted in transgene expression an order of magnitude higher than Lipofectamine or Fugene. However, GFP transgene expression occurred in only 10% of the ExGen500 transfected cells, a transfection efficiency comparable to electroporation. This report conflicts with that of others (Siemen et al., 2005), which state that ExGen 500 was less effective than other reagents such as Lipofectamine and Fugene. One study compared several chemical transfection reagents and confirmed the poor efficiency of ExGen 500 and determined that GeneJammer was the most effective method (Anderson et al., 2007). A recent report describes the development of a class of polymers, poly(β-amino esters), which have higher gene delivery efficiencies because of their ability to condense DNA into nanoparticles, thus enhancing cellular uptake (Green et al., 2008). These nanoparticles have gene delivery efficiencies 4 times higher than other transfection reagents such as Lipofectamine. It is important to note that the reported differences between the efficacies of these reagents could be attributed to differences in the culture systems and individual hESC lines tested. Therefore, some reagents may be better suited for use in specific hESC lines.

In order to achieve stable transfection into hESCs using chemical-based transfection methods, a gene encoding antibiotic-resistance or fluorescent protein must be present in the construct so that stable transfecants can be selected. HESC lines developed employing such methods have been utilized to monitor the differentiation status of hESCs. Eiges et al. used lipofection to transfect hESCs with GFP under the control of a murine Rex1 promoter, a gene that is rapidly downregulated upon differentiation of hESCs (Eiges et al., 2001). The transfected hESCs could be sorted to homogeneity using fluorescent-activated cell sorting (FACS). The purified cell population showed high levels of GFP expression when in an undifferentiated state. As the cells differentiated this expression was significantly reduced. Along similar lines, lipofection was used to generate a hESC line with GFP driven by an OCT4 promoter. Such OCT4-GFP lines have proven useful in monitoring the exit from the undifferentiated state and acquisition of differentiation markers.
Chemical-based transfection methods have also been used to tag and purify specific hESC derivatives. For example, chemical transgenic methods were used to enrich cardiomyocytes using a reporter system in which the human $\alpha$-myosin promoter drove expression of puromycin-N-acetyltransferase (PAC)(Anderson et al., 2007). Positive selection with puromycin of hESCs differentiated towards cardiomyocytes resulted in a cardiomyocytes enrichment of 14.5 fold (change from 7% to 92% cardiomyocytes). Enriched cardiomyocytes expressed high levels of cardiac specific markers and displayed cardiac-specific action potentials, demonstrating that functionality was not compromised by drug selection. In another example, lipofection was used to transfect hESCs with plasmids encoding GFP placed under the transcriptional control of a motor neuron specific enhancer within the 5'-regulatory region of the gene encoding the transcription factor Hb9. Motor neurons induced in vitro with the growth factor SHH and retinoic acid activated Hb9-driven GFP expression, allowing their isolation by FACS. Sorted cells achieved physiological and functional maturation in vitro, demonstrating the feasibility of promoter/enhancer-based FACS for the isolation of specific derivatives from hESCs. Chemical transfection has also been used to generate hESC lines with fluorescent reporter transgenes under transcriptional control of murine albumin (Lavon and Benvenisty, 2005; Lavon et al., 2004) and Pdx1 (Lavon et al., 2006) genes in order to identify hESC differentiating toward the hepatic and pancreatic lineages, respectively.

2.4 Lentiviral systems
Lentiviral vectors offer another strategy to stably introduce DNA into hESCs. Pfeifer et al. first reported the use of lentiviral vectors for high levels of transduction efficiency without transgene silencing in hESCs (Pfeifer et al., 2002). Specifically, almost 100% efficiency was reported at a multiplicity of infection of 50, and transgene expression was maintained over several passages. Several additional groups have reported high-level sustained transgene expression in hESCs and hESC derivatives using similar viral vectors. For example, Gropp et al. used lentiviral vectors to stably transduce a GFP encoding transgene under the control of an EF1$\alpha$ promoter (Gropp et al., 2003). Transduction of hESCs with these lentiviral vectors allowed for stable transgene expression for long-term (38 weeks) undifferentiated culture. Furthermore, transgene expression was not silenced upon differentiation as demonstrated by continuous GFP expression throughout differentiation both in vitro (EB formation) and in vivo (teratoma formation).

Lentiviral vectors have also been used to identify and select hESC derivative cell populations. For example, Huber et al. generated stable transgenic hESC lines using lentiviral vectors and isolated single-cell clones that expressed a GFP transgene under the transcriptional control of a cardiac specific promoter, human myosin light-chain-2V promoter (Huber et al., 2007). As hESCs differentiated, GFP expressing cells were isolated and purified to near homogeneity using FACS. The GFP-positive cells stained for cardiac specific markers, expressed cardiac-specific genes, displayed cardiac-specific action-potentials, and demonstrated long-term engraftment in rat hearts. A recent report expanded on these methods to engineer hESCs with lentiviral vectors that combined blasticidin, neomycin, and puromycin resistance based drug selection of pure populations of stem cells and cardiomyocytes with constitutive or lineage-specific promoters that control expression of fluorescent proteins. This allowed for cardiomyocytes and their progenitors to be visualized and tracked (Kita-Matsuo et al., 2009). Specifically, hESCs were infected with
lenti vectors carrying Brachyury (T) and α-myosin heavy chain (αMHC) promoters driving expression of fluorescent or drug-resistance proteins. HESCs differentiated to early mesoderm and cardiomyocytes were enriched using cell sorting or drug selection. Moreover, a drug selection differentiation protocol yielded 96% pure cardiomyocytes that could be cultured for over 4 months. Additionally, these drug-selected cardiomyocytes exhibited a gene expression profile similar to that of adult human cardiomyocytes and generated force and action potentials consistent with normal fetal cardiomyocytes. Lentiviral approaches have also been used to engineer hESCs with glial fibrillary acidic protein (GFAP) (Dhara et al., 2009) and alpha-fetoprotein (AFP) (Chiao et al., 2008) promoters driving expression of genes encoding fluorescent proteins to enrich for hESCs differentiating towards the neural and hepatic lineages, respectively.

2.5 Comparison of random transgenic methods

Very few studies have directly compared the non-viral and viral gene delivery methods discussed above. A recent study compared the transduction efficiencies and effect on cell viability of each method in several hESC lines using a fusion construct with an ubiquitin promoter driving GFP and firefly luciferase along with a neomycin selection marker (Cao et al., 2009). Lentiviral transduction demonstrated the highest efficiency (range: 22.4-25.3%) with >95% cell viability. Nucleofection demonstrated significantly lower efficiency (range: 5.8-16.1%) with moderate cell survival (range: 70-75%). Minimal transfection efficiency was observed with electroporation (range: 1.9-2.1%) and lipofection (range: 1.3-1.5%). Electroporation resulted in the lowest cell survival (range: 38-58%) while cell survival with lipofection (>90%) was comparable to that of lentiviral infection. Moreover, lentiviral transduction resulted in the greatest number of hESCs stably expressing the fusion reporter gene.

Even though lentiviral transduction results in the highest efficiency and cell survival of all transgenic approaches, it has several drawbacks that may make use of other transgenic methods more attractive. One major weakness of lentiviral methods is that constructs larger than 6-8 kb package poorly, thereby reducing the efficiency of transduction and limiting the size of DNA that can be used (Kumar et al., 2001). Another disadvantage of lentiviral vectors is that they integrate randomly and thus have the ability to activate or inactivate nearby genes (Kohn et al., 2003). A final concern that may limit the use of lentiviral vectors is that viral transgenes may be silenced in the infected cells. Several studies have shown that gene silencing can occur during propagation of undifferentiated ESCs (Cherry et al., 2000) as well as during differentiation (Laker et al., 1998). Such silencing is likely due to epigenetic modifications of the viral DNA following its integration. Furthermore, studies have demonstrated that transgenes delivered by lentiviral vectors are suppressed in ESCs in a promoter-dependent manner (Hong et al., 2007; Xia et al., 2007). For example, Xia et al. examined the silencing of four ubiquitous promoters in lentiviral vectors driving expression of GFP in hESCs: CMV, hybrid CMV enhancer/chicken β-actin (CAG), phosphoglycerate kinase (PGK) and EF1α (Xia et al., 2007). By calculating the ratio between the percentage of nonfluorescent cells and the GFP copy number per cell, the authors showed that more than 95% of the GFP copies driven by the CMV or CAG promoter and 75% of the GFP copies driven by the EF1α promoter were inactive. Meanwhile, GFP driven by the PGK promoter showed the least extent of suppression with ~55% inactive. Results such as these clearly demonstrate that genes delivered through the use of lentivirus are highly susceptible to
gene silencing and that the choice of promoter is critical for long term ubiquitous expression of transgenes. Whether similar gene silencing occurs with cell- and tissue-specific promoters remains to be explored.

3. Site-specific recombination

One way to avoid the problems caused by random integration of transgenes into the genome is the creation of hESCs lines where large, complex genetic elements can be stably introduced into cells at defined chromosomal locations in a single copy. Several studies have identified sites on chromosomes 2, 6, 10, 12, 13, 17, and 21 that are transcriptionally active in undifferentiated hESCs and remain transcriptionally active upon differentiation into all tested mature cell populations (Costa et al., 2005; Irion et al., 2007; Thyagarajan et al., 2008). Two strategies that have been used for site-specific recombination to introduce transgenes into hESCs are the Cre/loxP and phiC3I recombination systems.

3.1 Cre/loxP

The Cre protein is a site-specific DNA recombinase that catalyzes the recombination of DNA between loxP sequences which contain binding sites for Cre. When cells that have loxP sites in their genome express Cre, a recombination event occurs between the loxP sites (Bouhassira et al., 1997). Depending on the orientation of the loxP sites relative to each other, the recombination reaction can result in an insertion, deletion (direct repeat of two loxP sites), inversion (two inverted loxP on the same chromosome), or translocation (two loxP sites on different chromosomes). Nolden et al. first reported the Cre-mediated recombination of a chromosomally integrated loxP-modified allele in hESCs and hESC derived neural progenitors (Nolden et al., 2006). In this system, a constitutively active CAG promoter drove transcription of the HcRed (an enhanced version of red fluorescent protein [RFP]) while a neomycin-resistance gene (Neo) was under transcriptional control of the PGK promoter, thus allowing for selection of stable clones. Cre-mediated recombination caused deletion of HcRed and Neo resulting in the expression of a second reporter GFP. In another study, Cre-mediated recombination was used to insert genes at the human homolog of the mouse Rosa26 locus (hROSA26) (Irion et al., 2007). Specifically, a vector containing a loxP-flanked promoterless Neo-resistance gene followed by an inverted RFP variant, tandem dimer RFP (tdRFP), flanked by mutant loxP2272 was targeted to the hROSA26 locus using homologous recombination (see below for further discussion of homologous recombination). The loxP and mutant loxP2272 sites were positioned so that after Cre expression, the neomycin resistance cassette was deleted and the tdRFP is inverted, placing it under control of the endogenous hROSA26 promoter. The use of wildtype and mutant loxP allowed for the exchange of the tdRFP cassette with any cDNA of interest at the hROSA26 locus using Cre recombinase-mediated cassette exchange.

One drawback of the Cre-loxP system is that its efficiency depends on the level of Cre expression, which is typically achieved by either transfection or viral methods. As a result, recombination only occurs in a low percentage of the cell population (10-15%). As an alternative strategy to introduce Cre, several groups have reported the use of cell-permeable versions of Cre recombinase that induce recombination by direct protein delivery and lead to higher levels of recombination (>90%) (Jo et al., 2001; Joshi et al., 2002; Will et al., 2002).
3.2 phiC3I
Another method to carry out site-specific recombination in hESCs is the use of bacteriophage phiC3I integrase. Unlike Cre recombinase, the phiC3I catalyzes recombination between non-identical sites which makes recombination unidirectional, ensuring that integrated constructs do not act as substrates for the reverse excision reaction. PhiC3I carries out site-specific recombination between its attachment site (attB) and an attachment site (attP) in host DNA. Moreover, phiC3I integrases target a small number of loci (pseudo-attP sites) in transcriptionally active regions (“hot spots”) of the human genome.

Recently, hESC lines were developed in which phiC31 integrase was used to place a target site for the R4 integrase into a pseudo attP site, and the R4 integrase was used to place genes of interest into the specific R4 target site (Figure 1) (Liu et al., 2009; Thyagarajan et al., 2008). Specifically, hESCs were co-transfected with a plasmid with a CMV promoter driving expression of phiC3I integrase and a plasmid containing the R4 attP target site upstream of a promoter-less Zeocin-resistance gene, as well as a hygromycin-resistance marker and a phiC3I attB site (pJT1/Zeo). Individual colonies carrying a R4 attP site upstream of the Zeocin selectable marker was achieved with selection with hygromycin and screened for insertion into one of the genome’s pseudo-att sites (e.g. chromosomal location 13q32 represents a hot spot). Retargeting these “R4” cell lines was achieved by co-transfecting a plasmid expressing the R4 integrase with a plasmid that had the human EF1α promoter upstream of the complementary R4 attB recognition site (pER4B-EG and pER4B-hOG). Recombination between the R4 attP and attB sites positioned the EF1α promoter upstream of the Zeocin-resistance gene. Figure and legend adapted from Liu et al., 2009.

Fig. 1. Site-specific recombination strategy using phiC3I integrase. In the first step a “platform line” is created using phiC3I integrase, which mediates site-specific recombination of the pJT1/Zeo plasmid into one of the host cell genome’s pseudo-att sites. In a second “retargeting” step, the R4 integrase mediates site-specific recombination of the pER4B vector with the previously inserted att site. Proper R4 recombination events lead to the transcriptional activation of the Zeocin-resistance gene. Figure and legend adapted from Liu et al., 2009.
the Zeocin-resistance gene. Thus, only clones that had undergone a site-specific recombination event were resistant to Zeocin. This method has been used to place cassettes with GFP driven by promoters such as EF1α and OCT4 in specific loci in hESCs. Retargeting was extremely efficient with an average of 17 colonies per 10^6 transfected cells with all Zeocin resistant clones containing a single integration event occurring at the chromosome 13q23 hot spot.

A major disadvantage of the phiC3I site-specific integration techniques is that it requires the generation of "platform" hESC lines (e.g. the "R4" cell line). Although retargeting efficiency is extremely high (100% after Zeomycin selection), initial targeting to one of the pseudo-att sites is inefficient and requires the screening of hundreds of clones. Given that hESC lines vary in their behavior and propensity to differentiate into various lineages (Osafune et al., 2008), multiple "platform" lines would need to be generated.

4. Homologous recombination

Homologous recombination (HR) has long been a powerful "reverse genetics" approach to interrogate gene function in mice. Methods for HR exploit the endogenous DNA repair and recombination machinery to alter specific sequences within the genome by aligning the engineered transduced DNA sequences with genomic sequences. The outcome can include single base pair mutations, deletion of entire exons or open reading frames, and insertion of sequences that encode protein tags or reporter proteins, such as fluorescent proteins. In contrast to other genetic modifications described above, HR faithfully reproduces expression patterns of genes with largely insignificant effects on expression of neighboring genes.

There are three main reasons to target genes by HR in hESCs. First, hESC lines can be designed to express a reporter gene under the control of an endogenous and developmentally regulated promoter. Such marker lines are useful to monitor and interrogate the developmental progression from undifferentiated to mature cell states and to optimize protocols for efficient and directed differentiation. By using cell isolation and purification methods, such as FACS, even extremely rare and specialized cell types can be isolated and separated from other cell types, especially from undifferentiated hESC that carry the potential to generate tumors. Such purified cell populations are essential for cell transplantation studies and eventual cell replacement therapies. Second, efficient HR can be used to attach tags onto genes so that the function of their protein products can be studied in vitro and in vivo. By modifying genes with tags that permit detection and purification, endogenous protein function can be explored without relying on current overexpression approaches that are prone to yield artificial results. Finally, HR can be employed to introduce specific mutations in disease-associated genes to develop models of human disease. Currently, the majority of research to study human diseases in cell culture utilize hiPSCs derived from patients with specific genetic defects or predispositions. However, methods for iPSC generation are still wrought with potential complications, e.g. the various transgenes required for reprogramming may influence disease progression. Efficient HR in hPSCs would provide a powerful alternative method to insert mutations in disease associated genes and model human diseases.

In comparison to murine ESCs, a limited, yet gradually expanding number of genes has been modified by HR in hESCs (see Table 1 for examples). One reason for the limited success of HR is that hPSCs, in contrast to murine ESCs, grow poorly as single cells, a necessary condition to isolate the cells carrying the rare HR events. This problem can be
partially overcome by treating dissociated cells with ROCK inhibitor, which has been shown to increase survival of dissociated hESCs (Damoiseaux et al., 2009; Watanabe et al., 2007). Most reports of HR in hESCs have targeted the X-chromosome linked HPRT1 (Hypoxanthine-guanine phosphoribosyltransferase) locus, a gene which is inherently easier to target since deletion of the single copy of HPRT in male cells yields cells that are resistant to the drug 6-Thioguine (6-TG). While these studies provide compelling proof-of-principle, methods for HR in hESCs are far from optimized. In this section, we discuss four methods that show great promise in generating HR in hESCs.

### 4.1 Conventional HR method

The conventional HR method in mESCs utilizes standard transduction methods, such as electroporation (see descriptions above), to introduce linearized DNA constructs comprised of homology arms of 3 to 5 kilo bases (kb) flanking drug selection cassettes, such as Neomycin or Hygromycin resistance genes. These drug resistance genes are placed between the homology arms to disrupt or delete one or more exons of the targeted gene. Positive drug selection alone yields both targeted and random insertions into the genome. To increase the frequency of targeted insertions, many strategies incorporate a negative selectable marker, generally Herpes simplex virus thymidine kinase (HSV-TK), which confers sensitivity to Gancyclovir (Ganc), inserted distally to one of the homology arms (Mansour et al., 1988). Cells in which a random insertion has occurred will have retained HSV-TK and will consequently be eliminated in the presence of Ganc. Cells that are resistant to Ganc have not integrated the HSV-TK gene, thus potentially representing a HR event. HR rates, which vary between <0.1-5%, are generally indicated as a percentage of cells carrying a targeted insertion (as determined by Southern blot or PCR) among all Neomycin resistant clones. To further increase targeting efficiencies it is critical to utilize isogenic DNA, i.e. the DNA used to create the homology arms in the targeting vector matches the DNA of the cells used for targeting. While this is possible with the various in-bred mouse strains, the large number of single nucleotide polymorphisms (SNP) in the human population (one every 1,000 base pairs) rule out an approach that utilizes isogenic DNA. However, it is noteworthy that two reports found that HR in humans is efficient even with non-isogenic DNA (Sedivy et al., 1999; Urbach et al., 2004). The need for isogenic DNA is overcome as the length of the homology arms increases, however, plasmids of greater than 15 kb are difficult to engineer and propagate. This size restriction is overcome with the use of bacterial artificial chromosomes (see below). Due to the low targeting frequencies observed for many loci, this conventional HR method rarely succeeds at disrupting both copies of a gene. To date only a small number of genes have been successfully targeted in hESCs using positive-negative selection methods (see Table 1).

<table>
<thead>
<tr>
<th>Gene</th>
<th>hESC line</th>
<th>HR Method</th>
<th>Efficiency</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATM</td>
<td>HUES9, H9</td>
<td>BAC</td>
<td>21% (3/14) for first allele, 27% (10/37) for second allele</td>
<td>(Song et al., 2010)</td>
</tr>
<tr>
<td>FEZF2</td>
<td>HUES9</td>
<td>Conventional positive (Neo)-negative (HSV-TK) selection</td>
<td>1.5% (2/130 clones)</td>
<td>(Ruby and Zheng, 2009)</td>
</tr>
<tr>
<td>HMGA1</td>
<td>BG01, BG02, H1</td>
<td>AAV</td>
<td>96% (24 of 25 HygR clones)</td>
<td>(Khan et al., 2010)</td>
</tr>
<tr>
<td>Gene</td>
<td>Cell Line</td>
<td>Method</td>
<td>Percentage</td>
<td>Reference</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>--------</td>
<td>------------</td>
<td>-----------</td>
</tr>
<tr>
<td>HPRT1</td>
<td>H1</td>
<td>Conventional positive (Neo)-negative (6TG) selection</td>
<td>2% (7 6-TG&lt;sup&gt;R&lt;/sup&gt; out of 350 G418&lt;sup&gt;R&lt;/sup&gt;)</td>
<td>(Zwaka and Thomson, 2003)</td>
</tr>
<tr>
<td></td>
<td>H13</td>
<td>Conventional positive (Hyg)-negative (6TG) selection</td>
<td>~1-2%</td>
<td>(Urbach et al., 2004)</td>
</tr>
<tr>
<td></td>
<td>H1</td>
<td>Conventional positive (Neo)-negative (6TG) selection</td>
<td>2.3% (6 6TG&lt;sup&gt;R&lt;/sup&gt; out of 260 G418&lt;sup&gt;R&lt;/sup&gt;)</td>
<td>(Di Domenico et al., 2008)</td>
</tr>
<tr>
<td></td>
<td>KhES-1</td>
<td>HDAdV</td>
<td>45% (14 of 31 Ganc&lt;sup&gt;R&lt;/sup&gt; and G418&lt;sup&gt;R&lt;/sup&gt; clones)</td>
<td>(Suzuki et al., 2008)</td>
</tr>
<tr>
<td></td>
<td>HUES3, HUES 8</td>
<td>BAC</td>
<td>HUES3: 1 of 3 Puro&lt;sup&gt;R&lt;/sup&gt; clones HUES8: 1 of 6 Puro&lt;sup&gt;R&lt;/sup&gt; clones</td>
<td>(Song et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>BG01, iPSCs</td>
<td>AAV</td>
<td>BG01: 26% (8 of 31 G418&lt;sup&gt;R&lt;/sup&gt; clones) iPSCs: 19-29%</td>
<td>(Khan et al., 2010)</td>
</tr>
<tr>
<td>IL2RG</td>
<td>HUES1, HUES3</td>
<td>ZFN and IDLV</td>
<td>13-39%</td>
<td>(Lombardo et al., 2007)</td>
</tr>
<tr>
<td>MIXL1</td>
<td>HES3</td>
<td>Conventional positive (Neo) selection</td>
<td>Not indicated</td>
<td>(Davis et al., 2008)</td>
</tr>
<tr>
<td>OLIG2</td>
<td>BG01</td>
<td>Conventional positive (Neo)-negative (HSV-TK) selection</td>
<td>5.7% (6/106 clones)</td>
<td>(Xue et al., 2009)</td>
</tr>
<tr>
<td>p53</td>
<td>HUES9, H9</td>
<td>BACs</td>
<td>3.3% (2/60) for first allele, 22% (7/32) for second allele</td>
<td>(Song et al., 2010)</td>
</tr>
<tr>
<td>PIG-A</td>
<td>H1, H9, iPSC lines</td>
<td>ZFN</td>
<td>50% (6 out 12 Hyg&lt;sup&gt;R&lt;/sup&gt; clones)</td>
<td>(Zou et al., 2009)</td>
</tr>
<tr>
<td>PITX3</td>
<td>BG01, iPSCs</td>
<td>ZFN</td>
<td>11%</td>
<td>(Hockemeyer et al., 2009)</td>
</tr>
<tr>
<td>POU5F1 (OCT4)</td>
<td>H1</td>
<td>Conventional positive (Neo) selection</td>
<td>27% (28/103 G418&lt;sup&gt;R&lt;/sup&gt; clones) 39% (22/56 G418&lt;sup&gt;R&lt;/sup&gt; clones)</td>
<td>(Zwaka and Thomson, 2003)</td>
</tr>
<tr>
<td></td>
<td>BG01</td>
<td>ZFN</td>
<td>39-100% depending on ZFN pair</td>
<td>(Hockemeyer et al., 2009)</td>
</tr>
<tr>
<td>ROSA26</td>
<td>HES2</td>
<td>Conventional positive (Neo)-negative (D-TA) selection</td>
<td>2.3% (2/88 G418&lt;sup&gt;R&lt;/sup&gt; clones)</td>
<td>(Irion et al., 2007)</td>
</tr>
</tbody>
</table>

Table 1. Examples of genes targeted by homologous recombination in hPSCs. Abbreviations: 6-TG<sup>R</sup> = resistance to drug 6-thioguanine; BAC = Bacterial artificial chromosome; D-TA = diphtheria toxin A-fragment; G418<sup>R</sup> = resistance to drug G418; Ganc<sup>R</sup> = resistance to drug ganciclovir; HSV-TK = herpes simplex virus thymidine kinase; Hyg<sup>R</sup> = resistance to drug Hygromycin; IDLV = integration defective lenti virus; iPSC = induced pluripotent stem cell; Neo = Neomycin; Puro<sup>R</sup> = resistance to drug Puromycin; ZNF = Zinc finger nuclease.
4.2 BAC-based HR
A major reason for the low HR efficiency observed for conventional methods in both mouse and human ESCs is related to the limited length of homology arms. It has been documented that with an increase in homology arm lengths from 1.3 to 6.8 kb, gene targeting frequencies dramatically increase (Hasty et al., 1991), and further lengthening homology regions increases targeted gene insertion frequencies. The upper limit for the length of conventional targeting vectors is approximately 20 kb, with standard targeting vectors carrying approximately 10 kb of homologous DNA (e.g. a short and long arm of 4 and 6 kb, respectively). Bacterial artificial chromosomes (BACs) with usual insert sizes of 150-350 kb (and up to 700 kb) provide a unique tool to overcome the length constraint of conventional targeting vectors. Manipulation of these large DNA constructs requires recombinant technologies (Copeland et al., 2001), which exploits the HR system of Escherichia coli to generate large recombinant DNA constructs. Importantly, the large homology regions overcome the need for isogenic DNA, thus permitting HR in hESCs of various genetic backgrounds.

A recent report described successful BAC-based HR in hESCs at three loci, p53, ATM and HPRT, with targeting efficiencies of 21 to 27% (Song et al., 2010). The linearized BAC-targeting vectors were transduced into hESCs, both HUES9 and H9, by electroporation. Using the same targeting vectors with an alternative drug selection cassette, the authors were able to generate homozygous mutants in ATM and p53. One frequently cited drawback of BAC-based gene targeting is the difficulty in distinguishing HR from random integration. Song et al. overcame this challenge by shortening one homology arm to less than 10 kb while retaining the other at 80 to 100 kb. This allowed confirmation of HR by Southern blotting and a ligation-mediated PCR.

4.3 Zinc finger nucleases mediated HR
It has been well documented that double-stranded DNA breaks significantly enhance HR in human cells (Porteus and Carroll, 2005; Porteus et al., 2003). Such double-stranded breaks (DSB) are then repaired either by nonhomologous end joining (NHEJ), which is highly error prone thereby disrupting the targeted gene, or by homology-directed repair (HDR), where the sequence of a homologous strand of DNA (generally the intact homologous chromosome) is accurately copied. However, introduction of a single DSB in the genome is impossible with conventional restriction enzymes.

Zinc-finger nucleases (ZFN) offer a unique technology for cutting at single sites within the genome. In this approach a Zinc-finger DNA binding domain of the C2H2 class is engineered to recognize a single DNA sequence of interest and fused to an endo-nuclease, generally Fok1. Since Fok1 needs to dimerize to cleave DNA, two Zinc-finger DNA binding proteins with oppositely oriented binding sites are engineered. These two ZFNs are designed to be highly specific so that the Fok1 nuclease introduces a single DSB in the entire genome. One study demonstrated that design of the ZFN can greatly affect targeting efficiency, with rates varying between 39 to 100% for three distinct ZFN pairs designed to the OCT4 locus (Hockemeyer et al., 2009). In addition to the difficult and time-consuming nature of ZFN design and construction, it should be noted that site-specificity is difficult to control and off-target cleavage events have been observed (Radecke et al., 2010). As the ZFN introduces a single DSB, the DNA repair machinery will either introduce a deletion at the break site by NHEJ, or, if an exogenous DNA fragment is provided, a genetic modification near the DSB by HDR.
The ZFN technology has been successfully applied in multiple organisms, including Zebrafish, and in mouse and human cell culture, including hESC and hiPSCs (Davis and Stokoe, 2010; Hockemeyer et al., 2009; Lee et al., 2010a, b; Lombardo et al., 2007; Porteus and Carroll, 2005; Zou et al., 2009). A major obstacle to successfully applying this method is that two ZFNs and the homologous DNA have to be co-delivered into the cells of interest. Lombardo et al. addressed this issue by utilizing integration defective lentiviral vectors (IDLV) to co-deliver the ZFNs and the donor sequence. While this method yielded HR at the targeted site, the IL-2 receptor common γ-chain gene (IL2RG), in hESCs (Lombardo et al., 2007), such IDLVs are potentially capable of integrating randomly into the genome, thereby potentially disrupting critical genes.

An alternative virus free application of ZFNs in hESCs and hiPSCs was recently reported to target the PIG-A locus and a chromosomally integrated GFP gene (Zou et al., 2009). In this study, the donor DNA (2 kb homology arm) was transduced with or without the DNA encoding the ZFNs, into hESCs under optimized conditions that included the use of an immortalized cell line (W3R) that expresses Wnt3a, a growth factor that promotes hESC growth (Cai et al., 2007). The Amaxa Nucleofector® Technology was employed to achieve 50% transfection efficiency and high stable integration rates (10⁻⁵ cells). In the presence of the ZFNs, HR frequencies were increased 200 to 2,000 fold.

4.4 Adenoviral vector- and adeno associated viral-mediated HR.

Adenoviral vectors (AdVs) have been used extensively to transduce a broad range of cell types, including hPSCs, and are used in pre-clinical studies involving gene therapy. AdVs have been modified to remove all viral genes from the viral genome, thereby producing helper-dependent AdVs (HDAdVs) that are less cytotoxic than their wildtype counterparts. Suzuki et al. used HDAdV to successfully target the HPRT1 locus in hESCs at high efficiency (Suzuki et al., 2008).

Adeno-associated virus (AAV) is a single-stranded DNA virus, which derives its name from the fact that it is often found in cells that are simultaneously infected with adenovirus. However, in contrast to adenovirus, AAV does not elicit a host immune response or stimulate inflammation and can infect non-dividing cells. Wildtype AAV predominately integrates into one location on human chromosome 19. Like adenovirus, AAV can be engineered to remove all viral genes, including those needed for integration, and has been utilized for HR in mammalian cell culture (Hirata et al., 2002; Porteus et al., 2003; Russell and Hirata, 1998; Vasileva and Jessberger, 2005; Vasileva et al., 2006), including hESCs and hiPSCs (Khan et al., 2010).

Upon endocytosis, a wildtype AAV particle moves to the nucleus via motor proteins in the cytoplasm and nuclear tubular structures (Seisenberger et al., 2001), thereby “protecting” the single stranded genome on its journey to the nucleus. Once within the nucleus, “the linear single-stranded monomers mimic DNA damage and induce DNA-repair pathways” (Vasileva and Jessberger, 2005). In Escherichia coli it is well established that single stranded DNA triggers DNA repair via the SOS system. Consequently, it has been postulated that in eukaryotic cells, single stranded DNA would trigger a similar DNA repair process. The various methods described above involve donor DNA that is double stranded, which only efficiently triggers a DNA response pathway and homology-based repair when double stranded breaks exist in the genome. Thus, with its single stranded DNA genome, recombinant AAV may represent an ideal substrate for HR.
With a genome of 4.8 kb, AAV provides limited capacity of targeting vector design. At the same time, this limited genome size greatly simplifies cloning strategies relative to the conventional methods of BAC recombineering. Targeting vectors can be designed to replace all viral sequences except two essential palindromic inverted terminal repeats (ITR) flanking the donor DNA. In contrast to the conventional HR methods and BAC-mediated HR, left and right homology arms of 1 kilobase length have been shown to be sufficient to promote high HR rates using AAV. To permit selection of infected cells, the recombinant AAV can be constructed to contain drug selection genes under control of a constitutive promoter. HESCs can be efficiently infected with AAV serotype 2, 4 and 5 (Smith-Arica et al., 2003), and among drug selected clones HR frequencies approaching 100% have been reported (Hirata et al., 2002; Khan et al., 2010; Porteus et al., 2003; Suzuki et al., 2008; Vasileva et al., 2006).

5. Concluding remarks

Manipulation of gene expression using one or more of the various transgenesis methods described here are critical in advancing hPSC science and accelerating applications of hPSC in regenerative medicine. While many methods of transgenesis developed in other cell culture systems are available, it is clear that hPSCs exhibit unique properties and characteristics so that all methods require significant refinement and optimization. The current differences in gene transduction efficiencies among the various hPSC lines will likely evaporate as culture conditions for hPSC improve and become standardized. Recent advances and improvements in culture substrates and media formulations specifically designed for hPSCs will greatly augment and enhance the current methods to modify the genome of hPSCs. Additionally, further development of technologies for single cell isolation and purification, such as flow cytometry and cell sorting, will yield transduced and targeted cell populations in sufficiently large numbers to permit further analysis of transgenic cell lines and to enable transplantation studies.

6. References


Methodological Advances in the Culture, Manipulation and Utilization of Embryonic Stem Cells for Basic and Practical Applications


Zou, J., Maeder, M.L., Mali, P., Pruett-Miller, S.M., Thibodeau-Beganny, S., Chou, B.K.,
related gene in human induced pluripotent stem and embryonic stem cells. Cell
Stem Cell 5, 97-110.

Pluripotent stem cells have the potential to revolutionise medicine, providing treatment options for a wide range of diseases and conditions that currently lack therapies or cures. This book describes methodological advances in the culture and manipulation of embryonic stem cells that will serve to bring this promise to practice.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.