A UTF1-based selection system for stable homogeneously pluripotent human embryonic stem cell cultures

Shen Mynn Tan1, Siew Tein Wang2, Hannes Hentze2 and Peter Dröge1,*

1School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, 637551 and 2ES Cell International Pte Ltd, 11 Biopolis Way, #05-06 Helios, 138667 Singapore

Received March 2, 2007; Revised August 7, 2007; Accepted August 24, 2007

ABSTRACT

Undifferentiated transcription factor 1 (UTF1) was identified first in mouse embryonic stem cells and is also expressed in human embryonic and adult stem cells. UTF1 transcription ceases at the onset of differentiation, which clearly distinguishes it from less sensitive pluripotency markers, such as Oct4 or Nanog. We present here two transgenic hESC lines, named ZUN. Each line harbors one copy of the UTF1 promoter/enhancer driving a resistance gene and yielded highly homogeneous cultures under selection pressure, with a larger proportion of Oct4 and Sox2 positive cells. While ZUN cultures, like parental HUES8 cultures, retained the capacity to differentiate into tissues of all three germ layers using a SICD mouse teratoma model, they surprisingly exhibited an increased refractoriness to various differentiation cues in vitro. Together with its small size of only 2.4 kb for the entire cassette, these features render our selection system a powerful novel tool for many stem cell applications and human somatic cell reprogramming strategies.

INTRODUCTION

The transcriptional co-activator ‘undifferentiated transcription factor 1’ (UTF1) is expressed in both mouse and human embryonic stem cells (mESCs and hESCs, respectively) (1–3). UTF1 interacts with ATF-2 and TBP-containing complexes, and seems to be involved in cell proliferation control (4,5). UTF1 is regarded as one of the most important pluripotency marker genes defining the hESC niche (6,7). Importantly, its transcriptional regulation is very sensitive to differentiation cues and switched off faster and with a larger magnitude than either Oct4 or Sox2 during embryoid body formation from both mouse and human ESCs (8–10). A similar response is seen when mouse embryonic carcinoma (EC) and mESCs were triggered to differentiate by retinoic acid (RA) (11,12). In a comparison of hESCs cultured in medium containing either fetal calf serum or defined serum replacement, UTF1 was found to be downregulated to a larger extent in serum medium compared to Oct4 or Sox2, suggesting strongly that it is a very reliable marker for early differentiation (13). Furthermore, RNA interference (RNAi) of Oct4 transcripts in mESCs has been shown to cause downregulation of UTF1 transcripts (14), and more than 90% of all cells in undifferentiated mESC populations stain positively for both these markers (15). These two experiments further allude to the prevalence of UTF1 and Oct4 co-expression in undifferentiated ESCs.

Human embryonic stem cell cultures are always heterogeneous and usually contain a subset of less pluripotent and, hence, more differentiated cells (16–18). This is also true for mouse and medaka ESCs (19,20). It was shown that the presence of minor populations of differentiated cells compromises the pluripotency of entire stem cell cultures (21). These features represent a severe problem for many hESC applications which can be addressed either via fluorescence-activated cell sorting (FACS) in order to isolate pluripotent hESCs or by antibiotic selection as a means to eliminate differentiating cells.

FACS strategies generally utilize transgenes, such as EGFP, under the control of genetic elements of a pluripotency marker. For example, mouse Rex1-EGFP and Oct4-EGFP combinations have been used in hESCs to screen for undifferentiated cells via flow cytometry (17,22,23). A recent improvement in semi-automated FACS and clonal recovery utilized a mouse Oct4-EGFP construct to sort for pluripotent hESCs (24). However, this procedure has to be repeated routinely in order to ensure more homogeneous hESC cultures over longer periods of time.

Antibiotic selection has the advantage of being less intrusive to hESCs. Expression of the neomycin resistance (Neo) gene and G418 antibiotic selection, along with EGFP expression, have been shown to be compatible with
hESC pluripotency (23,25). In the mouse, the Oct4-Neo combination has been used to isolate ES and EC cell lines from embryos. (19,21,26). Thus far, only one report utilized the twin markers of EGFP and Neo driven by an Oct4 promoter for the maintenance of pluripotency in hESCs (22). However, the ability to select against partially differentiated cells as well as the derivation of more homogeneous pluripotent hESC cultures was not established in these studies.

Here we demonstrate that UTF1 is an exceptionally sensitive pluripotency marker. We found that MI, a genetic element comprising a conserved octamer sequence important for Nanog expression (27), is also present and functional in the UTF1 3' enhancer, but not in the UTR of Oct4 or Sox2. As a representative example, we present two stable ZUN hESC lines in which Neo is under the control of UTF1 promoter/enhancer. Application of G418 selection results in ablation of differentiated hESCs and homogeneously pluripotent ZUN cultures, containing a significantly larger number of Oct4 and Sox2 positive cells compared to parental cultures. While ZUN cultures retain the capacity to differentiate into progenitors of all three germ layers in vivo, they surprisingly exhibited an increased refractoriness to differentiation cues under various culture conditions.

**MATERIALS AND METHODS**

**Plasmids**

The full-length human UTF1 gene was amplified from HeLa genomic DNA using primers UTF1-F (CCGGA ATTCAGGCACGGACCCCTTTA) and UTF1-R (TGCTCTAGAGGGTTCAGCACTTCTGCCTC) via PCR. The product was cleaved with EcoRI and XbaI, and ligated into phagemid pTZ-18R. The UTF1 coding sequence was subsequently replaced with that of EGFP or Neo in the following way. KpnI and NcoI restriction sites were introduced between the end of the UTF1 promoter and before the ATG codon by amplifying the promoter region of pTZ-UTF1 via PCR with primers UTF1-F and UTF1-MR (CGGGGTACCGGGGCTGGGGCTGCGGC) and UTF1-BR (AGAATACTCAGCTGCCCGGCGCGG). The resulting fragment was cleaved with EcoRI and NeoI, and ligated to a cleaved pCMV-EGFP which replaces the CMV promoter. The plasmid was then digested with EcoRI and BsrGI to produce an (EcoRI) UTF1 promoter-EGFP (BsrGI) sticky-ended fragment. pTZ-UTF1 was used as a PCR template to amplify a segment of the UTF1 3' UTR region and inserting a BsrGI RE site near the end of UTF1 transcription unit using primers UTF1-MF (AGCTGTACAAGTGAGTCCTCTGCCTC) and UTF1-BR (AGAATACTCAGCTGCCCGGCGCGG). This produced another BsrGI-BamHI sticky-ended 1 kb fragment. pTZ-UTF1 was then digested with EcoRI and BamHI, and this 3.2 kb fragment was ligated to the above two 1 kb fragments, yielding pTZ-UTF1-EGFP. The Neo gene was amplified with flanking KpnI and BsrGI RE sites via PCR with pPGK-Neo as template and Neo-F (CTGGGTACCGGCGGCATGATGGA) and Neo-R (ACTGTGTA CATCAGAGAAGCTCGTCAAGAA) primers.

The pTZ-UTF1-EGFP backbone was digested with KpnI and BsrGI to remove EGFP, and ligated with the cleaved Neo fragment to yield pTZ-UTF1-Neo. The promoter-less version of pTZ-UTF1-EGFP was generated using EcoRI and KpnI to remove the UTF1 promoter. The ends were blunted using Klenow Fragment (NEB) followed by self-ligation. The enhancer-less version of pTZ-UTF1-EGFP was generated using PstI to remove the UTF1 enhancer followed by self-ligation.

pTZ-UTF1-EGFP was used as a template to introduce the mutant M1 sequence (GTAGTGGT; original sequence GTCTGGGTT; (27). Flanking arms 200 bp from M1 were amplified using two primer pairs, UTF1-M1F (GACCGGATGGTACTCAGGTAGCATG) and UTF1-M1mR (TATGGCCCAACACTACAGAGCCAC) (near the 5' XcmI RE site), and UTF1-M1mF (GTGGCTCTGTAGTGGTGCGGCTAG) and UTF1-M1R (CATTTTATCGTACTGCGG) (near the 3' NheI RE site). These were combined via assembly PCR using UTF1-M1F and UTF1-M1R as primers. The resulting 370 bp fragment was digested with XcmI and NheI, and ligated to a similarly cleaved pTZ-UTF1-EGFP in order to replace the original M1 sequence.

Plasmids were amplified in Escherichia coli strain DH5α. We used standard PCR conditions with Pfu polymerase (Promega). Restriction enzymes were from New England Biolabs.

**Cell culture, transfection and differentiation**

F3 human foreskin fibroblasts (28), A549 and HeLa carcinoma cell lines were cultured according to standard procedures. hESC line hES2 (karyotype: 46XX) was obtained from ES Cell International Pte Ltd. HUES8 (karyotype: 46XY) and HUES9 (karyotype: 46XX, inv9) was provided by the Harvard Stem Cell Institute. hESCs were cultured on G418-resistant MEFs (PMEF-N, Chemicon), as previously described (29,30).

F3, A549 and HeLa cells were transfected with 10 μg of reporter plasmid via electroporation as described previously (31). hESCs were transfected with 2 μg of reporter plasmid and Effectene (Qiagen) as described previously (29).

Differentiation medium was prepared by removing basic fibroblast growth factor and exchanging knockout serum replacement with fetal bovine serum (all three from Gibco). Spontaneous differentiation was conducted on matrigel (BD Biosciences) or gelatin surfaces without MEFs. Induced differentiation on MEFs was conducted either with 10 μM (RA) or 1% (DMSO) (Sigma) which was added to differentiation medium (32). Embryoid body formation was induced by plating hESCs onto low-cell binding plates (Nunc).

**SCID mouse teratoma assay**

To test the differentiation potential of hESC lines in vivo, serially passaged hESCs were manually harvested and injected as clumps with an approximate cell dose of 2–4 × 10⁶ cells in a volume of 50 μl into the quadriceps of the right hind limb of a male SCID mouse (3 mice/hESC line). Mice were maintained under controlled conditions...
in accordance with the National Institutes of Health (NIH) and National Advisory Committee for Laboratory Animal Research (NACLR) guidelines, and with approval of the Biopolis Institutional Animal Care and Use Committee (Biopolis IACUC approval 050008, National University of Singapore Institutional Review Board 05-020). Teratoma formation was monitored visually using a simple grading system that was confirmed by caliper measurements (grade 0—no detectible enlargement, grade 1—enlargement just detectible, grade 2—obvious enlargement, grade 3—enlargement impedes locomotion). When teratomas reached grade 3 (after 6–8 weeks), teratomas weighing 1–2 g were excised, fixed (Bouin’s solution), paraffin-embedded, sectioned (after 6–8 weeks), teratomas weighing 1–2 g were excised, and histologically analyzed following staining with hematoxylin and eosin. Additionally, teratomas were reproducibly led to a significantly smaller number of M1 sequence. The results demonstrate hESC-specific expression of the UTF1 promoter and enhancer regions were essential for EGFP expression; a response that was not observed for Nanog and Oct4 (Figure 1A). We next gathered the following three UTF1 transgenic constructs: pTZ-UTF1, pTZ-UTF1-EGFP and pTZ-UTF1-Neo (Figure 1B). In addition, three variants of pTZ-UTF1-EGFP were designed: one without the minimal 5’ promoter, one lacking the 3’ enhancer, and one containing an altered M1 sequence. The M1 octamer sequence was found to be conserved in the Nanog promoter of many species (27), and we identified it recently upstream of the Oct4/Sox2 binding site within the 3’ UTF1 enhancer.

The results demonstrate hESC-specific expression of the full-length pTZ-UTF1-EGFP after transfection into various human cell lines and hESC line HUES8 (Figure 1C). We next introduced the three pTZ-UTF1-EGFP versions into HUES8 cells and showed that the 5’ promoter and 3’ enhancer regions were essential for EGFP expression (Figure 1C). In addition, the altered M1 element reproducibly led to a significantly smaller number of EGFP expressing hESCs, which is similar in magnitude to its reported effect on Nanog transcription (27). Together, these data confirm that UTF1 transcription is a very sensitive pluripotency marker specific for hESCs and suggests that it’s regulation in hESCs involves M1-binding factors in addition to Oct4 and Sox2.

**RESULTS**

**UTF1 is a sensitive pluripotency marker for hESCs**

We evaluated first the sensitivity of UTF1 expression with respect to induced differentiation using quantitative real-time PCR (qRT-PCR). Various hESC lines were triggered to differentiate via retinoic acid (RA) or dimethyl sulfoxide (DMSO) treatment. The results show that this almost completely abolished UTF1 expression; a response that was not observed for Nanog and Oct4 (Figure 1A).

We next gathered the following three UTF1 transgenic constructs: pTZ-UTF1, pTZ-UTF1-EGFP and pTZ-UTF1-Neo (Figure 1B). In addition, three variants of pTZ-UTF1-EGFP were designed: one without the minimal 5’ promoter, one lacking the 3’ enhancer, and one containing an altered M1 sequence. The M1 octamer sequence was found to be conserved in the Nanog promoter of many species (27), and we identified it recently upstream of the Oct4/Sox2 binding site within the 3’ UTF1 enhancer.

The results demonstrate hESC-specific expression of the full-length pTZ-UTF1-EGFP after transfection into various human cell lines and hESC line HUES8 (Figure 1C). We next introduced the three pTZ-UTF1-EGFP versions into HUES8 cells and showed that the 5’ promoter and 3’ enhancer regions were essential for EGFP expression (Figure 1C). In addition, the altered M1 element reproducibly led to a significantly smaller number of EGFP expressing hESCs, which is similar in magnitude to its reported effect on Nanog transcription (27). Together, these data confirm that UTF1 transcription is a very sensitive pluripotency marker specific for hESCs and suggests that its regulation in hESCs involves M1-binding factors in addition to Oct4 and Sox2.
Characterization of transgenic pTZ-UTF1-Neo hESC lines

We established stable pTZ-UTF1-Neo cell lines from HUES8 cells via random genomic transgene insertion. Two single-copy lines with normal karyotype, dubbed ZUN1 and ZUN2, were used in subsequent experiments (Supplementary Figure S1). Careful characterization of both lines revealed that they have similar proliferation rates (Supplementary Figure S2), exhibit comparable alkaline phosphatase (AP) staining patterns (Supplementary Figure S3), and express Nanog at a similar level as parental HUES8 cells (Figure 2A). However, an elevated level of Oct4, Sox2 and UTF1 in ZUN cultures was reproducibly detected by qRT-PCR (Figure 2B). This correlation of expression supports the notion that the control of UTF1 involves factors Oct4 and Sox2, but most likely not Nanog (35).

Figure 1. UTF1 expression in hESCs. (A) qRT-PCR Expression analysis of RA- (for 12 days) and DMSO-induced (for 7 days) differentiation in hESCs, normalized to that of respective undifferentiated hESC lines, with two biological replicates for each sample. (B) Schematic drawing of vectors. Full-length UTF1 was cloned into phagemid pTZ-18R yielding pTZ-UTF1. Its coding region was subsequently replaced with enhanced green fluorescent protein (EGFP) or neomycin (Neo), yielding pTZ-UTF1-EGFP and pTZ-UTF1-Neo, respectively. Red text indicates alterations in pTZ-UTF1-EGFP. (C) Flow cytometry of UTF1-driven EGFP expression in various cell lines following transient transfection of pTZ-UTF1-EGFP. Transfection efficiencies were normalized to those determined in parallel using pCMV-EGFP and are the average of three biological replicates. See Supplementary Figure S7 for representative dot plots.

Figure 2. Characterization of ZUN hESC lines. (A) Expression analysis of ZUN hESCs. qRT-PCR comparison of pluripotency markers across the three hESC cultures normalized to HUES8 samples, performed with biological triplicates. (B) Oct4 and Sox2 expression analysis of ZUN hESCs. Percentage of hESCs expressing Oct4 and Sox2. ZUN cultures consist of significantly more Oct4- and Sox2-expressing cells compared to HUES8, based on the two-tailed paired Student’s t-test (P-values < 0.05). (C) Surface marker analysis of ZUN hESCs. Percentage of positively stained hESCs for TRA-1-60 and TRA-1-81. Both ZUN lines consist of significantly more positively stained cells compared to parental HUES8, based on the two-tailed paired Student’s t-test (P-values < 0.05). (D) qRT-PCR analysis of removal and restoration of G418 selection pressure on ZUN1 and ZUN2 cultures. ZUN lines were cultured in the presence of G418 for 90 days (+G418); G418 removed for 30 days (−G418) or 60 days (−/−G418); G418 removed for 30 days and re-introduced for 30 days (−/+G418). Expression levels were normalized to +G418 values and presented as an average of three experimental repeats.
Flow cytometry revealed that the observed higher level of global Oct4, Sox2 and UTF1 expression was due to the presence of a larger fraction of cells that express these markers. A statistically significant higher percentage of ZUN cells stained positive for Oct4 and Sox2 compared to HUES8 cultures (Figure 2B). We could also demonstrate that ZUN cultures contained more cells which stain positive for surface markers TRA-1-60 and TRA-1-81 (Figure 2C). These results indicate that HUES8 cultures contained more hESCs with a low expression level of pluripotency markers. In contrast, such cells have been ablated in ZUN cultures via G418 selection.

We next analyzed the effect of sequential removal and restoration of G418 selection on both ZUN lines and employed qRT-PCR of pluripotency markers Nanog, Oct4, Sox2 and UTF1 for the detection of the onset of differentiation (9). We also included the transgene UTF1-Neo in this analysis because its regulation should be linked to that of the endogenous UTF1. The results show that spontaneous differentiation can be detected in cultures during 60 days of G418 withdrawal. This is indicated by the progressive reduction in signals obtained from all markers, with the exception of Sox2 (Figure 2D). However, upon re-application of antibiotic selection, original expression levels could be restored in these cultures. Using a transgenic hESC line in which Neo is expressed from the elongation factor 1 alpha (EF-1α) promoter as a control, we confirmed that G418 withdrawal has no effect on the expression of this set of markers which clearly links this effect to the presence of the UTF1-Neo transgene in hESCs (Supplementary Figure S4). Together, these data show that ZUN cultures are more homogeneously pluripotent than the parental cultures and that G418 addition and subsequent ablation of differentiated ZUN cells is required to sustain homogeneous cultures over longer periods of time.

ZUN cultures are refractory to various differentiation cues

We demonstrated above that ZUN cultures are more homogeneously pluripotent than the parental cultures. We were interested, therefore, in investigating the differentiation behavior of ZUN cells. First, we induced differentiation by plating them on matrigel or gelatin without supporting mouse embryonic fibroblasts (MEFs) and in the absence of G418. In the matrigel protocol, HUES8 cells quickly initiated differentiation as indicated by the 5-fold downregulation of our panel of markers (Figure 3A). However, both ZUN lines maintained significantly higher overall expression levels even after 12 days of culture. Similarly, the gelatin protocol induced a 2-fold downregulation of Oct4 and UTF1 in HUES8, but not in ZUN populations. We observed further that under these conditions, ZUN cultures have a larger number of colonies consisting of cells which display a more compact morphology that stain stronger for alkaline phosphatase (AP), and a reduced occurrence of differentiating cells at colony edges. Upon subsequent addition of G418 for 2 days, only pluripotent cells that stained for AP survived in ZUN cultures, demonstrating again that differentiating ZUN cells are rapidly and efficiently ablated by G418 selection (Figure 3B).

We analyzed next the effects of RA or DMSO treatment on ZUN cultures in the absence of G418 selection. After 12 and 7 days, respectively, qRT-PCR showed that ZUN cultures exhibited a lower rate of differentiation in comparison to HUES8 cultures (Figure 4A). Notably, UTF1 and UTF1-Neo expression diminished by the largest magnitude among the pluripotency markers. Certain differentiation markers also indicated a slower onset of differentiation in ZUN cultures. HUES8 cultures showed a greater (2- to 10-fold) upregulation of Sox9, Nestin and ID2 due to RA treatment, and of HAND1, IGF2 and Nestin after exposure to DMSO (Figure 4B). The expression of surface markers of pluripotency,
TRA-1-60 and TRA-1-81, complements these transcript analyses, showing that HUES8 cells exhibited a more muted staining pattern compared to ZUN colonies (Figure 4C). These results confirm our conclusion that under various conditions, ZUN cultures exhibit a more robust phenotype against the onset of differentiation than parental cultures.

**ZUN cells retain the capacity to differentiate into progenitors of all three germ layers in vitro and in vivo**

The observed phenotype of increased robustness against the onset of differentiation led us to investigate ZUN cultures under long term differentiation conditions. The results show that ZUN and HUES8 cells cultured under RA and DMSO differentiation protocols for 30 days show similar morphologies and lack of AP staining (Supplementary Figure S5). These cells were completely ablated upon addition of G418, as were those in equally long-term gelatin and matrigel cultures (data not shown). In addition, 21 days of embryoid body formation in suspension cultures caused a significant upregulation of differentiation markers from all three germ layers (Figure 4D). Both ZUN lines and HUES8 show very high expression of HAND1, IGF2 and ID2 as indicated by the large fold increase relative to their undifferentiated counterparts. They also show relatively high expression of α1 anti-trypsin (α1AT) and Sox9. The morphological characteristics and growth rates of HUES8 and ZUN cell-derived EBs were indistinguishable (Supplementary Figure S6). Furthermore, we showed by a teratoma assay that the parental HUES8 line as well as ZUN1 and
ZUN2 formed tissue of all three germ layers (i.e. ectodermal structures, mesoderm such as cartilage and bone, as well as mesodermal glandular structures) when transplanted into immune-deficient SCID mice (Figure 5). We conclude, therefore, that the robustness which our ZUN cultures displayed against the onset of differentiation induced by various protocols does not compromise their potential to differentiate, like parental cultures, into all three germ layers in vitro and in vivo.

**DISCUSSION**

Expression of *UTF1* is controlled by a 5′ TATA-less promoter consisting of four GC boxes. The 3′ enhancer element harbors a twin octamer sequence where the synergistic binding of Oct4 and Sox2 is essential for *UTF1* expression in both mouse and human ESCs (8,35). Sp1-like transcription factors that bind to GC boxes present in the *UTF1* promoter are most likely involved in the regulation of *UTF1*, since they are regarded as regulators of embryonic development in vertebrates (36) and shown to be involved in the transcriptional control of both Oct4 and Nanog (37,38).

We showed here that *UTF1* is indeed expressed in hESCs, but not in human primary fibroblasts and carcinoma cell lines. Further, both 5′ promoter and 3′ enhancer are needed for expression. A novel finding is that an octamer sequence (M1) which is conserved in the Nanog promoter of many species (27) is also present upstream of the analogous human Oct4/Sox2 cognate sequence in the 3′ *UTF1* enhancer. The fact that M1 mutations resulted in equally muted expression levels of *UTF1* and Nanog, strongly suggests that similar M1-binding factor(s) are involved in the control of both pluripotency marker genes. Our results also confirm previous reports showing that endogenous *UTF1* is downregulated faster than Oct4 or Nanog at the onset of differentiation to nearly undetectable levels. In addition, detailed analysis of SymAtlas data from the Genomics Institute of the Novartis Research Foundation showed that both *UTF1* and Nanog expression are greatly diminished at day 8.5 of mouse embryo formation, unlike Oct4 or Sox2 (http://symatlas.gnf.org). Together, these results firmly establish *UTF1* as a sensitive and reliable pluripotency marker for hESCs.

Recent reports revealed that *UTF1* is also an important marker expressed in other stem cell types. For example, spermatogonial stem cells isolated from adult mouse testis were shown to exhibit features reminiscent of pluripotent ESCs. These cells express Oct4, Nanog, Rex1 and *UTF1* and, interestingly, also downregulate *UTF1* significantly faster than these other markers upon embryoid body formation (39). Further, pluripotent stem cells derived from re-programmed primary mouse fibroblasts through co-expression of transgenes encoding Oct4, Sox2, c-Myc and Klf4 also express *UTF1*. It is noteworthy that the expression level of *UTF1* in pluripotent stem cells derived from different sources is always more tightly linked with
that of the key marker Oct4, and not with other markers like Cripto or Nanog (40). These data, in conjunction with results presented here, indicate that our UTF1-based selectivity tool should be widely applicable to select for pluripotent human stem cell lines derived from different sources.

We demonstrated that the control of both endogenous and exogenous UTF1 in ZUN cells is linked, as expected, and that UTF1-driven Neo expression in hESCs in conjunction with G418 selection can be used to efficiently ablate differentiating cells in standard culture conditions. This resulted in ZUN cultures which display a higher global expression level of key pluripotency markers. We showed that this elevated level is most likely due to an increase in the fraction of Oct4 and Sox2 positive hESCs. Furthermore, the number of cells staining positive for surface markers TRA-1-60 and TRA-1-81 was also elevated which, if taken together, indicates that ZUN cultures are more homogeneously pluripotent than the parental cultures (Figure 2A–C). Upon removal of selection pressure from ZUN cultures, the overall expression level of these markers declined, which indicates loss of pluripotency and, consequently, more heterogeneous hESC cultures. Interestingly, the absence of selection pressure for 60 days did not cause widespread silencing of the exogenous UTF1-Neo. This enabled us to add G418 only periodically in order to select against differentiating hESCs in ZUN cultures. Hence, the UTF1 promoter/enhancer combination could be a generally applicable control element for sustained expression of transgenes in pluripotent hESCs, even in the absence of selection.

We showed that the generation of ZUN lines harnessed the sensitivity of UTF1 as a pluripotency marker which resulted in rapid and efficient ablation of differentiating cells. An unexpected finding was the refractoriness of ZUN cultures to the onset of differentiation. We attribute this phenotype to the increased fraction of pluripotent hESCs in homogeneous ZUN cultures compared to typical, more heterogeneous hESC cultures. Given the propensity for heterogeneous cultures to spontaneously differentiate (21), the increased stability of human stem cell cultures with respect to the uniformity in pluripotency combined with efficient ablation of differentiating cells offered by the ZUN system could significantly improve scale growth for many future hESC applications. In addition, ZUN cultures show a higher expression of pluripotency markers when compared to HUES8 cultures following thawing after standard slow cooling cryopreservation procedures (data not shown), thereby counteracting a reported negative effect associated with cryopreservation (41).

An intrinsic advantage of the UTF1-based selection system over other pluripotency marker-based strategies is its enhanced sensitivity towards many differentiation pathways. Although the well-characterized pluripotency marker Oct4 has been a component in strategies for both mouse and human ESCs (21,22), the use of Oct4 in this context can be problematic. Apart from being less sensitive to the onset of differentiation, its expression level is also very tightly controlled. A less than 2-fold increase in Oct4 expression alone leads to differentiation into primitive endodermal and mesodermal lineages in mouse ESCs (42). This implies that the Oct4-Neo system is potentially less sensitive to these differentiation pathways, and, thus, will not be able to efficiently ablate cells that are undergoing differentiation into these lineages. In addition, the <2 kb length of the entire human UTF1 promoter/enhancer element represents a significant advantage for genetic manipulation of hESCs, or other human cell types, if one compares this to an unwieldy 8 kb of the Oct4 control element.

Our ZUN system has a large range of potential applications. In its current context as a stable transgene in hESCs, it can be used for the maintenance of homogeneously pluripotent cultures. In combination with gene targeting technologies, this application can be upscaled to meet the larger cell quantities ultimately needed for stem cell-based therapies. ZUN lines can also be employed to screen and characterize synthetic factors, media components or surface matrices that permit maintenance of pluripotency, and it can be used to test 3D scaffolds or aid in the derivation of new hESC lines.

As indicated above, the ZUN system shows potential in selection protocols in order to isolate adult stem cells from primary cultures. It can also be used in hESCs in conjunction with cell differentiation strategies as a component of positive-negative selection. The system can also function as a tool in somatic cell reprogramming experiments to identify successfully reprogrammed pluripotent cells. It can also be used to improve future human somatic cell nuclear transfer protocols. With our first demonstration of its utility here and its many possible applications, we envision that our ZUN system will become a valuable selectivity tool for many facets of stem cell research.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank R.J. Zhou for help with the Southern blot analysis, C.-T. Lee for the construction of pTZ-UTF1-EGFP(M1) and K. Neef for hESC line EPN14. Special thanks go to C.A. Davey for critical comments on the manuscript. Funding to pay the Open Access publication charges for this article was provided by NTU special investigator grant.

Conflict of interest statement. None declared.

REFERENCES

1. Ginis,I., Luo,Y., Miura,T., Thies,S., Brandenberger,S., Gerecht-Nir,S., Amit,M., Hokc,A., Carpenter,M.K. et al. (2003) Differences between human and mouse embryonic stem cells. Dev. Biol., 269, 360–380.
2. Brimble,S.N., Zeng,X., Weiler,D.A., Luo,Y., Liu,Y., Lyons,I.G., Freed,W.J., Robins,A.J., Rao,M.S. et al. (2004) Karyotypic stability, genotyping, differentiation, feeder-free maintenance, and gene expression sampling in three human embryonic stem cell lines derived prior to August 9, 2001. Stem Cells Dev., 13, 585–597.
3. Sun, Y., Li, H., Liu, Y., Shin, S., Mattson, M.P., Rao, M.S. and Zhan, M. (2007) Cross-species transcriptional profiles establish a functional portrait of embryonic stem cells. *Genomics*, **89**, 22–35.

4. Fukushima, A., Okuda, A., Nishimoto, M., Seki, N., Hori, T.A. and Muramatsu, M. (1998) Characterization of functional domains of an embryonic stem cell coactivator UTF1 which are conserved and essential for potentiation of ATF-2 activity. *J. Biol. Chem.*, **273**, 25840–25849.

5. Nishimoto, M., Miyagi, S., Yamagishi, T., Sakaguchi, T., Niwa, H., Muramatsu, M. and Okuda, A. (2005) Oct-3/4 maintains the proliferative embryonic stem cell state via specific binding to a variant octamer sequence in the regulatory region of the UTF1 locus. *Mol. Cell. Biol.*, **25**, 5084–5094.

6. Bhattacharya, B., Miura, T., Brandenberger, R., Meijido, J., Loo, Y., Yang, A.X., Joshi, B.H., Ginis, I., Thies, R.S. et al. (2004) Gene expression in human embryonic stem cell lines: unique molecular signature. *Blood*, **103**, 2956–2964.

7. Trounson, A. (2006) The production and directed differentiation of human embryonic stem cells. *Endocr. Rev.*, **27**, 208–219.

8. Nishimoto, M., Fukushima, A., Okuda, A. and Muramatsu, M. (1999) The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. *Mol. Cell. Biol.*, **19**, 5453–5465.

9. Cai, J., Chen, J., Liu, Y., Miura, T., Luo, Y., Loring, J.F., Freed, W., Jiang, H., Xiao, S., Zeng, X. (2006) Assessing self-renewal and differentiation in human embryonic stem cell lines. *Stem Cells*, **24**, 516–520.

10. Wei, C.L., Miura, T., Robson, P., Lim, S.K., Xu, Q.X., Lee, M.Y., Gupta, S., Stanton, L., Luo, Y. et al. (2005) Transcriptome profiling of human and murine ESCs identifies divergent paths required to maintain the stem cell state. *Stem Cells*, **23**, 166–185.

11. Eifert, C., Sangster-Guity, N., Yu, L.M., Yi, S., Chang, J., Tine, J.A. and McCormick, P.J. (2006) Global gene expression profiles associated with retinoic acid-induced differentiation of embryonal carcinoma cells. *Mol. Reprod. Dev.*, **73**, 796–824.

12. Ivanova, N., Dobrin, R., Lu, R., Kostenchok, L., Levorse, J., DeCote,C., Schaffer, X., Sun, Y. and Lemschka, I.R. (2006) Dissecting self-renewal in stem cells with RNA interference. *Nature*, **442**, 533–538.

13. Skottman, H., Stromberg, A.-M., Matilainen, E., Inzunza, J., Hovatta, O. and Lahesmaa, R. (2005) Unique gene expression signature by human embryonic stem cells cultured under serum-free conditions correlates with their enhanced and prolonged growth in an undifferentiated stage. *Stem Cells*, **24**, 151–167.

14. Sen, G., Wehrman, T.S., Myers, J.W. and Blau, H.M. (2004) Restriction enzyme-generated siRNA (REGS) vectors and libraries. *Nat. Genet.*, **36**, 183–189.

15. Hoof, D.V., Passier, R., Oostwaard, D.W.-V., Pinkse, M.W.H., Driever, A., Passier, R. and McWhir, J. (2000) Effects of eight growth factors on the proliferation of murine embryonic stem and embryonic germ cells by selective ablation. *Transgenic Res.*, **24**, 33–44.

16. Schuldiner, M., Yanuka, O., Itskovitz-Eldor, J., Melton, D.A. and Benvenisty, N. (2000) Establishment of human embryonic stem cell lines from the medaka fish, Oryzias latipes. *Mech. Dev.*, **119**, 821–830.

17. Gupta, S., Stanton, L., Luo, Y. et al. (2006) Characterization of functional domains of an embryonic stem cell coactivator. *Mol. Cell. Biol.*, **26**, 183–189.

18. Patel, M.R., Gupta, S., Loring, J.F., Noldé, J., Wolf, F., Li, M. et al. (2003) Discovery of the transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, **122**, 947–956.

19. Wu, D.Y. and Yao, Z. (2006) Functional analysis of two Sp1/Sp3 binding sites in murine Nanog promoter. *J. Cell Biochem.*, **97**, 11307–11312.

20. Wang, S., Li, J., Zeng, X. et al. (2006) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.*, **34**, 372–376.