Accepted Manuscript

Transgenic human ES and iPS reporter cell lines for identification and selection of pluripotent stem cells *in vitro*

Dmitry A. Ovchinnikov, Drew M. Titmarsh, Patrick R.J. Fortuna, Alejandro Hidalgo, Samah Alharbi, Deanne J. Whitworth, Justin J. Cooper-White, Ernst J. Wolvetang

PII: S1873-5061(14)00065-8
DOI: doi: 10.1016/j.scr.2014.05.006
Reference: SCR 443

To appear in: *Stem Cell Research*

Received date: 25 November 2013
Revised date: 22 May 2014
Accepted date: 31 May 2014

Please cite this article as: Ovchinnikov, Dmitry A., Titmarsh, Drew M., Fortuna, Patrick R.J., Hidalgo, Alejandro, Alharbi, Samah, Whitworth, Deanne J., Cooper-White, Justin J., Wolvetang, Ernst J., Transgenic human ES and iPS reporter cell lines for identification and selection of pluripotent stem cells *in vitro*, *Stem Cell Research* (2014), doi: 10.1016/j.scr.2014.05.006

Transgenic human ES and iPS reporter cell lines for identification and selection of pluripotent stem cells *in vitro*.

Dmitry A. Ovchinnikov[1], Drew M. Titmarsh[1], Patrick R. J.Fortuna[1], Alejandro Hidalgo[1], Samah Alharbi[1], Deanne J. Whitworth[2], Justin J. Cooper-White[1], Ernst J. Wolvetang[1,*].

[1]Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, Brisbane, QLD 4072, Australia; [2] School of Veterinary Sciences, The University of Queensland, Gatton, QLD 4343, Australia

***Correspondence should be addressed to Assoc. Prof. Ernst Wolvetang**

Australian Institute for Bioengineering and Nanotechnology

Corner College and Cooper Roads (Bldg 75),

University of Queensland, St Lucia, QLD 4072, Australia

E-mail address: e.wolvetang@uq.edu.au

Tel:+61-7-33463894

Running title:

Selectable stable pluripotency reporter hPSC lines in pluripotency and differentiation

Highlights:

- stable selectable pluripotency reporter lines in hES and transgene-free iPS cells
- more sensitive pluripotency readout than conventional surface markers
- pluripotency enrichment facilitates directed differentiation protocols
- reveals heterogeneity in conventional embryoid body assays

Abstract:

Optimization of pluripotent stem cell expansion and differentiation is facilitated by biological tools that permit non-invasive and dynamic monitoring of pluripotency, and the ability to select for an undifferentiated input cell population. Here we report on the generation and characterization of clonal human embryonic stem (HES3, H9) and human induced pluripotent stem cell lines (UQEW01i-epifibC11) that have been stably modified with an artificial EOS(C3+) promoter driving expression of EGFP and puromycin resistance-conferring proteins. We show that EGFP expression faithfully reports on the pluripotency status of the cells in these lines, and that antibiotic selection allows for an efficient elimination of differentiated cells from the cultures. We demonstrate that the extinction of the expression of the pluripotency reporter during differentiation closely correlates with the decrease in expression of conventional pluripotency markers, such as OCT4 (POU5F1), TRA-1-60 and SSEA4 when screening across conditions with various levels of pluripotency-maintaining or differentiation-inducing signals. We further illustrate the utility of these lines for real-time monitoring of pluripotency in embryoid bodies and microfluidic bioreactors.

1. Introduction:

Human embryonic stem (ES) and induced pluripotent stem (iPS) cells are expected to revolutionize regenerative medicine and are potentially powerful tools for therapeutic drug screening (Inoue, H. and Yamanaka, S., 2011; Rowntree, R.K. and McNeish, J.D., 2010). Despite recent rapid advances in our ability to expand these cells as pluripotent cells (Ungrin, M.D. et al., 2008) and direct their differentiation into a range of desired cell lineages (Keller, G., 2005), much is to be learned about the extrinsic and endogenous factors produced that control these processes, and how these are modulated by cell-cell contact signalling and other features of the microenvironment. The ability to genetically manipulate mouse ES cells and the generation of mouse models with constitutive or inducible tissue-specific, cell-specific or gene-specific fluorescent reporters has greatly accelerated our understanding of pluripotent stem cell biology. However, in the case of human pluripotent cells the development of such tools has been much slower, mainly because of the inefficient nature of homologous recombination in human pluripotent cells, the difficulty in obtaining clonal populations of cells, and their inherent genetic and epigenetic instability in culture. The derivation of transgenic, clonal, and karyotypically normal human pluripotent stem cell lines is therefore a non-trivial endeavour. Furthermore, there is the obvious inability to perform human blastocyst injection of genetically-tagged ES cells, an assay that is widely used in mice to exemplify the efficacy of ES and iPS cell reporters. Given the intrinsic differences between mouse and human ES cells in terms of molecular pathways maintaining pluripotency and inducing lineage-specific differentiation, and the need to isolate pure populations of lineage-specific progenitors (or the identification of specific cell surface markers thereof), a number of laboratories have undertaken the arduous task of genetically tagging promoters and genes of interest using either classical homologous recombination or zinc finger/TALEN-based genome editing techniques (Liu, Y. et al., 2011; Zou, J. et al., 2009; Zwaka, T.P. and Thomson, J.A., 2003). It is, however, becoming increasingly clear that human embryonic stem cell lines differ intrinsically in their ability to grow and respond to pluripotent cell culture conditions and also possess different biases in their propensity to undergo differentiation into specific cell types (Lepski, G., 2012). This is most evident when comparing human iPS cell lines derived from different cell types or generated through different methods (Alvarez, C.V. et al., 2012). Understanding and harnessing these different propensities of human pluripotent cell lines, properties most likely influenced by both genetic background and epigenetic parameters, is one of the major current challenges in the stem cell field (Tobin, S.C. and Kim, K., 2012).

With the recent advent of iPS cells and their wide potential in the field of regenerative medicine, the propensity of differentiated, iPS-derived cells to revert back to a pluripotent phenotype is particularly topical (Pera, M.F., 2011). Hence, the ability to identify, and ultimately eliminate, the cells that retain these undesired pluripotent properties after directed differentiation will be essential for regenerative medicine. One of the most viable and attractive approaches to this end is the genetic tagging of cells which, in turn, requires the identification of the best promoters capable of reporting on such cells.

There is thus a clear need for tools that allow for the rapid and facile delivery of gene reporters to human pluripotent stem cells that faithfully report on pluripotency and differentiation. Lentiviral delivery of fluorescent promoter reporters is clearly an attractive solution, and a family of selectable

fluorescent reporter lentiviruses has been described by Hotta *et al.* (Hotta, A. et al., 2009a; Hotta, A. et al., 2009b). Various promoters have been trialled for their ability to drive expression specifically in pluripotent cells, and the artificial *OCT4*/*POU5F1* promoter sequences-based promoter driving expression of both EGFP and puromycin N-acetyltransferase (conferring puromycin resistance), used in this study, appears to provide the highest efficiency and specificity in human pluripotent ES (hES and iPS) cells (Hotta, A. et al., 2009a). Here we report on the generation and verification of the pluripotent stem cell lines with the lentivirally-delivered EOS(C3+)-EGFP-IRES-PURO transgene that both reports on pluripotency and, of key importance, allows for the selection of pluripotent stem cells.

2. Materials and Methods

2.1 *Pluripotent and differentiated cell culture conditions and cell lines*

hES and iPS cells were cultured in feeder-free conditions in xeno-free defined medium, mTeSR™1 (StemCell Technologies, Vancouver, BC) or NutriStem (Stemgent) on Matrigel™ (BD Biosciences at 1/100 dilution) coated tissue culture plastic surfaces. NutriStem medium was supplemented with 10 ng/mL FGF2 (Invitrogen/Life Technologies) for hES and 50 ng/mL for iPS cell line propagation. Cells were passaged as clumps using dispase (Gibco/Life Technologies) digestion at 1.5 mg/mL. For a typical 2D differentiation protocol, cells were grown in a 6-well plate to ~50% confluence in the presence of 2 µg/mL of puromycin, and the medium changed to selection-free KSR medium (Chung, T.L. et al., 2010), with or without inducers of differentiation, either 10% FCS or 5 µM retinoic acid (RA). Cells were harvested at 2-6 days after induction of differentiation. HES3 and H9 hES cell lines were provided by the UQ AIBN StemCore (www.stemcore.com.au). The UQEW01i-epifibC11 iPS episomally-derived cell line (referred to as WT11TF for brevity from now on) was described previously (Briggs, J.A. et al., 2012; Nayler, S. et al., 2012). At least three independent clones from hES and iPS parental cell lines with representative characteristics (levels of reporter expression in pluripotent state, reporter inactivation dynamics) were used for analysis.

2.2 *Lentiviral transduction of the pluripotent stem cells*

Lentiviral and packaging vectors, pCMVΔR8.2 and pVSV-G, were co-transfected into the 293FT cell line (Invitrogen) for viral particle production as described in (Whitworth, D.J. et al., 2012). HES3 and H9 hES cells adapted to passaging as a single-cell suspension, as described previously (Hudson, J. et al., 2012), as well as the hESC and UQEW01i-epifibC11 lines maintained as bulk cultures on Matrigel™ substrate(BD Biosciences), were transduced with lentiviral particles and maintained in mTeSR™1/Matrigel cultures under selection in the presence of 2.5 µg/mL of puromycin. Putative single cell-derived colonies homogeneous for moderate to high EGFP expression were generated from HES3 and H9 cell lines and transgene-free iPS clone WT11TF (UQEW01i-epifibC11). The following plasmids were obtained from Addgene (MA, U.S.A.) and used for lentivirus production: Addgene plasmid 21312 PL-SIN-PGK-EiP, 21313 PL-SIN-EOS-C(3+)-EiP (used to produce the described EOS(C3+)-EiP reporter lines), 21314 PL-SIN-EOS-S(4+)-EiP.

2.3 *Analysis of the reporter and pluripotency marker expression*

2.3.1 Flow cytometry: Analysis of cell surface markers was carried out essentially as described in (Ovchinnikov, D.A. et al., 2012; Whitworth, D.J. et al., 2012). Briefly, cells were brought to a single cell suspension using TrypLE or Cell Dissociation Buffer (Invitrogen) and stained live with anti-TRA-1-60 1/300 (MAB4360, Millipore), anti-SSEA4 1/400 (MAB4304, Millipore), anti-TG30 1/400 (MAB4427, Millipore), or with anti-OCT4 (MAB4419,Millipore) after 10 min fixation in 2% neutralized PFA and permeabilization with TritonX-100. For blocking, antibody incubation and permeabilization (when applicable) steps, PBS containing 10% heat-inactivated goat serum (Hyclone) and 0.1% TritonX-100 (0.4% for permeabilization) was used. For flow cytometry secondary antibodies anti-mouse IgG (H+L) AlexaFluor488 or anti-mouse IgM AlexaFluor488 (Molecular Probes/Invitrogen, Carlsbad, USA) were used at a 1/1000 dilution. Isotype control antibodies at

identical concentrations as the primary antibodies were used to set the gates. Flow cytometry data were collected using BD LSRII or AccuriCFlow C6 cytometers, and analysed using AccuriCFlow.

2.3.2 Immunofluorescence: The primary antibodies used for immunofluorescence were: anti-Oct3/4 (C10, sc5279 SantaCruz) at 1/75, anti-Sox2 (AB5603, Millipore) at 1/100, anti-NANOG (9220 Millipore) at 1/150. Secondary antibody: anti-mouse IgG (H+L) AlexaFluor488 (Molecular Probes/Invitrogen) was used at a 1/1000 dilution. Isotype controls at similar dilutions to the primary antibodies were used to assess non-specific binding. No significant labelling was detected in isotype control-incubated samples (data not shown). Counter-staining of the nuclei was performed using Hoechst 33342 nuclear dye (Invitrogen/Life Technologies) at 0.1 µg/mL in PBS.

2.4 *Statistical analysis of the reporter and surface marker correlation*

To correlate the extinction of pluripotency reporter-driven EGFP fluorescence with the surface pluripotency markers, flow cytometric analysis was performed on cells from 2D differentiation cultures at their intermediate stages: i.e. when the majority of the cells possessed an intermediate level of surface marker expression and differentiation phenotype. Fluorescence readings corresponding to single-cell events (gated on forward and side scatter plots and peak width) were exported in text (*.csv) format using Weasel v2.7 software (WEHI, Melbourne, Australia). Spearman's rank test analysis was performed as described in (McDonald, J.H., 2009). For calculations n>1000 was used, thus providing high confidence in the significance of the correlation ($p<0.01$).

2.5 *Reporter cell line evaluation in multifactorial microbioreactor array-based assays*

Data from full-factorial microbioreactor array screening experiments using a previously described system (Hudson, J. et al., 2012) were used to compare EOS-GFP expression with TG30 and TRA-1-60 expression levels. Briefly, microbioreactors were fabricated and sterilised, surface-coated with 400 µg/mL purified human fibronectin (BD Biosciences), and seeded with HES3::EOS(C3+)-EiP hESCs at ~4×10^4 cells/cm^2 in mTeSR^{TM}1 without FGF2 and TGFβ1 (StemCell Technologies) with inclusion of 10 µM ROCK inhibitor (Sigma-Aldrich). Cells were allowed to attach for 8-10 h then cultured for 6.5 d under 60 µL/h continuous fluid flow, under the panel of conditions shown in Figure 3E. At the endpoint, arrays were stained for TRA-1-60 and with Hoechst33342, and then imaged with a Zeiss LSM 710 laser scanning confocal microscope, with optical sections processed into a maximum intensity projection for image analysis. Total fluorescence intensities ($T_{EOS\text{-}GFP}$, for example) were extracted from array images with AGScan (https://mulcyber.toulouse.inra.fr/projects/agscan), and individual spot intensities in each channel were linearly transformed about the mean and standard deviation for all spots in that channel in an individual array, by $I_{EOS\text{-}GFP} = (T_{EOS\text{-}GFP} - \mu_{EOS\text{-}GFP})/\sigma_{EOS\text{-}GFP}$, where $I_{EOS\text{-}GFP}$ is termed the expression index of EOS-GFP, and $\mu_{\text{EOS-GFP}}$ is the mean and $\sigma_{\text{EOS-GFP}}$ the standard deviation of all spot intensities ($T_{\text{EOS-GFP}}$)for a particular array.

2.6 *Assessment of the differentiation potential of the pluripotency reporter cell lines*

The differentiation potential of EOS(C3+)-EiP cell lines was assessed using both *in vitro* embryoid body formation and *in vivo* teratoma formation assays. For embryoid body formation, cells were digested using dispase with subsequent resuspension in KSR medium (20% KnockOut^{TM} Serum

Replacement in DMEM/F12 medium, (both from LifeTechnologies), full description in (Chung, T.L. et al., 2010), in a well of a 6-well ultra-low attachment polystyrene plate (Falcon). For the teratoma formation assay, a pellet of $5x10^5$ cells was formed in the presence of Matrigel matrix at 1:50 dilution, and injected intramuscularly into the thigh muscle of a NOD/SCID mouse. Teratomas were harvested within 4-8 weeks, fixed and processed for paraffin embedding and histological analysis. Haematoxylin/eosin-stained 5 μm sections were permanently mounted, microscopically analysed and imaged on an Olympus IX51 inverted microscope equipped with MicroPublisher 3.3 RTV CCD camera (QImaging). For directed differentiation, we utilized previously described protocols for neuronal (Briggs, J.A. et al., 2012) differentiation from the hPSCs.

2.7 *Analysis of the pluripotency and lineage-specific marker expression in embryoid bodies derived from EOS(C3+)-EiP human pluripotent cell lines*

Downregulation of the expression of the conventional markers of pluripotency (*POU5F1* and *SOX2*) and selected markers of differentiation into the three major germ layers and their derivatives (*PAX6, EOMESODERMIN* and *CDX2)* were assayed using a qPCR-based approach. Briefly, total RNA was isolated using the Nucleospin RNA II kit (Macherey-Nagel) according to the manufacturer's protocol. 1 μg of isolated total RNA was utilized to synthesize cDNA using the iScript cDNA synthesis kit (Bio-Rad) according to the manufacturer's protocol. qPCR reactions comprised of Ssofast Evagreen (Bio-Rad) reaction mix and cDNA template were performed using the C1000 Thermal Cycler (Bio-Rad). All measured genes had their corresponding quantitative levels normalized to endogenous GAPDH. Primer sequences can be found in Table 1.

3.1 *Generation and characterisation of EOS reporter lines*

Human embryonic stem cell lines HES3 and H9 and the non-viral, integration-free human iPS cell line WT11TF (UQEW01i-epifibC11 (Briggs, J.A. et al., 2012)) were cultured under feeder-free conditions in mTeSR^{TM}1 or MEF-conditioned medium, and transduced with high titre lentiviral preparations produced using the EOS(C3+)-EGFP-IRES-PURO (from this time referred to as EOS(C3+)-EiP) constructs (Hotta, A. et al., 2009a). By day 5, up to 5 % of hES/iPS cells were found to express GFP. Such cultures were next allowed to expand for 3 more days before selection for stable integration of the transgene through the addition of initially 2 µg/ml puromycin, with a further increase to 2.5 µg/ml puromycin after day 6. After 12-16 days, we observed an almost homogeneous expression of EGFP in some colonies of surviving hESC and iPSC (Supplementary figure 1F). These colonies were next harvested, expanded and cryopreserved. During continued passaging of EOS(C3+)-EIP hES and iPS cells we observed the well-known sporadic inactivation of lentiviral transgenes previously observed in hPSCs (Liew, C.G. et al., 2007; Xia, X. et al., 2007), as indicated by the "salt-and-pepper" pattern of EGFP–expressing cells interspersed with those that have lost the expression (best illustrated in Supplementary Figure 1C), most likely due to epigenetically-driven lentiviral transgene inactivation. However, the addition of puromycin to such cultures allowed facile deletion of the vast majority of such cells (Supplementary Figure 1C, L, also discussed further in section 3.3). Clones displaying highly-homogeneous and penetrant expression of the reporter transgene were picked for further characterisation (Figure 1B, F, H and Supplementary Figure 1G-I). EOS(C3+)-EIP lentivirally-transduced hESC and iPSC were expanded for more than 20 weekly passages and exhibited normal karyotypes (Supplementary figure 1) and expressed high levels of the pluripotency-controlling transcription factors POU5F1/OCT4, SOX2 and NANOG (Figure 2A) by immunofluorescence and Oct4 and SOX2 by FACS (Table 2). When under constant selection, the fraction of EGFP-expressing resistant cells in single cell-derived clones was found to be consistently high in all hPSCs, specifically 90.14 ± 3.2% (5 clones) for HES3 hESC and 93.5 ± 2.7% (5 clones) for WT11TF iPSC (also see Table 1). The pluripotency of stable EOS lines was confirmed by RT-PCR analysis of embryoid bodies (EBs), which demonstrated expression of markers of all three germ layers and their derivatives (Supplementary figure 2). Furthermore, both hESC and iPSC EOS lines formed teratomas in NOD-SCID mice that exhibited cell types representative of the three germ layers (Figure 2 E-G for hESC- and Supplementary figure 2 for iPSC-derived EOS clones). In situ analysis of GFP expression in cultured hESC and iPSC lines indicated an excellent correlation between the loss of EGFP fluorescence and morphological signs of differentiation (Fig 1 I-J), such as those typically occurring at the centres of hES and iPS colonies.

3.2 *Enrichment for pluripotent cells improves directed differentiation protocols*

Because differentiated cells excrete paracrine factors and impose confounding cell-cell interactions on human pluripotent stem cells, homogeneous and efficient differentiation protocols generally require (or at least greatly benefit from) a pure pluripotent starting population, *e.g.* (Hudson, J. et al., 2012). The presence of the EOS promoter-driven puromycin antibiotic selection marker allows for a rapid deletion of any differentiated cells from such cultures (*e.g.* Fig 1K where the EGFP-IRES-puro

cassette is driven by a ubiquitous PGK promoter compared to a colony in Fig 1L utilising the EOS(C3+) promoter). Nearly-complete ablation could be achieved in as little as 2-3 days due to the potent nature of the puromycin selection, causing rapid death and detachment of non-resistant cells (Supplementary Figure 1L and data not shown).

To exemplify the benefits of this selection, we cultured three WT11TF::EOS(C3+)-EiP clones for 3 weeks in parallel in conditioned medium containing 50 ng/mL bFGF with and without puromycin selection. Subsequent flow-cytometric analysis showed that the antibiotic-selected population contained 1.7 ± 0.5% of POU5F1/OCT4$^-$ cells (N=3) while the similarly cultured non-selected populations contained 9.3 ± 2.3% (N=3) POU5F1/OCT4$^-$ cells (Figure 2H). We next assessed whether this difference in starting population affects the efficiency of neuronal differentiation (See Methods section). Indeed, after 6 days of directed neuronal differentiation, puromycin-selected cultures displayed 31 ± 4.2% PAX6+ neuroepithelial progenitors, as measured by flow cytometry with anti-PAX6 antibody, whereas unselected cultures show 20 ± 5.1% PAX6+ cells (N=3, Figure 2H and Supplementary Figure 2D).

3.3 *Validation of the functionality of the pluripotency reporter lines*

In time-lapse videos of differentiating colonies of the reporter clones a very good concordance between the EGFP extinction and morphological changes associated with differentiation is apparent (Supplementary videos 1-3), with central areas of colonies inactivating the reporter and acquiring a differentiated morphology first. Complete extinction of reporter fluorescence was observed after 6 days of withdrawal of the pluripotency-maintaining factors, addition of foetal calf serum (FCS) or supplementation with retinoic acid at 5 µM (Figure 3A).

Since the conventional variant of the EGFP protein was used in the construction of the transgene it was important to verify the dynamic response range and kinetics of EGFP downregulation during differentiation of the reporter lines in comparison with the expression of other widely-used pluripotency markers. To this end we allowed reporter-containing hESC and iPSC to undergo spontaneous differentiation, through the withdrawal of FGF and TGFβ from the culture medium, or induced differentiation through addition of retinoic acid to the medium, and assessed EGFP and pluripotency marker expression at days 0, 2, 4 and 6. Downregulation of EGFP occurred more rapidly and more completely than downregulation of the widely-used pluripotency markers TRA-1-60, TG30 and SSEA4 (Figure 3C, D, Supplementary Figure 3B, C), and correlated very closely with the downregulation of OCT4 expression (Figure 3C, Supplementary Figure 3C) for both hESC-(Figure 3) and iPS-derived (Supplementary Figure 3) reporter clones. Downregulation of the reporter followed the decrease in levels of OCT4 protein with a ~24 hours lag period resulting in a population of cells with reduced OCT4 expression but still retaining detectable levels of EGFP reporter protein expression (Figure 3C). This observation is consistent with the stability of the EGFP mRNA (Cubitt, A.B. et al., 1995), the half-life of EGFP protein of around 26 hours in mammalian cells (Corish, P. and Tyler-Smith, C., 1999) and the fact that OCT4 is itself is a direct regulator of the reporter promoter.

Flow-cytometric analysis revealed a high concordance between EGFP expression and expression of pluripotency cell surface markers such as TRA-1-60 and SSEA-4 (representative raw data plots shown for TRA1-60 in Figure 3B, and for SSEA4 in Supplementary Figure 3A). The Spearman's rank-based test is most suitable for assessment of general trends of this type, as it is not based on any assumption of the nature of the sample's distribution. The correlation appeared to be very good with Spearman's rho values ≥0.7 for EGFP and surface markers, >0.8 for TRA-1-60 (r_s=0.81 ± 0.9%, n=3) and r_s >0.7 for SSEA4 (Supplementary Figure 3A and data not shown).

Collectively these data demonstrate that EGFP expression in lentivirally-generated hESC and hiPSC EOS(C3+)-EiP pluripotency reporter lines faithfully and dynamically reports on pluripotency and that such lines behave as normal human pluripotent cell lines.

3.4 Pluripotency screening in a multifactorial microbioreactor array

The pluripotency reporter cell line HES3:EOS-C(3+)-EiP was seeded into the microbioreactor array (Titmarsh, D.M. et al., 2013) and subjected to factorial combinations of the pluripotency maintenance factors FGF2 and TGFβ1 and the differentiation-inducing factor retinoic acid (RA) using the concentrations listed in the panel in Figure 3E. After continuous slow perfusion culture for 6.5 days under these conditions, cells had proliferated and showed morphological differences between various treatments, and expressed GFP under the control of the EOS(C3+)-EIP promoter construct. GFP expression in the HES3::EOS-C(3+)-EiP reporter line exhibited a high dynamic range (bright expression and efficient extinction) and sensitive response. The presence of 5 or 10 μM RA reduced EOS(C3+)-EIP expression to a baseline level regardless of FGF2 or TGFβ1 treatment. Also, treatments lacking maintenance factors (FGF2 and TGFβ1, Column 1) or with insufficient factors (0.25 ng/mL TGFβ1, Column 4) had similar EOS(C3+)-EIP expression to RA-treated conditions, showing sensitivity to absent or insufficient pro-maintenance signals. Supply of sufficient maintenance factors (column 25, for example) maintained high-level expression. The highest response for all 3 pluripotency markers occurred at Column 25 (100 ng/mL FGF2, 0.5 ng/mL TGFβ1), which corresponds to reconstituted composition of original mTeSR^{TM}1 maintenance medium.

TRA-1-60 exhibited a lower dynamic range than EOS(C3+)-EIP, and had higher residual expression, resulting in "false-positive" readouts in pluripotency non-maintaining environments in some differentiating conditions in the array (Figure 3D and figure legend, and Figures 3C and Supplementary Figure 2). As a result of this, levels of TRA-1-60 expression, both when normalized across the microbioreactor (Figure 3E) or against the chamber's DNA contents (data not shown), varied more widely than EOS(C3+)-EIP and thus were less sensitive as an indicator of differentiation, while the reporter expression appeared to correlate well with the conditions favouring maintenance of pluripotency, and showed robust down-regulation in all RA-containing environments (Figure 3E).

3.5 *Persistence of the EOS(C3+)-EIP reporter uncovers heterogeneity amongst embryoid bodies in differentiation assays*

To investigate extinction dynamics of the pluripotency reporter in a 3D setting, we next subjected hESC and iPSC reporter lines to embryoid body (EB) differentiation protocols. Conventional analysis of lineage-specific marker expression on pooled EBs after 8 weeks shows the expected overall

decrease in expression of both *OCT4*/*POU5F1*and *SOX2*, and up-regulation of the neuronal (*PAX6, TUBB3*), posterior endoderm (*CDX2)* and mesoderm (*EOMESODERMIN*) lineage markers that is routinely reported in the literature (Supplementary figure 2). However, in situ examination of EGFP expression in EOS(C3+)-hESC and iPSC reporter clone derived suspension EBs revealed an unexpected persistence and remarkable inter-EB diversity (Figure 4 A-F). While there was a marked and expected overall downregulation of reporter expression after 2 weeks in culture when compared to freshly generated EBs (Figure 4C), a large proportion of the EBs retained low levels of EGFP expression (Figure 4C,C'). After 6 weeks of suspension culture in KSR medium without pluripotency-maintaining factors a number of EBs (~70%, increasing with time in culture) lost all, or most, of the EOS(C3+)-EIP reporter-driven EGFP fluorescence (Figure 4F). These EBs were typically smaller, and started to form cavities or cyst-like structures (Supplementary Figure 4B,C). Pooling of the 6 week-old EGFP-low and EGFP-negative EBs followed by a qPCR analysis revealed that EBs retaining reporter fluorescence maintained higher levels of residual *OCT4*/*POU5F1* (tested using 2 distinct qPCR primer sets), its downstream target *LEFTYA*, and higher levels of the mesodermal marker *EOMESODERMIN* (Supplementary Figure 4A). On the other hand, low EGFP-expressing EBs showed similar expression levels of *SOX2* (a direct regulator of the EOS promoter, also tested using 2 distinct qPCR assays, one of which is shown in Supplementary Figure 4A), *DNMT3A* (a pluripotency-associated DNA methyl-transferase), *PAX6* (a neuroectodermal marker) and *TUBB3* (a neuronal marker*)* (Supplementary Figure 4A). By week 12 virtually no EGFP expression was detected in any of the EBs (Figure 4F, Supplementary Figure 4C,D), indicating that this persistence of EGFP expression in a subset of 6 week-old EBs that appears to undergo mesodermal differentiation is not due to an intrinsic inability to down-regulate the reporter. Identification of this unexpected heterogeneity within EB pools would not have been possible without the use of the pluripotency reporter and is one example of its utility.

4. Discussion

The ability for live monitoring of the pluripotency status of stem cells is important for both optimisation of growth and differentiation conditions, and assessment of the retention and extinction of pluripotency by stem cell lines, in particular iPSCs, during differentiation. The EOS promoter tested is available in 2 variants, with tandemly concatenated *OCT4*/*SOX2* binding sites derived from either the *OCT4/POU5F1* promoter (C3, 3 copies) or *SOX2* promoter (S4 variant, 4 copies) (Hotta, A. et al., 2009a; Hotta, A. et al., 2009b). Clonally-derived hES lines were originally generated using both variants, but preliminary analysis showed that the former (C3+) variant tended to produce lines with more consistent reporter expression and responsiveness across a wider range of basal fluorescence levels, and produced a larger fraction of puromycin-resistant clones that were $EGFP^+$ (78% vs. 62%, n=50 for HES3).

It is likely that a pluripotency reporter construct could be a useful tool to drive "suicide gene"-based cassettes and delete unwanted stem cells that either retain or re-acquire pluripotency in iPSC-derived grafts. In this study we show that the reporter is rapidly extinguished during various directed and spontaneous differentiation protocols, with more efficient extinction in neurally-biased embryoid body differentiation, but can maintain residual expression for extended periods in cells that differentiate into embryoid bodies biased towards the mesodermal lineage. Reporter expression is eventually lost in the majority of cellular aggregates (EBs) after 6 weeks and from all cells after 12 weeks of EB-based differentiation, identifying the time window when such deletion strategies should commence *in vivo*. One could further envisage that the lentiviral delivery of the reporter to newly generated iPS cell lines could be used as a tool to screen such lines for their propensity to re-acquire pluripotency following differentiation, *e.g.* as EBs, and study the molecular basis of this unwanted property.

In the shorter term the greatest utility of these lines lies in their ability to report on pluripotency in real time. We exemplify that human pluripotent stem cell lines that stably express the reporter allow monitoring of the pluripotency states of individual cells in a wide spectrum of experimental settings (*e.g.* Titmarsh, D.M. et al., 2013). These lines are particularly valuable for microbioreactor arrays, bioengineering applications, and other applications requiring non-intrusive and continuous monitoring regimes such as screening of culture surfaces.

The reporter lines allowed us to observe in real time that colonies with greater than 300 cells displayed slightly higher expression levels of the reporter at the colonies' outermost few cell layers (Figure 1A, F, fluorescence profiles across colonies shown in Supplementary figure 2M). This is in agreement with previous observations of higher levels of TGFβ/activin signalling and higher expression of pluripotency markers on the edges of the human pluripotent stem cell colonies (Hough, S.R. et al., 2009).

The ability to efficiently select for cells of a consistently highly-pluripotent phenotype in cultures, including those on a large scale, is crucial for providing homogeneously-pluripotent starting material for directed differentiation protocols, for instance in the generation of cardiomyocytes and certain neuronal cell types.

We demonstrate that removal of the puromycin selection causes the emergence of a population of differentiated, OCT4-negative cells (Figure 2H). It appears that these spontaneously-differentiated cells have a cell non-autonomous effect on the directed neuroepithelial conversion of the whole population, resulting in a 33% drop in the numbers of neuroectodermally-converted cells as measured by the expression of PAX6 protein (Figure 2H and Supplementary Figure 2D).

The difference in correlation between EGFP and TRA160 as compared to that between EGFP and SSEA4 or TG30 is consistent with the observations that extinction dynamics of TRA-1-60 with differentiation are substantially faster than that of TG30 or SSEA4 (Figure 3D and supplementary figure 2). The dynamic response range and sensitivity of the *EOS(C3+)-EGFP-IRES-Puro* reporter is in fact superior to TRA-1-60, as can be judged from the more rapid and pronounced downregulation of EGFP than TRA-1-60 following RA-induced differentiation (Figure 3D and Supplementary Figure 2C). Similarly, when titrating concentration combinations of FGF2 and TGFβ1, the expression of reporter EGFP corresponded much more dynamically to variations in these pluripotency-maintaining factors than TRA-1-60 (Figure 3E).

In conclusion we have generated and characterized multiple hES and iPS cell lines that should prove a useful tool for exploration of the maintenance of pluripotency, pluripotency extinction and re-emergence in real time, enabling simple and scalable selection of hPSC populations that are near homogeneously pluripotent, and undergo more homogeneous and efficient directed differentiation .

Acknowledgements

We would like to thank StemCore (http://www.stemcore.com.au) at the University of Queensland for provision of hES cell lines and logistical support with teratoma and karyotype analyses. We also would like to acknowledge the funding provided by the Australian Stem Cell Centre and the ARC Centre of Excellence "Stem Cells Australia".

Alvarez, C.V., Garcia-Lavandeira, M., Garcia-Rendueles, M.E., Diaz-Rodriguez, E., Garcia-Rendueles, A.R., Perez-Romero, S., Vila, T.V., Rodrigues, J.S., Lear, P.V., and Bravo, S.B. (2012). Defining stem cell types: understanding the therapeutic potential of ESCs, ASCs, and iPS cells. J Mol Endocrinol *49*, R89-111.

Briggs, J.A., Sun, J., Shepherd, J., Ovchinnikov, D.A., Chung, T.L., Nayler, S.P., Kao, L.P., Morrow, C.A., Thakar, N.Y., Soo, S.Y.*, et al.* (2012). Integration-free induced pluripotent stem cells model genetic and neural developmental features of Down syndrome etiology. Stem Cells *31*, 467-478.

Chung, T.L., Turner, J.P., Thaker, N.Y., Kolle, G., Cooper-White, J.J., Grimmond, S.M., Pera, M.F., and Wolvetang, E.J. (2010). Ascorbate promotes epigenetic activation of CD30 in human embryonic stem cells. Stem Cells *28*, 1782-1793.

Corish, P., and Tyler-Smith, C. (1999). Attenuation of green fluorescent protein half-life in mammalian cells. Protein Eng *12*, 1035-1040.

Cubitt, A.B., Heim, R., Adams, S.R., Boyd, A.E., Gross, L.A., and Tsien, R.Y. (1995). Understanding, improving and using green fluorescent proteins. Trends in Biochemical Sciences *20*, 448-455.

Hotta, A., Cheung, A.Y., Farra, N., Garcha, K., Chang, W.Y., Pasceri, P., Stanford, W.L., and Ellis, J. (2009a). EOS lentiviral vector selection system for human induced pluripotent stem cells. Nat Protoc *4*, 1828-1844.

Hotta, A., Cheung, A.Y., Farra, N., Vijayaragavan, K., Seguin, C.A., Draper, J.S., Pasceri, P., Maksakova, I.A., Mager, D.L., Rossant, J.*, et al.* (2009b). Isolation of human iPS cells using EOS lentiviral vectors to select for pluripotency. Nat Methods *6*, 370-376.

Hough, S.R., Laslett, A.L., Grimmond, S.B., Kolle, G., and Pera, M.F. (2009). A continuum of cell states spans pluripotency and lineage commitment in human embryonic stem cells. PLoS One *4*, e7708.

Hudson, J., Titmarsh, D., Hidalgo, A., Wolvetang, E., and Cooper-White, J. (2012). Primitive cardiac cells from human embryonic stem cells. Stem Cells Dev *21*, 1513-1523.

Inoue, H., and Yamanaka, S. (2011). The use of induced pluripotent stem cells in drug development. Clin Pharmacol Ther *89*, 655-661.

Keller, G. (2005). Embryonic stem cell differentiation: emergence of a new era in biology and medicine. Genes Dev *19*, 1129-1155.

Lepski, G. (2012). What do we know about the neurogenic potential of different stem cell types? Arq Neuropsiquiatr *70*, 540-546.

Liew, C.G., Draper, J.S., Walsh, J., Moore, H., and Andrews, P.W. (2007). Transient and stable transgene expression in human embryonic stem cells. Stem Cells *25*, 1521-1528.

Liu, Y., Jiang, P., and Deng, W. (2011). OLIG gene targeting in human pluripotent stem cells for motor neuron and oligodendrocyte differentiation. Nat Protoc *6*, 640-655.

McDonald, J.H. (2009). Handbook of Biological Statistics, 2nd edn (Baltimore, maryland, Sparky House Publishing).

Nayler, S., Gatei, M., Kozlov, S., Gatti, R., Mar, J.C., Wells, C.A., Lavin, M., and Wolvetang, E. (2012). Induced pluripotent stem cells from ataxia-telangiectasia recapitulate the cellular phenotype. Stem Cells Transl Med *1*, 523-535.

Ovchinnikov, D.A., Turner, J.P., Titmarsh, D.M., Thakar, N.Y., Sin, D.C., Cooper-White, J.J., and Wolvetang, E.J. (2012). Generation of a human embryonic stem cell line stably expressing high levels of the fluorescent protein mCherry. World J Stem Cells *4*, 71-79.

Pera, M.F. (2011). Stem cells: The dark side of induced pluripotency. Nature *471*, 46-47.

Rowntree, R.K., and McNeish, J.D. (2010). Induced pluripotent stem cells: opportunities as research and development tools in 21st century drug discovery. Regen Med *5*, 557-568.

Titmarsh, D.M., Hudson, J.E., Hidalgo, A., Elefanty, A.G., Stanley, E.G., Wolvetang, E.J., and Cooper-White, J.J. (2012). Microbioreactor arrays for full factorial screening of exogenous and paracrine factors in human embryonic stem cell differentiation. PLoS One *7*, e52405.

Titmarsh, D.M., Ovchinnikov, D.A., Wolvetang, E.J., and Cooper-White, J.J. (2013). Full factorial screening of human embryonic stem cell maintenance with multiplexed microbioreactor arrays. Biotechnology journal *8*, 822-834.

Tobin, S.C., and Kim, K. (2012). Generating pluripotent stem cells: differential epigenetic changes during cellular reprogramming. FEBS Lett *586*, 2874-2881.

Ungrin, M.D., Joshi, C., Nica, A., Bauwens, C., and Zandstra, P.W. (2008). Reproducible, ultra high-throughput formation of multicellular organization from single cell suspension-derived human embryonic stem cell aggregates. PLoS One *3*, e1565.

Whitworth, D.J., Ovchinnikov, D.A., and Wolvetang, E.J. (2012). Generation and characterization of LIF-dependent canine induced pluripotent stem cells from adult dermal fibroblasts. Stem Cells Dev *21*, 2288-2297.

Xia, X., Zhang, Y., Zieth, C.R., and Zhang, S.C. (2007). Transgenes delivered by lentiviral vector are suppressed in human embryonic stem cells in a promoter-dependent manner. Stem Cells Dev *16*, 167-176.

Zou, J., Maeder, M.L., Mali, P., Pruett-Miller, S.M., Thibodeau-Beganny, S., Chou, B.K., Chen, G., Ye, Z., Park, I.H., Daley, G.Q.*, et al.* (2009). Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. Cell Stem Cell *5*, 97-110.

Zwaka, T.P., and Thomson, J.A. (2003). Homologous recombination in human embryonic stem cells. Nature biotechnology *21*, 319-321.

Figure 1.Cellular distribution of reporter fluorescence in the EOS promoter-based reporter lines grown in feeder-free conditions. A,B - Corresponding phase-contrast and EGFP epifluorescence images of a HES3 hES cell line-based EOS(C3+)-EiP colony grown under puromycin selection. **C,D** - Flow cytometry histograms illustrating consistently high levels of EOS(C3+)-EiP expression in 3 puromycin-selected HES3 EOS(C3+) clones. Parental cell line HES3 (orig), was used as a control. **E,F** Phase contrast and epifluorescence images of a typical H9 hES EOS(C3)-EiP colony. **G,H** - Images of the transgene-free iPS line UQEW01i-epifibC11-based EOS(C3+)-EiP cell line colony. (A,B and E-H are all passage 2 after transduction and continuous puromycin selection) **I,J** - phase-contrast and epifluorescence images of the iPS (UQEW01i-epifibC11) EOS(C3+)-EiP cell lines' colonies following partial spontaneous differentiation in their central portions. Prominent drop in the reporter fluorescence correlates with morphological signs of differentiation. **K,L**- phase-contrast images of the edges of HES3::PGK-EiP (K) and HES3::EOS(C3+)-EiP (L) colonies, respectively, grown in presence of 2µg/mL of puromycin, illustrating efficient elimination of fibroblast-like differentiated cells from the colony's edges in EOS(C3+)- but not in the ubiquitous PGK promoter-driven EGFP and puromycin HES3 line. Scale bars = 20µm.

Figure 2. Validating pluripotency and differentiation potential of the EOS(C3+)-EIP pluripotency reporter in hESC and iPSC lines, and their utility in improving directed differentiation. **A**. An immunofluorescent staining for OCT4/POU5F1 and SOX2 in representative HES3::EOS(C3+)-EiP colonies grown under puromycin selection. Scale bar=20µm. **B-D**. Flow-cytometric analysis of EGFP reporter and surface marker expression (SSEA4 and TRA-1-60) in a HES3 EOS(C3+)-EiP hES cell line grown for 1 day without selection. **E-G.** Histological teratoma sections derived from HES3::EOS(C3+)-EiP cell line, demonstrating fully-differentiated tissues derived from the three germ layers: mesoderm (**E**, cartilage), ectoderm (**F**, retinal pigmented neuroepithelium-like) and endoderm (**G**, secretory epithelium and goblet-like cells lining a mucus-filled cavity, *). Hematoxylin/eosin staining on 5 µm-thick paraffin sections, scale bar 100 µm**. H,** Flow cytometric analysis of the difference in loss of pluripotency marker POU5F1/OCT4, and subsequent acquisition of expression of the neuroepithelial marker PAX6 in iPS-derived reporter clone grown with (green) and without (blue) selection.

Figure 3. Utility of the EOS(C3+)-EiP reporter lines for assessment of the pluripotency status. A. Histogram of the distribution of reporter fluorescence levels illustrating effects of various differentiation-inducing conditions on the expression of the reporter in HES3::EOS(C3+). **B.** Plot illustrating the correlation between EGFP fluorescence and TRA1-60 expression in HES3::EOS(C3+)-EiP after 3 days of FCS-induced differentiation as measured by flow cytometry. Only 300 events (for visual clarity) were plotted on a double decimal logarithm scale plot. Spearman's rho coefficient for this plot r_s= 0.82. **C.** Contour-plots of the flow-cytometric measurements illustrating kinetics of the decrease in OCT4 and EGFP expression in differentiating HES3::EOS(C3+)-EiP. **D**. Histograms illustrating dynamics of OCT4,EGFP, TRA1-60 and TG30 extinction in a HES3::EOS(C3+)-EiP clone during the time course of differentiation in the KSR medium **E.** Expression indices (see Methods, and (Titmarsh, D.M. et al., 2013) for definition) of EOS(C3+)-EIP reporter and TRA-1-60 in microbioreactor during factorial analysis of the effects of FGF2, TGFβ1 and RA concentration on maintenance of pluripotency in a HES3 reporter cell line. Reporter and surface marker fluorescence reading for each of the wells was normalized against DNA content. Fluorescence indices are plotted as mean ± SD.

Figure 4. Use of the EOS reporter reveals heterogeneity within pools of spontaneously-differentiating embryoid bodies (EBs). A. Fluorescent image of representative EBs 2 days after formation. **B.** EBs 1 weeks after formation retain high levels of EOS(C3+)-EiP reporter fluorescence. **C, C'.** Fluorescent and bright-field images of EBs 2 weeks into differentiation, when a number of EBs (<10%) fully inactivate the reporter. **D, D'.** Representative images of both types of EBs at 6 weeks of differentiation that clearly retain (EGFP+) or extinguish (EGFP-) the reporter expression in majority of their cells, and were used for the differential expression analysis. **E, E'.** By 8 weeks, only low level of EGFP expression persists in some of the EBs, and virtually all expression is lost by week 12 (See Supplementary Figure 4). **F.** Representation of the quantification of EGFP+ and EGFP- fractions in iPSC- reporter derived EBs. Scored EB number N>40 for each time point. **G.** qPCR analysis of the expression of pluripotency marker *POU5F1/OCT4*, and mesodermal (*EOMESODERMIN*) and neuroepithelial (*PAX6*) markers in $EGFP^+$ and $EGFP^-$ EBs at 6 weeks. Persistence of EGFP correlates with high levels of residual *POU5F1*/*OCT4* expression. See Discussion for more detail. Data for all genes is normalised against the internal *GAPDH* control, error bars-SEM values (N=3). Scale bar = 500µm.

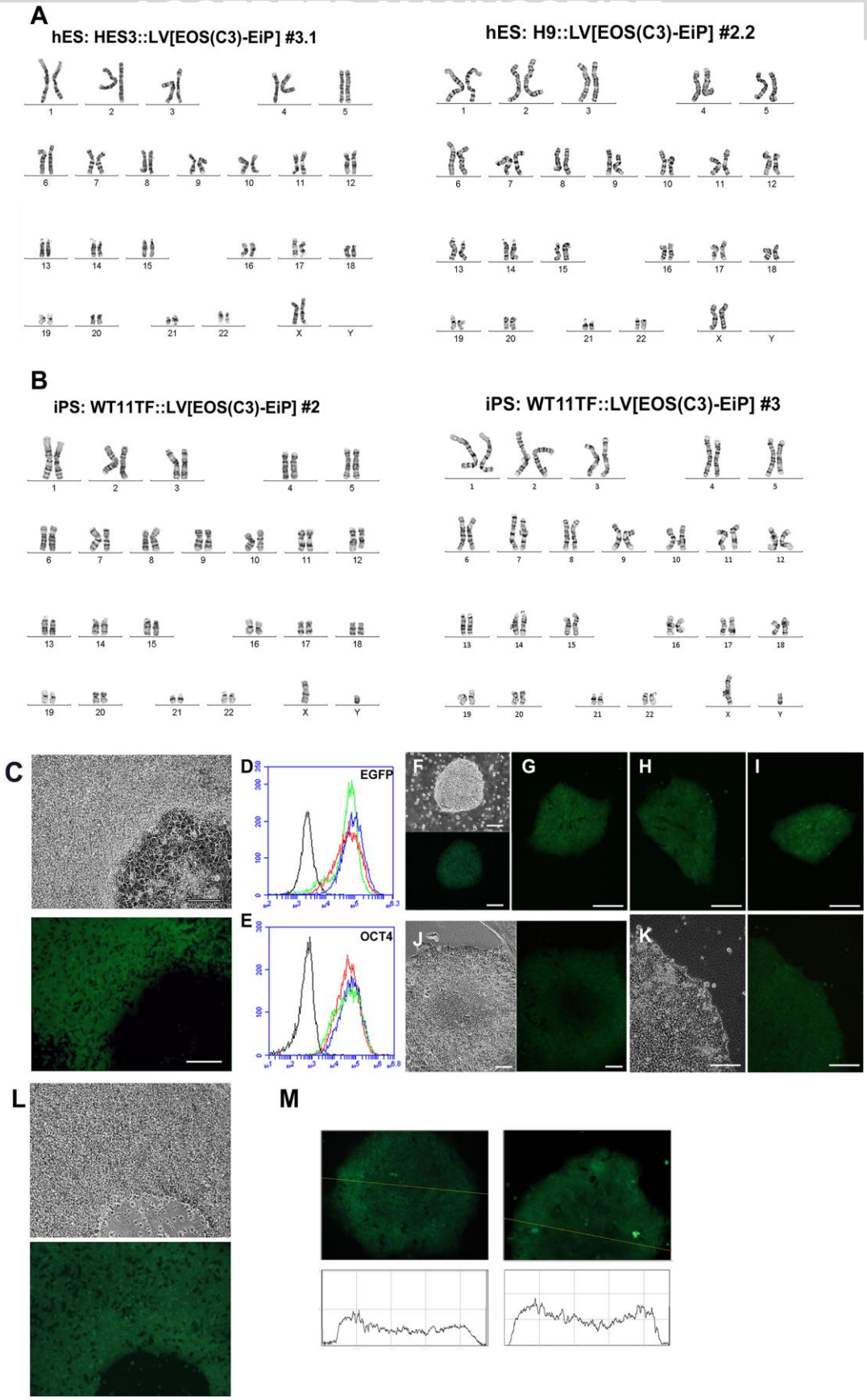
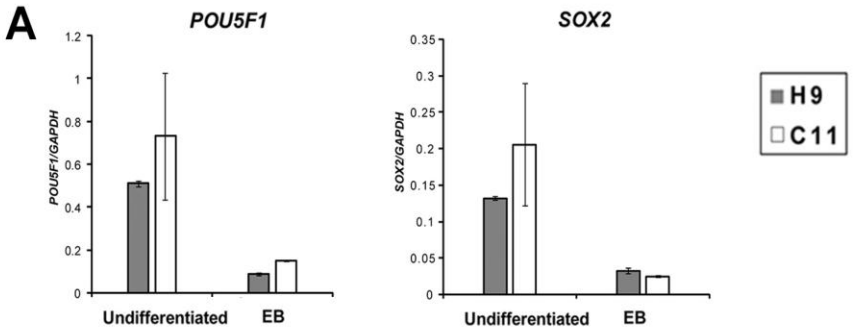
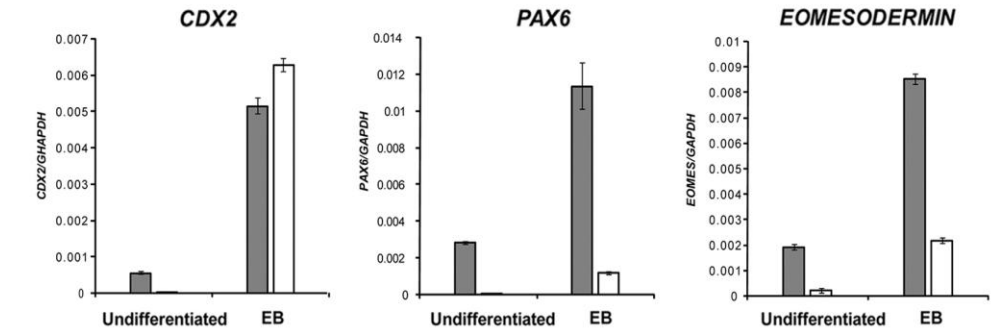
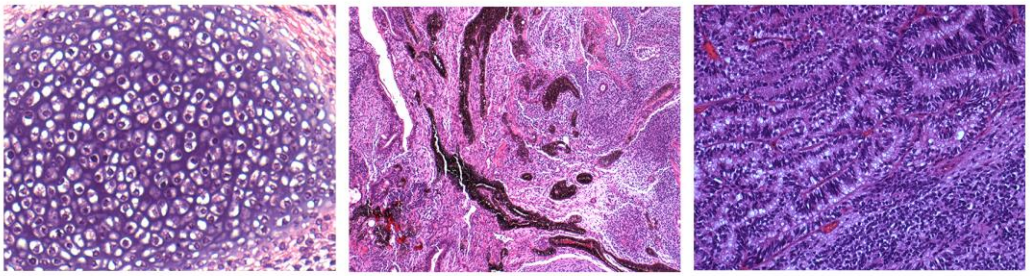

A
hES: HES3::LV[EOS(C3)-EiP] #3.1
hES: H9::LV[EOS(C3)-EiP] #2.2
B
iPS: WT11TF::LV[EOS(C3)-EiP] #2
iPS: WT11TF::LV[EOS(C3)-EiP] #3
C
D
EGFP
E
OCT4
F
G
H
I
J
K
L
M

Figure 1







D

Clone	%OCT4+ before diff	% PAX6+ after 6d diff
C11#1 -puro	9.1	19.9
C11#1 +puro	1.3	33.8
C11 #2 -puro	9.7	19.1
C11 #2 +puro	2.2	29.4
C11 #3 -puro	9.0	21.2
C11#3 +puro	1.6	30.4

Figure 2

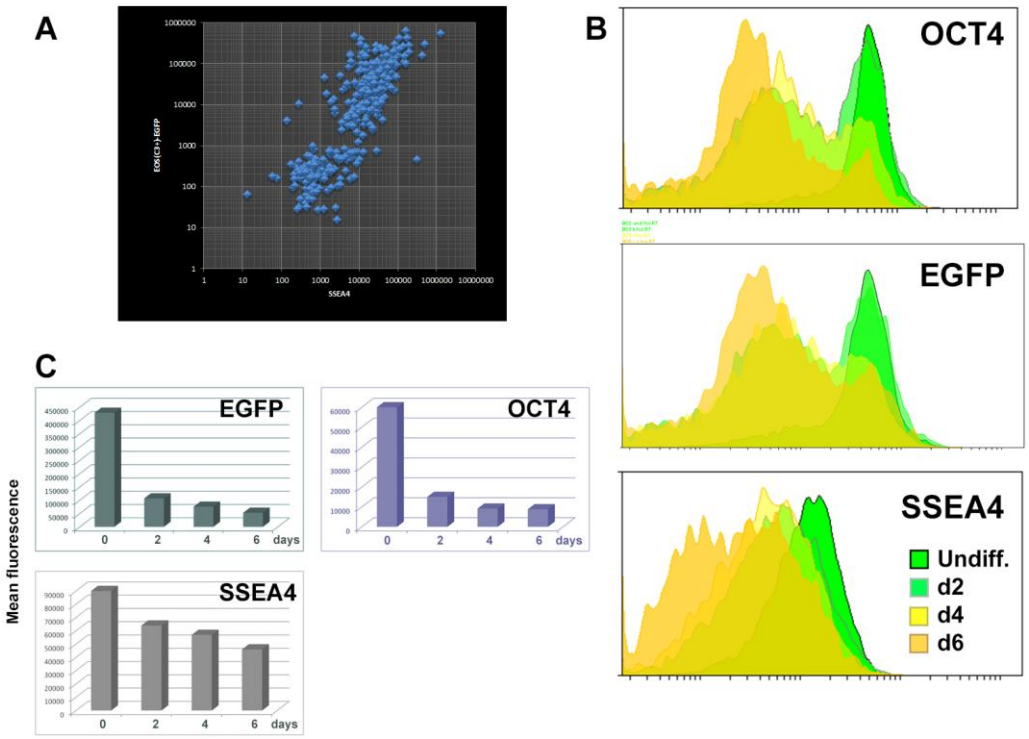


Figure 3

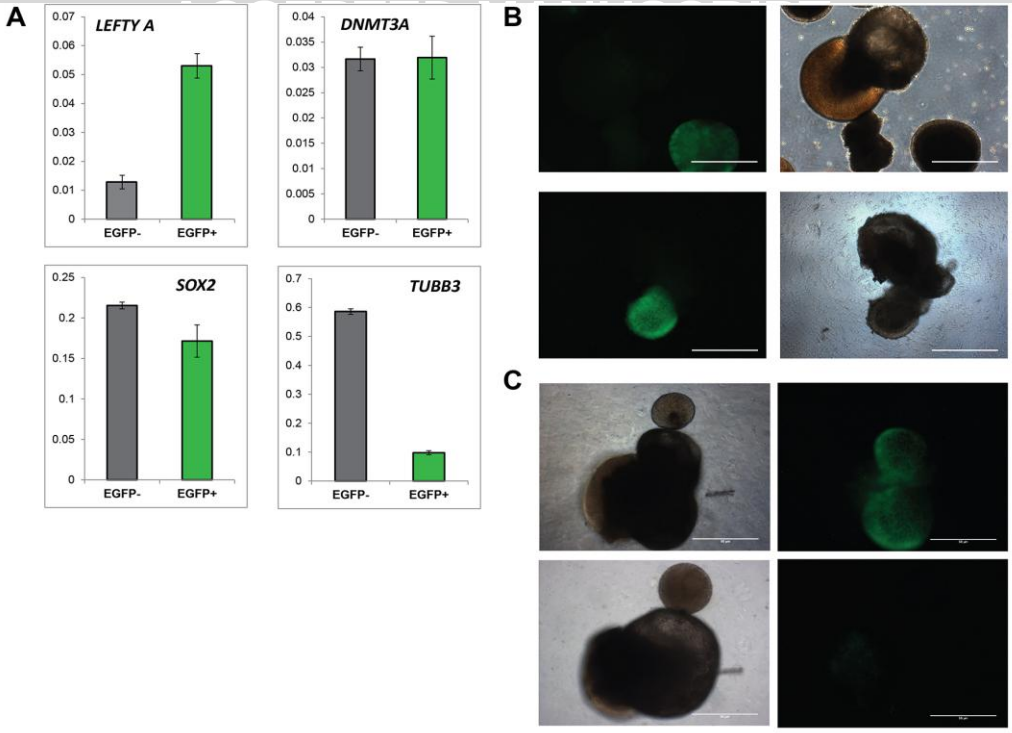


Figure 4

Table 1. Gene targets and their corresponding primer sequences utilized for qPCR analysis.

Gene Target	Forward Primer	Reverse Primer
POU5F1	TGAAGCTGGAGAAGGAGAAG	ATCGGCCTGTGTATATCCC
SOX2	CCACCTACAGCATGTCCTACTCG	GGGAGGAAGAGGTAACCACAGG
CDX2	GGAGCTGGAGAAGGAGTTT	TGATTTTCCTCTCCTTTGCTC
PAX6	CAGCACCAGTGTCTACCAACCA	CAGATGTGAAGGAGGAAACCG
EOMES	CAAATTCCACCGCCACCAAACTGA	TTGTAGTGGGCAGTGGGATTGAGT
GAPDH	ATGGGGAAGGTGAAGGTCG	TAAAAGCAGCCCTGGTGACC
SOX17	GGCGCAGCAGAATCCAGA	CCACGACTTGCCCAGCAT
SOX2-2	GCTACAGCATGATGCAGGACCA	TCTGCGAGCTGGTCATGGAGTT

Table 2. Characterisation of stable EOS(C3+)-EiP hPSC clones used in this study

Cell type	Cell line	Clone	Karyotype	% EGFP^{+*}	% OCT4^{+*}	% TRA1-60*	NANOG^{+}(IF$^{¥}$)
hES	HES3	3.1	Normal ♀	**91.2**	97.1	97.9	Yes
		3.2	Normal ♀	**89.7**	96.9	98.2	Yes
		3.3	N/A	**86.8**	96.2	98.0	Yes
	H9	#1	Normal ♀	**91.7**	96.7	97.1	Yes
		#2	Normal ♀	**94.2**	98.6	99.3	Yes
		#3	Normal ♀	**92.3**	99.3	98.4	Yes
iPS	WT11TF	#1	Normal ♂	**93.8**	98.1	98.9	Yes
		#2	Normal ♂	**94.1**	98.3	97.9	Yes
		#3	Normal ♂	**92.3**	97.8	98.4	Yes

*** in cultures kept under continuous selection with 2µg/mL puromycin for 2 weeks**

¥ >95% positive by immunofluorescent staining for nuclear NANOG protein

N/A not available

- stable selectable pluripotency reporter lines in hES and transgene-free iPS cells
- more sensitive pluripotency readout than conventional surface markers
- pluripotency enrichment facilitates directed differentiation protocols
- reveals heterogeneity in conventional embryoid body assays