THE CELL SURFACE MARKER SUSHI CONTAINING DOMAIN 2 FACILITATES ESTABLISHMENT OF HUMAN NAÏVE PLURIPOTENT STEM CELLS

Nicholas Bredenkamp[1], Giuliano Giuseppe Stirparo[1], Jennifer Nichols[1,3], Austin Smith[1,2*] and Ge Guo[1*]

[1]Wellcome–MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge CB2 1QR, United Kingdom.

[2]Department of Biochemistry, University of Cambridge, Cambridge, CB2 1GA, United Kingdom

[3]Department of Physiology, Development and Neuroscience, University of Cambridge, CB2 1GA United Kingdom

 *Corresponding authors: Ge Guo, gg251@ cam.ac.uk;

                                Austin Smith, austin.smith@cscr.cam.ac.uk

Running title: SUSD2 identifies human naïve PSC

SUMMARY

Recently naïve human pluripotent stem cells (hPSC) have been described that relate to an earlier stage of development than conventional hPSC. Naïve hPSC remain challenging to generate and authenticate, however. Here we report that Sushi Containing Domain 2 (SUSD2) is a robust cell surface marker of naïve hPSC in the embryo and in vitro. *SUSD2* transcripts are enriched in the pre-implantation epiblast of human blastocysts and immunostaining shows localisation of SUSD2 to KLF17 positive epiblast cells. *SUSD2* mRNA is strongly expressed in naïve hPSC, but is negligible in other hPSC. SUSD2 staining of live or fixed cells provides unambiguous discrimination of naïve versus conventional hPSC. SUSD2 staining or flow cytometry enable monitoring of naïve hPSC in maintenance culture, and their isolation and quantification during resetting of conventional hPSC or somatic cell reprogramming. Thus SUSD2 is a powerful non-invasive tool for reliable identification and purification of the naïve hPSC phenotype.

INTRODUCTION

Pluripotent cells are present in the human embryo for around 10 days, from emergence in the blastocyst until lineage commitment during gastrulation (Rossant and Tam, 2017). In vitro, two classes of human pluripotent stem cell (hPSC) have been described. Conventional hPSC, derived from inner cell mass (ICM) explants (O'Leary et al., 2012; Thomson et al., 1998) or generated by somatic cell reprogramming (iPSC) (Takahashi et al., 2007), share characteristics with late post-implantation gastrulating epiblast (Nakamura et al., 2016). Propagation of stem cells resembling naïve emergent epiblast has been reported more recently (Guo et al., 2017; Takashima et al., 2014; Theunissen et al., 2014). Naïve hPSC are generated by conversion from conventional PSC, a process termed resetting. In addition, embryonic stem cells with naïve features have been derived directly from dissociated ICM cells from day 6 human blastocysts (Guo et al., 2016). Naïve type hPSC have also been obtained following somatic cell reprogramming (Kilens et al., 2018; Liu et al., 2017). The availability of naïve and conventional hPSC provides two complementary systems for modelling early human development. In particular, the global relationship of naïve hPSC to pre-implantation embryo epiblast (Nakamura et al., 2016; Stirparo et al., 2018) provides an opportunity to study and dissect the progression of pluripotency over a time window of human embryogenesis inaccessible in utero.

Current conditions employed to generate naïve hPSC show variation in efficiency, especially when applied across different cell lines (Guo et al., 2017). Reliable cell surface markers would aid identification and establishment of naïve hPSC and facilitate optimisation of culture conditions and procedures. To date several surface markers have been reported that may distinguish naïve from conventional hPSC. Some of these are expressed by conventional hPSC and not by naïve hPSC (O'Brien et al., 2017; Pastor et al., 2016; Shakiba et al., 2015) while others show only differences in level between the two PSC types (Liu et al., 2017; O'Brien et al., 2017). On the other hand, markers claimed to be specific for naïve PSC (Collier et al., 2017) show either no expression or broad

expression in early human embryos, challenging their relevance for identification of the naïve pluripotent phenotype.

Here we present Sushi Containing Domain 2 (SUSD2) (Sugahara et al., 2007) as a robust cell surface marker of naïve pluripotency in the human embryo and hPSC in vitro. We identified high enrichment of *SUSD2* in naïve pre-implantation epiblast through analysis of differential gene expression in human embryos. We evaluated SUSD2 protein expression by antibody staining of human blastocysts and of naïve and conventional PSC cultures. Finally we investigated the applicability of SUSD2 live cell staining and flow cytometry during resetting and reprogramming to naïve PSC status.

RESULTS

Sushi Containing Domain 2 (SUSD2) is a marker for naïve pluripotency

To identify candidate markers for human naïve pluripotent cells we scanned integrated single cell RNA-seq datasets from early human embryos (Stirparo et al., 2018) for transmembrane proteins differentially expressed in the pre-implantation epiblast. We observed that *Sushi Containing Domain 2 (SUSD2)* is highly enriched in the ICM and with a mean level in embryonic day (E) 6-7 epiblast, 8-fold higher than in primitive endoderm or trophectoderm (p-value < 1.18^{-10}) (Figure 1A). In contrast, other recently reported surface markers for naïve cells are either not detectable or not specific to epiblast (Figure S1).

SUSD2 is a type I membrane protein with a large extracellular domain (Sugahara et al., 2007). Several commercial antibodies are available against SUSD2 (Sivasubramaniyan et al., 2013). We therefore examined whether SUSD2 protein expression reflects transcript distribution. We immunostained E7 human embryos using a monoclonal antibody against SUSD2. Intense cell surface staining was observed on a subset of cells within the ICM (Figure 1B, S2A). These SUSD2 positive cells co-express the transcription factor KLF17, denoting human naïve epiblast identity (Blakeley et al., 2015; Guo et al., 2016; Stirparo et al., 2018). In contrast SUSD2 staining was faint in trophectoderm cells and absent in GATA4-positive hypoblast cells.

We then inspected publicly available hPSC transcriptome data (Gafni et al., 2013; Guo et al., 2017; Guo et al., 2016; Stirparo et al., 2018; Takashima et al., 2014; Theunissen et al., 2014; Ware et al., 2014; Yang et al., 2017). We found that *SUSD2* transcript levels are appreciable only in cells cultured in either t2iLGö or 5iLAF medium that satisfy stringent criteria for naïve pluripotent features (Davidson et al., 2015; Huang et al., 2014; Nakamura et al., 2016; Stirparo et al., 2018; Takashima et al., 2014; Theunissen et al., 2016; Theunissen et al., 2014) (Figure 1C). *SUSD2* mRNA is very low or absent in conventional or other hPSC, including cultures in NHSM (Gafni et al., 2013) and so-called

extended pluripotent stem (EPS) cells (Yang et al., 2017). These observations indicate that SUSD2 expression is a discriminating marker for naïve hPSC.

We therefore investigated the utility of SUSD2 antibodies for discriminating human PSC phenotypes. Flow cytometry analysis showed no detectable expression in conventional PSC (Figure 1D). In contrast, SUSD2 was expressed unimodally at high levels in embryo-derived HNES1 naïve PSC (Guo et al., 2016). This was the case both for cultures in the original t2iLGö formulation (Takashima et al., 2014), and in a modified version, PXGL (Guo et al., 2017), including the tankyrase inhibitor XAV939 and omitting GSK3 inhibition (for details see Methods) (Figure 1D & S2B). SUSD2 was also highly expressed in chemically reset (cR) naïve hPSC in PXGL medium (Figure S2C). Comparative flow cytometry analysis with other reported naïve cell surface markers (Collier et al., 2017) revealed only CD75 to have a similar profile to SUSD2, while other markers did not effectively discriminate naïve from conventional PSCs, or were weakly expressed in naïve PSC (Figure S2C).

We noted strong in situ cell surface staining of naïve PSC using a conjugated SUSD2 monoclonal antibody (Figure 1E). Importantly, live staining did not perturb cell viability or morphology and naïve cells could subsequently be expanded without consequence. With the exception of heterogeneous staining for CD7, in situ staining was not detected using conjugated antibodies for CD75 or other reported naïve markers (Collier et al., 2017) (Figure S2C). We also evaluated SUSD2 immunostaining after parafomaldehyde fixation of cR-S6 and HNES1 naïve cells (Figure 1F). We detected no signal on conventional cells but intense surface staining of naïve PSC. SUSD2 immuno-positive cells co-expressed the naïve transcription factor TFCP2L1 and the primate-specific naïve transcription factor KLF17 (Blakeley et al., 2015; Guo et al., 2016; Takashima et al., 2014).

To initiate multi-lineage differentiation, naïve PSC must be converted to a state approaching conventional PSC, a process we have termed capacitation (Smith, 2017). For human naïve PSC,

capacitation is achieved by withdrawal of t2iLGö or PXGL and culture for 8-10 days in N2B27 medium supplemented with XAV939 (Rostovskaya et al., accepted) (Figure S3A). SUSD2 expression is progressively down regulated during capacitation and is absent by day 8 (Figure 1G). Change in cell identity after capacitation is confirmed by loss of colony forming ability in naïve cell medium (Figure S3B). We have shown elsewhere that naïve hPSC capacitation reflects in vivo progression of the epiblast from ICM to late post-implantation in the embryo of the non-human primate *Macaca fascicularis* (Nakamura et al., 2016; Rostovskaya et al., accepted). We examined expression of *SUSD2* in the *Macaca* dataset and found *SUSD2* mRNA expression in pre-implantation epiblast but not detected in post-implantation stages (Figure 1H). Thus SUSD2 expression in PSC is closely linked to naïve status both in cultured stem cells and in the primate embryo.

SUSD2 identifies naïve PSC after chemical resetting

Conventional hPSC can be reset to naïve status by short-term exposure to the histone deactylase inhibitor valproic acid followed by culture in naïve medium (Guo et al., 2017). Naïve hPSC are then enriched by continuous passaging. We investigated whether SUSD2 could be utilised to identify naïve cells upon chemical resetting (Figure 2A). We used SHEF6 (S6)-EOS and H9-EOS conventional hPSC in which the *PB-EOS-GFP* transgene provides a reporter of naïve status (Guo et al., 2017; Takashima et al., 2014). Live cell imaging on day 10 of resetting revealed co-expression of EOS-GFP and SUSD2 in emerging domed naïve-type colonies (Figure 2B).

We performed flow cytometry analyses with SUSD2 in combination with the conventional PSC marker, CD24 (Shakiba et al., 2015). SUSD2$^+$CD24$^-$ and SUSD2$^-$CD24$^+$ cell surface phenotypes unambiguously distinguish naïve from conventional hPSC (Figure 2C). At the start of resetting, all PSCs are SUSD2$^-$CD24$^+$. By day 10, a substantial proportion (20-30%) of cells are SUSD2$^+$ and either CD24 low or CD24 negative (Figure 2D). By day 14, most SUSD2$^+$ cells are negative for CD24 and

SSEA4 (Figure 2D, 2E), another marker of conventional PSC that is absent on naïve hPSCs (Pastor et al., 2016). The vast majority of SUSD2$^+$ (>95%) cells at this stage co-express EOS-GFP (Figure 2E).

Immunostaining of S6-EOS and H9-EOS cultures at day 14 of resetting showed co-expression of SUSD2 with transcription factors KLF17 and NANOG (Figure 2F). SUSD2 negative cells did not express either of these factors. We purified SUSD2$^+$CD24$^-$ populations at day 10 and 14 for further analysis of marker expression by RT-qPCR. Expression of naïve markers *KLF4, KLF17, TFCP2L1, DPPA3* and *DPPA5* was restricted to SUSD2$^+$CD24$^-$ cells and absent/low in the SUSD2$^-$CD24$^+$ fraction (Figure 2G).

Together these observations indicate that SUSD2 staining in combination with absence of CD24 or SSEA4 identifies reset naïve PSC. We therefore investigated whether SUSD2 antibody staining could be utilised to purify naïve PSC from resetting cultures. We fractionated SUSD2$^+$CD24$^-$ and SUSD2$^-$CD24$^+$ populations by flow cytometry on day 14 of resetting and plated in PXGL (Figure 2H, S4A). Analysis 5 days later showed that the great majority of cells retained a cell surface phenotype consistent with their sort profile (Figure 2I, S4A). Consistent with the flow profile, SUSD2$^+$CD24$^-$ sorted cells generated numerous EOS-GFP$^+$ dome-shaped colonies with few if any other colony types apparent. In contrast SUSD2$^-$CD24$^+$ cells yielded predominantly heterogeneously differentiated cells that were EOS-GFP$^-$, with only the occasional GFP$^+$ domed colony (Figure 2J, S4A).

The SUSD2$^+$CD24$^-$ sorted cultures remained morphologically homogeneous and maintained EOS-GFP expression upon passaging (Figure 2J, S4A). RT-qPCR analysis at passage (P) 1 and P3 after sorting showed that expression of naïve markers *KLF4, KLF17, TFCP2L1* and *DPPA5* was sustained in SUSD2$^+$CD24$^-$ derived cells but remained negligible in the SUSD2$^-$CD24$^+$ derivatives (Figure S4B). Flow cytometry analysis of SUSD2$^+$CD24$^-$ sorted cells at later passages showed maintenance of homogeneous SUSD2$^+$ cells (>98%). In contrast, parallel unsorted reset cultures retained sub-

populations of both SUSD2^{lo}CD24$^+$ and SUSD2$^-$CD24$^+$ cells in addition to the majority SUSD2$^+$CD24$^-$ population (Figure S4C).

SUSD2 sorting facilitates establishment of naïve PSC cultures

The efficiency of chemical or transgene-driven resetting from conventional to naïve PSC is variable and extended passaging may be required to establish homogeneous cultures (Guo et al., 2017; Takashima et al., 2014). We utilised SUSD2 to monitor efficiency and purify naïve cells during resetting of diverse ESC and iPSC lines. We applied the chemical resetting protocol to two conventional hESC, H1 and H7 (Thomson et al., 1998) and two iPSC, MeCP2-clone17 and NCRM2. Flow cytometry analysis of resetting cultures at day 10, day 14, P1 and P3 revealed that SUSD2$^+$CD24$^-$ populations appeared with different frequencies and kinetics (Figure 3A). Resetting efficiency, quantified by SUSD2$^+$CD24$^-$ cell fraction at P3, ranged from 22% to 98% across the lines (Figure 3A). The identity of SUSD2$^+$ cells was evaluated by RT-qPCR analysis. As previously this showed expression of naïve markers *KLF4, KLF17, TFCP2L1* and *DPPA5* is restricted to SUSD2$^+$CD24$^-$ cells (Figure 3B).

We purified SUSD2$^+$CD24$^-$ cells by flow cytometry at P3. Replating resulted in uniform cultures of naïve type colonies. In contrast, unsorted cultures displayed persisting heterogeneity (Figure 3C). Immunostaining showed co-expression of SUSD2, NANOG and KLF17 throughout naive cultures after sorting (Figure 3D). Post-sort cultures were expanded for 10 passages in PXGL and retained domed colony morphology.

MeCP2-clone17 cells are heterozygous for a loss-of-function mutation in the X-linked gene *MECP2* (Lee et al., 2001; Sahakyan et al., 2017). MeCP2 protein is not expressed in conventional MeCP2-clone17 cells because the wild type allele is on the silent X chromosome. Upon resetting, the silent X

is expected to be reactivated (Guo et al., 2017). Consistent with this, immunostaining revealed co-expression of MeCP2 with SUSD2 in sorted reset cultures, providing further evidence of naïve status (Figure 3E).

These findings demonstrate the utility of SUSD2$^+$ staining during chemical resetting, bypassing the requirement for a transgenic reporter or prolonged passaging to establish homogeneous naïve cultures from different starting hPSC lines, even when the initial frequency of resetting is poor.

SUSD2 identifies emerging naïve PSC during somatic cell reprogramming

Naïve hPSC can be generated from somatic cells by molecular reprogramming (Giulitti et al., 2018; Kilens et al., 2018; Liu et al., 2017). We used Sendai viral (SeV) vectors to reprogram human diploid fibroblasts (Figure 4A) (Fusaki et al., 2009). Around one week following transfection and culture in medium with basic fibroblast growth factor (FGF2), we observed small patches of cells undergoing mesenchymal to epithelial transition. On day 8 culture medium was exchanged to PXGL supplemented with ROCK inhibitor Y-27632 (Figure 4A). Six days following transfer to PXGL SUSD2 positive cells emerged (Figure 4B). By 10 days multiple SUSD2 positive domed colonies were apparent. Flow cytometry analysis showed that the majority of cells in the culture at day 10 are SUSD2$^+$ and CD24$^-$ (Figure 4C). These cultures can be bulk passaged and further propagated to establish naïve iPSC.

We also examined SUSD2 expression during reprogramming using episomal pCXLE vectors (Okita et al., 2011) and found that SUSD2 positive cells appeared with similar kinetics as in SeV reprogramming. However, the episomal system produced more heterogeneous cultures (Figure S4D). We therefore used EpCAM to exclude any mesenchymal cells that may express SUSD2 (Sivasubramaniyan et al., 2013). Flow cytometry analysis identified an EpCAM$^+$SUSD2$^+$CD24$^-$

population 7 days after transfer to PXGL that increased to 13% of the culture by day 14 (Figure S4E). Immunostaining of reprogrammed cultures on day 14 showed co-expression of SUSD2 with KLF17 in both SeV and episomal systems (Figure 4D, S4F). RT-qPCR analysis of isolated SUSD2$^+$CD24$^-$ populations in both cases confirmed expression of naive pluripotency markers at similar levels to embryo-derived HNES1 naïve PSC (Figure 4E). These results demonstrate that SUSD2 marks emerging naïve iPSC during somatic cell reprogramming.

DISCUSSION

SUSD2 is a type I membrane protein with a large extracellular region (Sugahara et al., 2007). The function of SUSD2 is uncertain. It is expressed by various cell types in development and adulthood, and has been used to fractionate both pancreatic and mesenchymal progenitors (Masuda et al., 2012; Ramond et al., 2017; Sivasubramaniyan et al., 2013). However, it has not previously been associated with pluripotent cells, *in vivo* or *in vitro*.

Our findings reveal that SUSD2 is highly enriched in naïve epiblast in the human blastocyst and uniformly expressed on the cell surface of naïve hPSC but absent from other hPSC. SUSD2 is lost as naïve hPSC transition toward differentiation, consistent with transcriptional down-regulation observed in the post-implantation cynomolgus embryo. Future investigation will determine whether the restricted expression of SUSD2 has functional consequence for naïve pluripotency.

SUSD2 antibody binding provides non-invasive labelling for live cell imaging and flow cytometric cell sorting with no evident deleterious effects. SUSD2 strongly stains naïve hPSC in situ and allows unambiguous discrimination of naïve from conventional hPSC by flow cytometry. Several candidate markers for distinguishing naïve hPSC have recently been reported (Collier et al., 2017; Liu et al., 2017; O'Brien et al., 2017; Pastor et al., 2016; Shakiba et al., 2015) and reviewed (Trusler et al.,

2018). Among these F11R is highly expressed by naïve hPSC (O'Brien et al., 2017), but is also present at substantial levels on conventional cells (Liu et al., 2017) (Figure S1). Of the other markers present on naïve hPSC, CD75, CD77 and CD130 display broad mRNA expression and antibody staining in E6 embryos while CD7 mRNA is not detected in the human embryo (Collier et al., 2017). CD7 and CD77 do not resolve naïve hPSC from conventional PSC by flow cytometry, while CD130 expression is weak. CD75, CD77 and CD130 did not stain naïve PSC in situ while CD7 stained live cells heterogeneously. In contrast, SUSD2 decisively labels naïve hPSC, and provides assurance of embryo lineage and hPSC classification not offered by current markers.

Live cell staining or flow cytometry analysis for SUSD2 provide simple and reliable means to routinely monitor hPSC culture status. This is useful because culture conditions for naïve hPSC have yet to be fully optimised and cells can become heterogeneous during maintenance. SUSD2 is particularly valuable for facilitating the establishment of naïve hPSC by resetting or reprogramming procedures. Current protocols are relatively inefficient and extended passaging may be required to enrich for the naïve phenotype. Sorting for SUSD2$^+$ and against a conventional hPSC marker such as CD24 or SSEA4 allows efficient purification from mixed resetting cultures and thus more rapid establishment of naïve PSC lines.

In summary, our findings illustrate the utility of SUSD2 antibody staining for classification, quantification and isolation of naïve hPSC. We envisage that SUSD2 staining will be a useful tool for optimising resetting and reprogramming conditions. It will also be of future interest to determine whether SUSD2 plays a significant biological role in naïve pluripotent cells.

EXPERIMENTAL PROCEDURES

Human embryos

Human embryo research was licensed by the UK Human Fertilization and Embryology Authority. Supernumerary embryos donated from *in vitro* fertilization programs with informed consent thawed at day 5 or 6 post fertilization directly into N2B27 and cultured until fixation at day 7 post fertilization. Blastocysts were fixed in 4% PFA, immunostained, and imaged as described (Takashima et al., 2014).

Cell culture

Conventional hPSC cultures were propagated on Geltrex (growth factor-reduced, Thermo Fisher, A1413302) in Essential 8 (E8) medium made in-house (Chen et al., 2011). Chemically reset (cR) or embryo-derived (HNES1) naïve PSC were propagated in N2B27 with PXGL Cells were cultured in 5% O_2, 7% CO_2 in a humidified incubator at 37°C and passaged by dissociation with Accutase (Thermo Fisher Scientific, A1110501) every 3-5 days. Lines used: S6, S6-EOS, H9-EOS, H1, H7, NCRM2, MECP2-clone17 (Sahakyan et al., 2017). Cell lines were confirmed free of mycoplasma contamination by periodic in-house PCR assay. For capacitation, cells were passaged once without feeders in PXGL medium then exchanged into N2B27 containing 2µM XAV939 (Rostovskaya et al., accepted). Chemical resetting were performed as described (Guo et al., 2017) with minor modifications. Somatic cell reprogramming were performed either with Sendai virus vector system (DNAVEC) (FUSAKI et al., 2009) or with episomal vectors, (*pCXLE-OCT4-shRNA(p53)*, *pCXLE--SOX2-KLF4* and *pCXLE-L-MYC-LIN28* (Okita et al., 2011). See detailed procedures in supplemental methods.

Author Contributions

Conceptualization, N.B., G.G and A.S.; Investigation, N.B., G.G., J.N.; Methodology, N.B, G.G.; Formal analysis, G.G.S.; Writing, A.S, N.B. and G.G.; Supervision G.G., J.N. and A.S.

Acknowledgements

Competing Interests

NB, AS and GG are inventors on a patent application relating to this work filed by the University of Cambridge.

REFERENCES

Blakeley, P., Fogarty, N.M.E., del Valle, I., Wamaitha, S.E., Hu, T.X., Elder, K., Snell, P., Christie, L., Robson, P., and Niakan, K.K. (2015). Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. Development.

Chen, G., Gulbranson, D.R., Hou, Z., Bolin, J.M., Ruotti, V., Probasco, M.D., Smuga-Otto, K., Howden, S.E., Diol, N.R., Propson, N.E.*, et al.* (2011). Chemically defined conditions for human iPSC derivation and culture. Nat Methods *8*, 424-429.

Collier, A.J., Panula, S.P., Schell, J.P., Chovanec, P., Plaza Reyes, A., Petropoulos, S., Corcoran, A.E., Walker, R., Douagi, I., Lanner, F.*, et al.* (2017). Comprehensive Cell Surface Protein Profiling Identifies Specific Markers of Human Naive and Primed Pluripotent States. Cell Stem Cell *20*, 874-890 e877.

Davidson, K.C., Mason, E.A., and Pera, M.F. (2015). The pluripotent state in mouse and human. Development *142*, 3090-3099.

Fusaki, N., Ban, H., Nishiyama, A., Saeki, K., and Hasegawa, M. (2009). Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. Proc Jpn Acad Ser B *85*, 348-362.

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A.*, et al.* (2013). Derivation of novel human ground state naive pluripotent stem cells. Nature *504*, 282-286.

Giulitti, S., Pellegrini, M., Zorzan, I., Martini, P., Gagliano, O., Mutarelli, M., Ziller, M.J., Cacchiarelli, D., Romualdi, C., Elvassore, N.*, et al.* (2018). Direct generation of human naive induced pluripotent stem cells from somatic cells in microfluidics. Nat Cell Biol.

Guo, G., von Meyenn, F., Rostovskaya, M., Clarke, J., Dietmann, S., Baker, D., Sahakyan, A., Myers, S., Bertone, P., Reik, W.*, et al.* (2017). Epigenetic resetting of human pluripotency. Development *144*, 2748-2763.

Guo, G., von Meyenn, F., Santos, F., Chen, Y., Reik, W., Bertone, P., Smith, A., and Nichols, J. (2016). Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. Stem Cell Reports *6*, 437-446.

Huang, K., Maruyama, T., and Fan, G. (2014). The naive state of human pluripotent stem cells: a synthesis of stem cell and preimplantation embryo transcriptome analyses. Cell Stem Cell *15*, 410-415.

Kilens, S., Meistermann, D., Moreno, D., Chariau, C., Gaignerie, A., Reignier, A., Lelièvre, Y., Casanova, M., Vallot, C., Nedellec, S.*, et al.* (2018). Parallel derivation of isogenic human primed and naive induced pluripotent stem cells. Nature Communications *9*, 360.

Lee, S.S., Wan, M., and Francke, U. (2001). Spectrum of MECP2 mutations in Rett syndrome. Brain Dev *23 Suppl 1*, S138-143.

Liu, X., Nefzger, C.M., Rossello, F.J., Chen, J., Knaupp, A.S., Firas, J., Ford, E., Pflueger, J., Paynter, J.M., Chy, H.S.*, et al.* (2017). Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. Nature Methods *14*, 1055-1062.

Masuda, H., Anwar, S.S., Buhring, H.J., Rao, J.R., and Gargett, C.E. (2012). A novel marker of human endometrial mesenchymal stem-like cells. Cell Transplant *21*, 2201-2214.

Nakamura, T., Okamoto, I., Sasaki, K., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Yamamoto, T., and Saitou, M. (2016). A developmental coordinate of pluripotency among mice, monkeys and humans. Nature *537*, 57-62.

O'Brien, C.M., Chy, H.S., Zhou, Q., Blumenfeld, S., Lambshead, J.W., Liu, X., Kie, J., Capaldo, B.D., Chung, T.L., Adams, T.E.*, et al.* (2017). New Monoclonal Antibodies to Defined Cell Surface Proteins on Human Pluripotent Stem Cells. Stem Cells *35*, 626-640.

O'Leary, T., Heindryckx, B., Lierman, S., van Bruggen, D., Goeman, J.J., Vandewoestyne, M., Deforce, D., de Sousa Lopes, S.M., and De Sutter, P. (2012). Tracking the progression of the human inner cell mass during embryonic stem cell derivation. Nat Biotechnol *30*, 278-282.

Okita, K., Matsumura, Y., Sato, Y., Okada, A., Morizane, A., Okamoto, S., Hong, H., Nakagawa, M., Tanabe, K., Tezuka, K.*, et al.* (2011). A more efficient method to generate integration-free human iPS cells. Nat Methods *8*, 409-412.

Pastor, W.A., Chen, D., Liu, W., Kim, R., Sahakyan, A., Lukianchikov, A., Plath, K., Jacobsen, S.E., and Clark, A.T. (2016). Naive Human Pluripotent Cells Feature a Methylation Landscape Devoid of Blastocyst or Germline Memory. Cell Stem Cell *18*, 323-329.

Ramond, C., Glaser, N., Berthault, C., Ameri, J., Kirkegaard, J.S., Hansson, M., Honore, C., Semb, H., and Scharfmann, R. (2017). Reconstructing human pancreatic differentiation by mapping specific cell populations during development. Elife *6*.

Rossant, J., and Tam, P.P.L. (2017). New Insights into Early Human Development: Lessons for Stem Cell Derivation and Differentiation. Cell Stem Cell *20*, 18-28.

Rostovskaya, M., Stirparo, G.G., and Smith, A. (accepted). Capacitation of human naïve pluripotent stem cells for multi-lineage differentiation. Development.

Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T.W., Kuoy, E., Langerman, J., Clark, A.T., Jaenisch, R.*, et al.* (2017). Human Naive Pluripotent Stem Cells Model X Chromosome Dampening and X Inactivation. Cell Stem Cell *20*, 87-101.

Shakiba, N., White, C.A., Lipsitz, Y.Y., Yachie-Kinoshita, A., Tonge, P.D., Hussein, S.M., Puri, M.C., Elbaz, J., Morrissey-Scoot, J., Li, M.*, et al.* (2015). CD24 tracks divergent pluripotent states in mouse and human cells. Nat Commun *6*, 7329.

Sivasubramaniyan, K., Harichandan, A., Schumann, S., Sobiesiak, M., Lengerke, C., Maurer, A., Kalbacher, H., and Buhring, H.J. (2013). Prospective isolation of mesenchymal stem cells from human bone marrow using novel antibodies directed against Sushi domain containing 2. Stem Cells Dev *22*, 1944-1954.

Smith, A. (2017). Formative pluripotency: the executive phase in a developmental continuum. Development *144*, 365-373.

Stirparo, G.G., Boroviak, T., Guo, G., Nichols, J., Smith, A., and Bertone, P. (2018). Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. Development *145*.

Sugahara, T., Yamashita, Y., Shinomi, M., Yamanoha, B., Iseki, H., Takeda, A., Okazaki, Y., Hayashizaki, Y., Kawai, K., Suemizu, H.*, et al.* (2007). Isolation of a novel mouse gene, mSVS-1/SUSD2, reversing tumorigenic phenotypes of cancer cells in vitro. Cancer Sci *98*, 900-908.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W.*, et al.* (2014). Resetting Transcription Factor Control Circuitry toward Ground-State Pluripotency in Human. Cell *158*, 1254-1269.

Theunissen, T.W., Friedli, M., He, Y., Planet, E., O'Neil, R.C., Markoulaki, S., Pontis, J., Wang, H., Iouranova, A., Imbeault, M.*, et al.* (2016). Molecular Criteria for Defining the Naive Human Pluripotent State. Cell Stem Cell *19*, 502-515.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L.*, et al.* (2014). Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell *15*, 471-487.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Trusler, O., Huang, Z., Goodwin, J., and Laslett, A.L. (2018). Cell surface markers for the identification and study of human naive pluripotent stem cells. Stem Cell Res *26*, 36-43.

Ware, C.B., Nelson, A.M., Mecham, B., Hesson, J., Zhou, W., Jonlin, E.C., Jimenez-Caliani, A.J., Deng, X., Cavanaugh, C., Cook, S.*, et al.* (2014). Derivation of naive human embryonic stem cells. Proc Natl Acad Sci U S A *111*, 4484-4489.

Yang, Y., Liu, B., Xu, J., Wang, J., Wu, J., Shi, C., Xu, Y., Dong, J., Wang, C., Lai, W.*, et al.* (2017). Derivation of Pluripotent Stem Cells with In Vivo Embryonic and Extraembryonic Potency. Cell *169*, 243-257 e225.

Figure legends

Figure 1. SUSD2 is expressed by human naïve PSCs in the embryo and in culture

(A) *SUSD2* transcript levels in human embryos at different stages and lineages shown, extracted from an integrated single cell RNA-seq dataset (Stirparo et al., 2018). (B) Immunostaining for KLF17, GATA4 and SUSD2 in an E7 human blastocyst. Scale bar, 50µm. (C) *SUSD2* transcript levels in

naïve and conventional PSCs (Stirparo et al., 2018). (D) Flow cytometry analysis of SUSD2 in conventional and naïve cells. (E) Images show bright field and SUSD2 immunostaining, using a SUSD2-PE antibody. Scale bar, 50µm. (F) Immunostaining for SUSD2, TFCP2L1 and KLF17 in conventional and naïve (cR-S6 and HNES1) cells. Scale bar, 100µm. (G) Flow cytometry analysis of SUSD2 expression during capacitation of cR-S6 and HNES1 cells. (H) *SUSD2* transcript levels in *Macaca fasciculari* embryos. Abbreviation: cMOR, compacted morula; eICM, early inner cell mass; TE, trophectoderm; Epi, epiblast; PrE, primitive endoderm. See also Figure S1, S2, S3.

Figure 2. SUSD2 identifies and purifies reset naïve PSCs

(A) Schematic of the chemical resetting protocol. HDACi, histone deacetylase inhibitor. (B) Images of cultures at day 10 of resetting. Scale bar, 50µm. (C) Flow cytometry analysis of SUSD2 and CD24 expression in conventional and naïve PSCs and (D) during resetting. (E) Flow cytometry analysis of GFP and SSEA4 on SUSD2$^+$CD24$^-$ cells at day 14 of resetting. (F) Immunostaining for SUSD2, NANOG and KLF17 at day 14 of resetting. Scale bar, 100µm. (G) RT-qPCR analysis of sorted SUSD2$^+$CD24$^-$ and SUSD2$^-$CD24$^+$ cells at day 10 and day 14 of resetting. Error bars indicate s.d. of three individual experiments. (H) Representative flow cytometry sort plot. (I) Flow cytometry analysis of SUSD2 and CD24 expression on cell populations sorted in (H) five days after sorting. SUSD2$^+$, SUSD2$^+$CD24$^-$; SUSD2$^-$, SUSD2$^-$CD24$^+$. (J) Bright field and GFP images of cell populations sorted in (H) at passage one and five after sorting. Scale bar, 50µm. See also Figure S4.

**Figure 3. SUSD2 identifies reset naïve PSC in multiple cell lines**

(A) Flow cytometry analysis of SUSD2 and CD24 expression during resetting of two ESC lines (H1 and H7) and two iPSC lines (MECP2-clone17 and NCRM2). (B) RT-qPCR analysis of marker expression in sorted SUSD2$^+$CD24$^-$ and SUSD2$^-$CD24$^+$ cells at passage (P) 3. Note: No SUSD2$^-$CD24$^+$ population is evident for the cR-H1 line. *DPPA5* expression is shown at 0.5x actual expression. Error bars indicate s.d. of two individual experiments. (C) Flow cytometry analysis of

16

SUSD2 and CD24 expression and bright field images of reset cultures at P5. Top row, unsorted reset cultures; bottom row, cultures sorted for SUSD2$^+$CD24$^-$ at P3. (D) Immunostaining for SUSD2, NANOG and KLF17 at P6 for cultures that were sorted for SUSD2$^+$CD24$^-$ at P3. Scale bar, 100µm. (E) Immunostaining for NANOG, MECP2 and SUSD2 on reset (P6) and parental MECP2-clone17 cells. Scale bar, 100µm.

Figure 4. SUSD2 identifies naïve PSC during somatic cell reprogramming

(A) Schematic of the reprogramming protocol. (B-D) Sendai vector reprogramming of human dermal fibroblasts. (B) Images of bright field and SUSD2 live staining. Scale bar, 50µm. (C) Flow cytometry analysis of SUSD2 and CD24 expression. (D) Immunostaining for SUSD2, NANOG and KLF17. Scale bar, 50µm. (E) RT-qPCR analysis of markers on sorted SUSD2$^+$CD24$^-$ at day 8+14 for Sendai vector- and episomal-mediated reprogramming, and on established naïve PSCs. Bars indicate mean of two individual experiments. See also Figure S4.

Figure 2

Figure 4

Click here to access/download;Figure;Bredenkamp Fig4 0328.tif ⬇
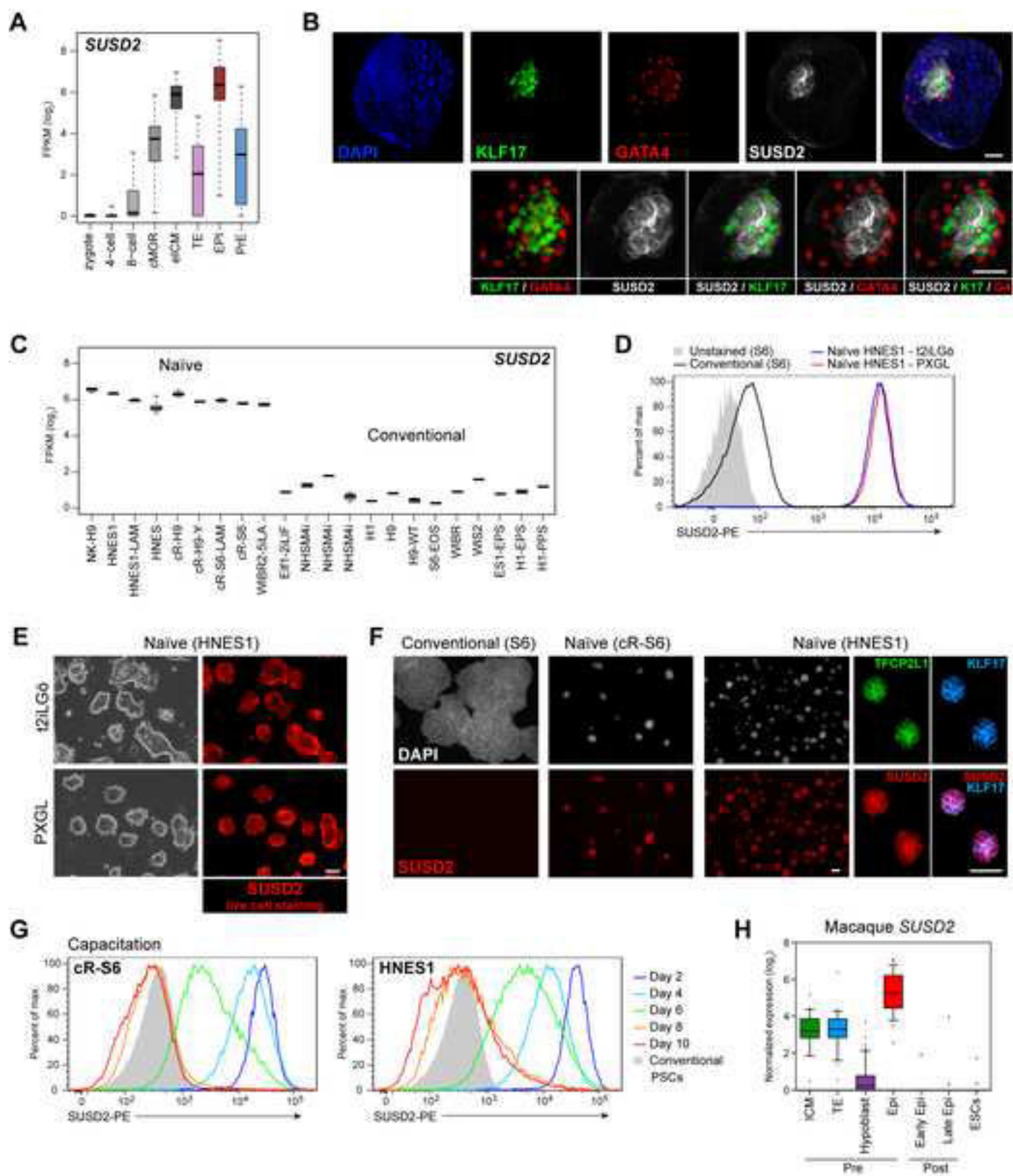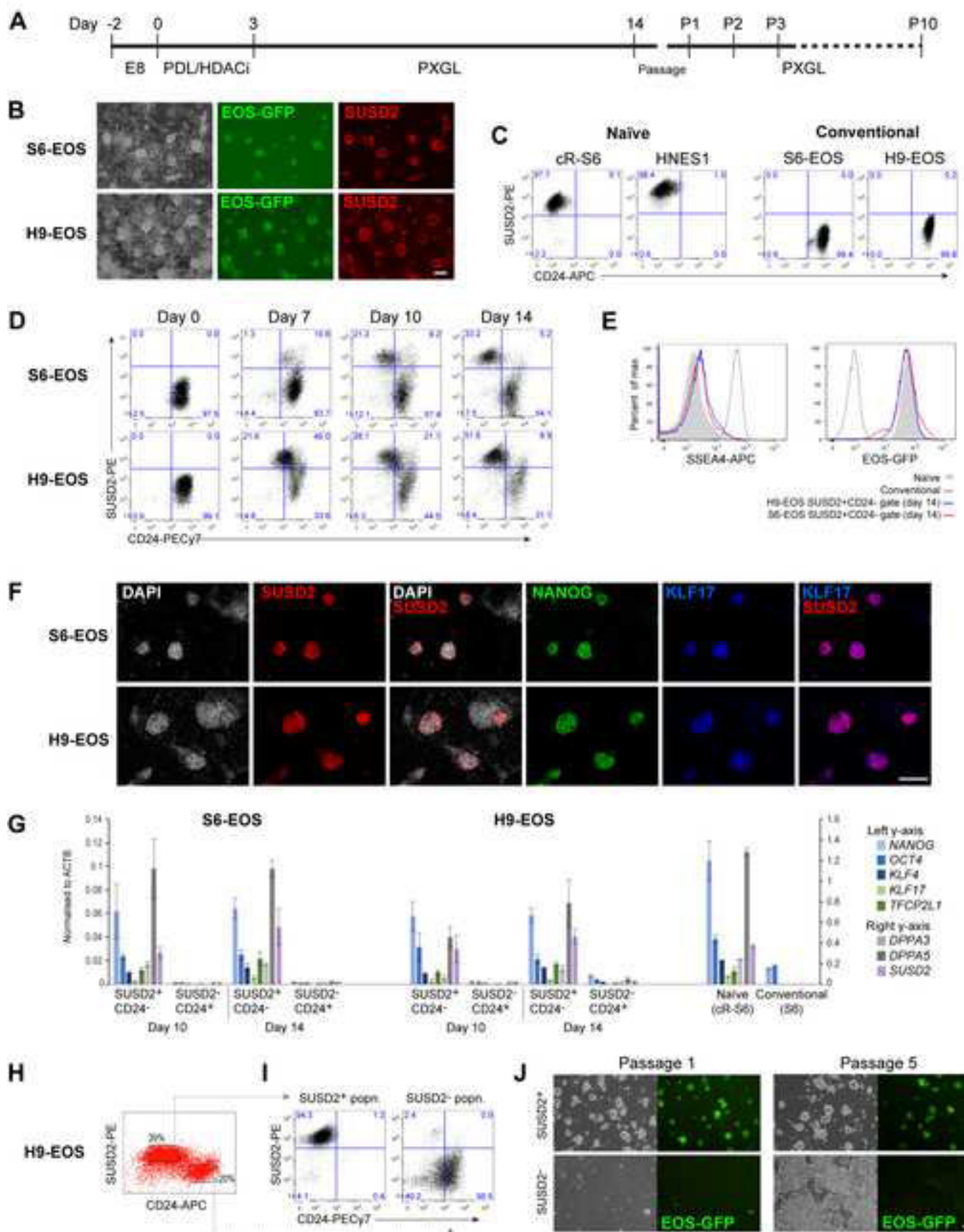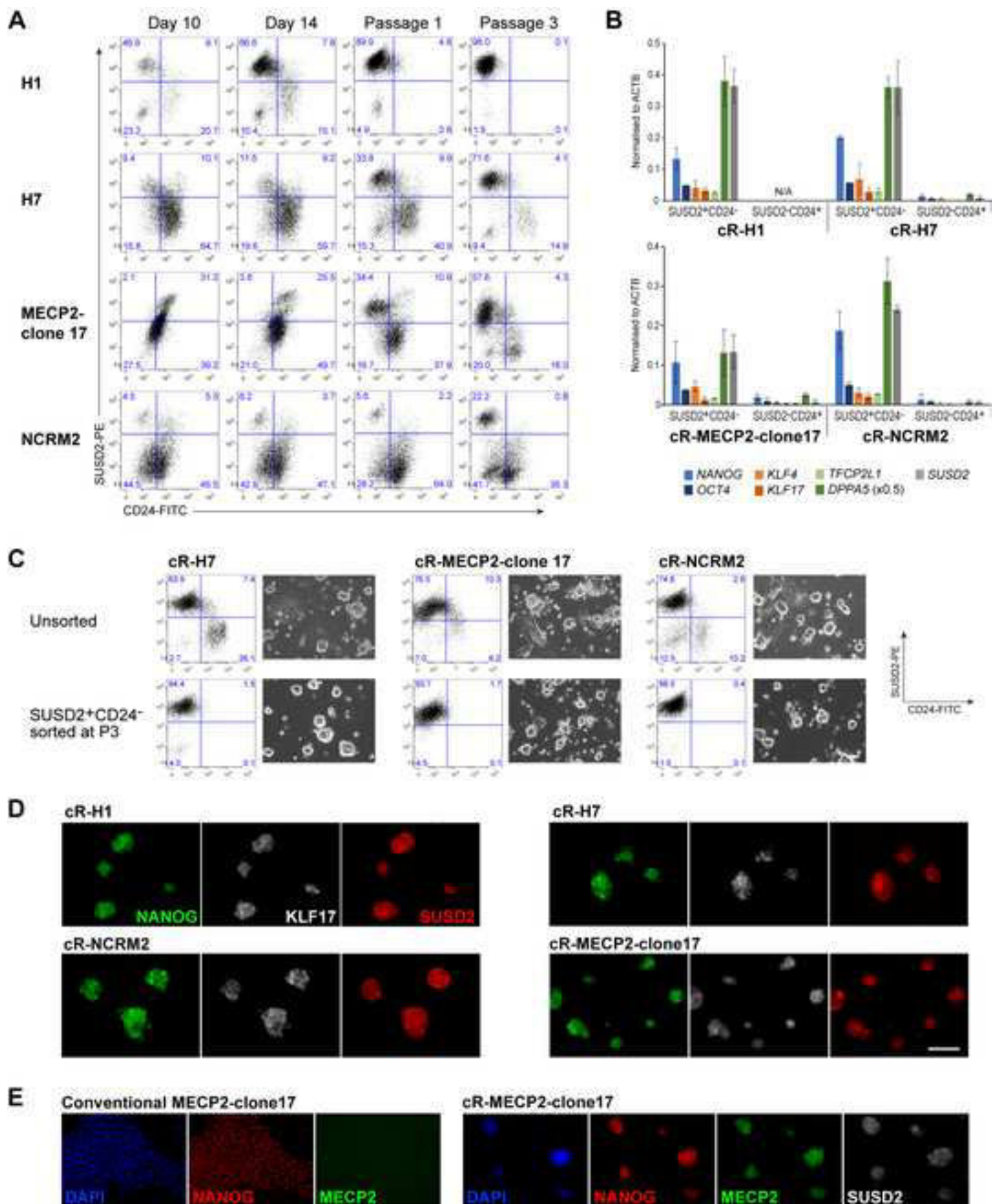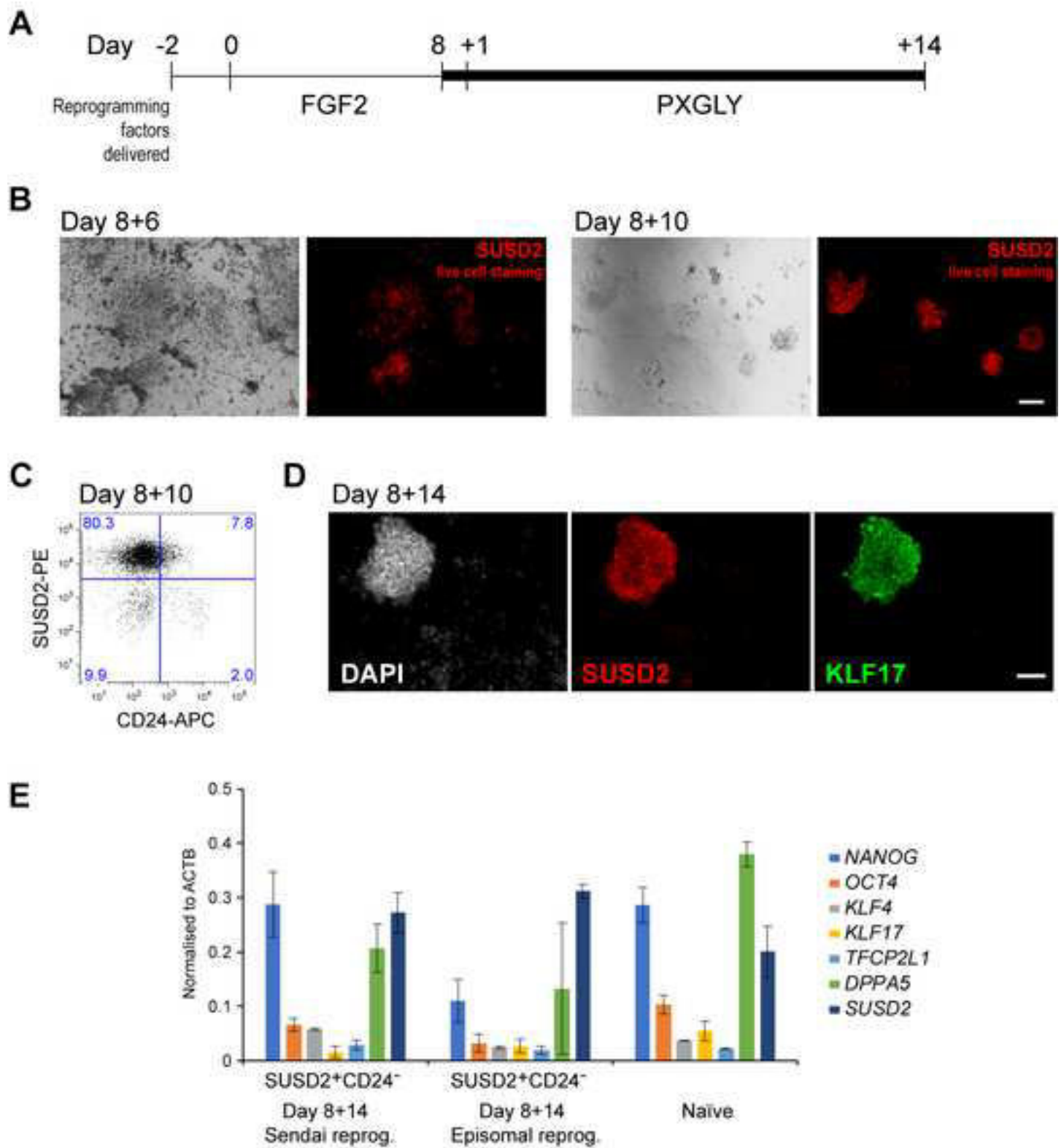
Inventory of Supplemental information

Supplemental methods

Cell culture procedures, reverse transcription and real-time PCR, Immunofluorescence staining and Flow cytometry are described.

Supplemental figure legend

Figure S1 Surface marker expression by RNA-seq analysis, Relates to Figure 1

Figure S2 Surface marker expression in human embryos and naïve PSC, Relates to Figure 1

Figure S3 Capacitation of human naïve PSC, Relates to Figure 1

Figure S4 Application of SUSD2 in identification and purification of human naïve PSC during resetting and reprogramming, Relates to Figure 2 and Figure 4

SUPPLEMENTAL METHODS

Human naïve PSC culture

Human naïve PSC were propagated in N2B27 with PXGL (1 µM PD0325901 (P), 2µM XAV939 (X), 2 µM Gö6983 (G) and 10 ng/ml human LIF (L),] on irradiated MEF feeders with ROCK inhibitor (Y-27632) added during replating. Geltrex was added to media during plating for feeder-free expansion.

Chemical Resetting

Conventional human PSC were seeded at $1 \times 10^4 / cm^2$ onto irradiated MEF feeders in E8 medium. Two days later (day 0) medium was changed to PDL/HDACi (N2B27 with 1µM PD0325901 (P), 10ng/ml human LIF (L) and 1mM valproic acid sodium salt (VPA, Sigma, P4543) (HDACi)). Following 3 days in PDL/HDACi, medium was changed to PXGL and refreshed daily for a further 10-11 days before passaging to establish naïve PSC cultures.

Somatic cell reprogramming

Human dermal fibroblast cells were seeded onto feeder layers at $0.5-1 \times 10^4$ cells/cm^2. The next day OSKM factors were delivered using the Sendai virus vector system (DNAVEC) (FUSAKI et al., 2009). Following 2 days of culture in GMEM/10%FCS, cells were passaged onto fresh feeders and cultured in modified Essential 7 (E7) with 10ng/ml FGF2 (prepared in-house) for 8 days. Cells were then switched into naïve media (PXGLY) to capture naïve iPSCs. Episomal reprogramming was performed with the same protocol as SeV reprogramming except that reprogramming factors (*pCXLE-OCT4-shRNA(p53)*, *pCXLE--SOX2-KLF4* and pCXLE-L-MYC-LIN28 (Okita et al., 2011) were transfected into HDFs by electroporation.

Reverse transcription and real-time PCR

Total RNA was extracted using a RNeasy Kit (Qiagen) and cDNA synthesized with SuperScript III reverse transcriptase (Thermo Fisher Scientific, 18080085) and oligo(dT) adapter primers. TaqMan assays and Universal Probe Library (UPL) probes (Roche Molecular Systems) were used to perform gene quantification.
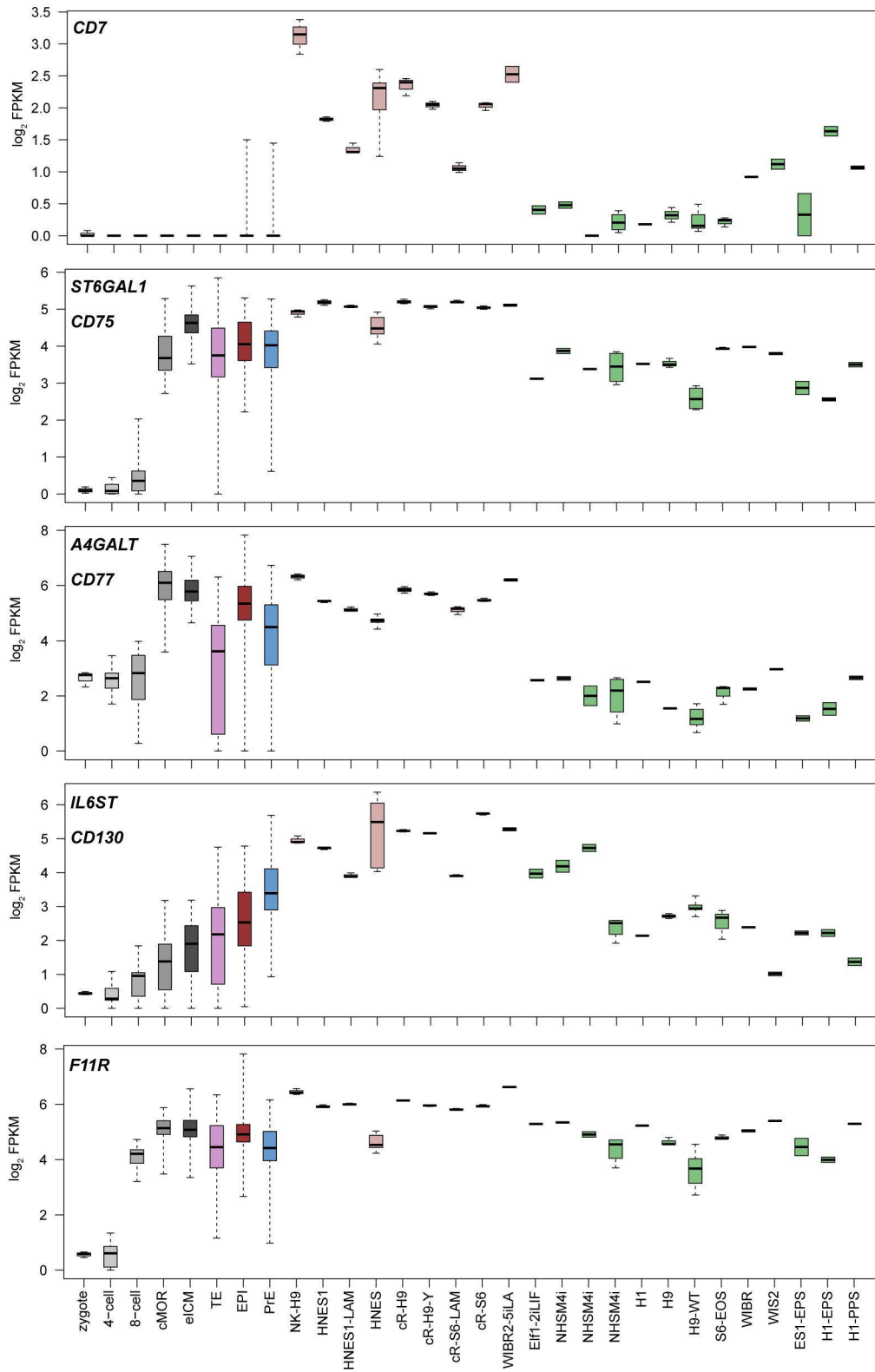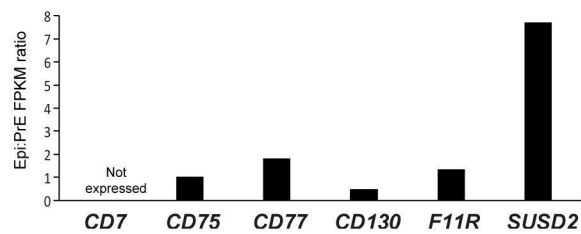
Immunostaining

Cells were fixed with 4% PFA for 15 min at room temperature and blocked/permeabilised with 0.1% Triton X-100 in PBS for 30 min. Incubation with primary antibodies KLF17 (Atlas Antibodies HPA024629), NANOG (R&D Systems AF1997), TFCP2L1 (R&D Systems AF5726), SUSD2 clone W5C5 (BioLegend 327401), GATA4 (Santa Cruz sc-1237), MECP2 (Cell Signalling Technology 3456T) was overnight at 4°C and secondary antibodies were added for 1 h at room temperature. For live cell staining of human cells incubation with conjugated SUSD2 clone W5C5 (SUSD2-PE, BioLegend 327406; SUSD2-FITC, Miltenyi Biotech 130-106-401), CD7-PE clone 6B7 (BioLegend 343105), CD75-FITC clone LN1 (BD Biosciences 555654), CD77-FITC clone 5B5 (BioLegend 357103), CD130-PE clone 2E1B02 (BioLegend 362003) in culture media was for 30 min before washing and imaging.

Flow cytometry

Flow cytometry analysis was carried out on a Fortessa instrument (BD Biosciences). Cell sorting was performed using a MoFlo high-speed instrument (Beckman Coulter). The following antibodies were used for flow cytometry: SUSD2 clone W5C5

(SUSD2-PE, BioLegend 327406; SUSD2-APC, BioLegend 327408) CD24-FITC (BioLegend 311104) or CD24-APC or CD24-PECy7 (eBioscience 17-0247-42 or 25-0247-41), SSEA4-APC (BioLegend 330418), CD7-APC (BioLegend 343108), CD75-eFluor 660 (eBioscience 50-0759-42), CD77-Alexa 647 (BD Biosciences 563632), CD130-PE (BioLegend 362003).

Figure S1

A

	DAPI	KLF17	GATA4	SUSD2	
Embryo 2					
Embryo 3					
Embryo 4					

B

C

Conventional / Naïve

CD7 — 63.3% / 57.8%
CD75 — 99.8% / 100%
CD77 — 95.1% / 59.3%
CD130 — 94.6% / 92.5%
SUSD2 — 100% / 100%

cR-S6 / HNES1

CD7-APC / CD75-eF660 / CD77-A647 / CD130-PE / SUSD2-PE

cR-H9 / HNES1

CD7-PE / CD75-FITC / CD77-FITC / CD130-PE / SUSD2-PE

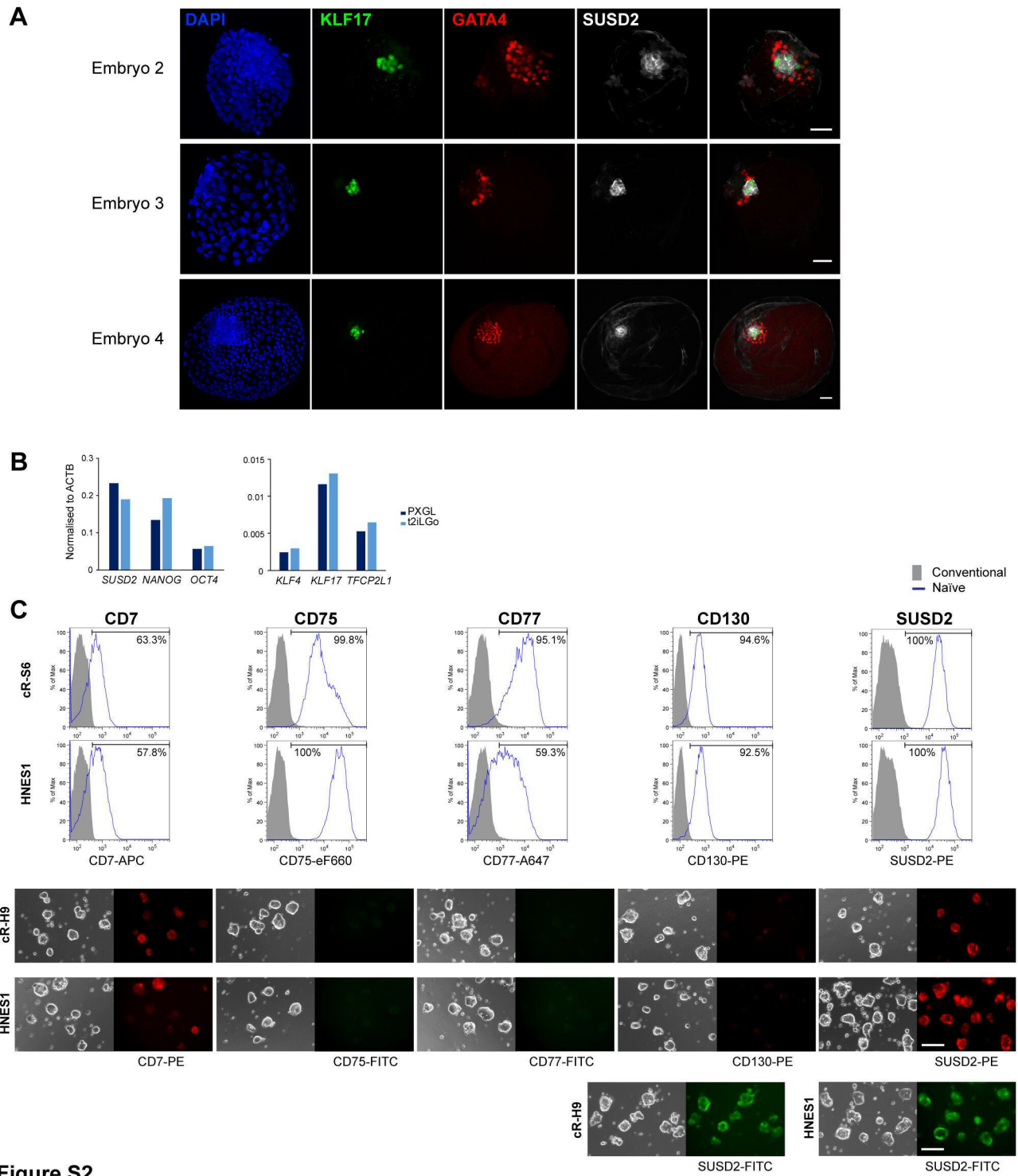cR-H9 / HNES1

SUSD2-FITC / SUSD2-FITC

Figure S2

A

Capacitation

B
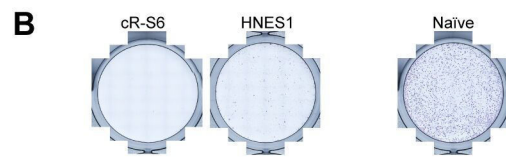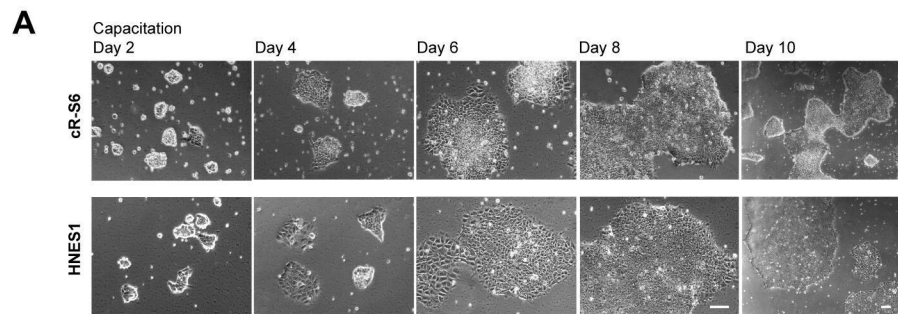
Capacitation for 10 days, replated to PXGL

Figure S3

A

FACS sort at day 14 of restting   5 days post-sort

S6-EOS

SUSD2-PE / CD24-APC

28%
38%

SUSD2+ popn.
95.8   0.5
3.4   0.3
SUSD2-PE / CD24-PECy7

SUSD2- popn.
1.0   1.3
35.4   62.3
CD24-PECy7

Passage 1 post-sort   Passage 5 post-sort

SUSD2+
SUSD2-

EOS-GFP   EOS-GFP

B

cR-S6-EOS   cR-H9-EOS

Normalised to ACTB

Left y-axis
NANOG
OCT4
KLF4
KLF17
TFCP2L1

Right y-axis
DPPA5
SUSD2

SUSD2+ CD24- SUSD2- CD24+ SUSD2+ CD24- SUSD2- CD24+ SUSD2+ CD24- sorted P3
P1   P3   P10

SUSD2+ CD24- SUSD2- CD24+ SUSD2+ CD24- SUSD2- CD24+ SUSD2+ CD24- sorted P3
P1   P3   P10

C

cR-H9-EOS

SUSD2+CD24- sorted at day 14   Not sorted

Passage 5
98.3   0.6
1.1   0.0
90.3   2.8
1.7   5.1

Passage 10
98.8   0.4
0.8   0.0
92.1   2.6
1.2   4.1

SUSD2-PE / CD24-APC

D

Day 8+6   Day 8+14

SUSD2 live cell staining   SUSD2 live cell staining

E

Day 8+7   Day 8+14

SUSD2-PE / EPCAM-PECy7

7.5   13.2

Percent of max.

— Day 8+7 SUSD2+EPCAM+ gate
— Day 8+14 SUSD2+EPCAM+ gate
— Conventional iPSC

CD24-FITC

F

Day 8+14

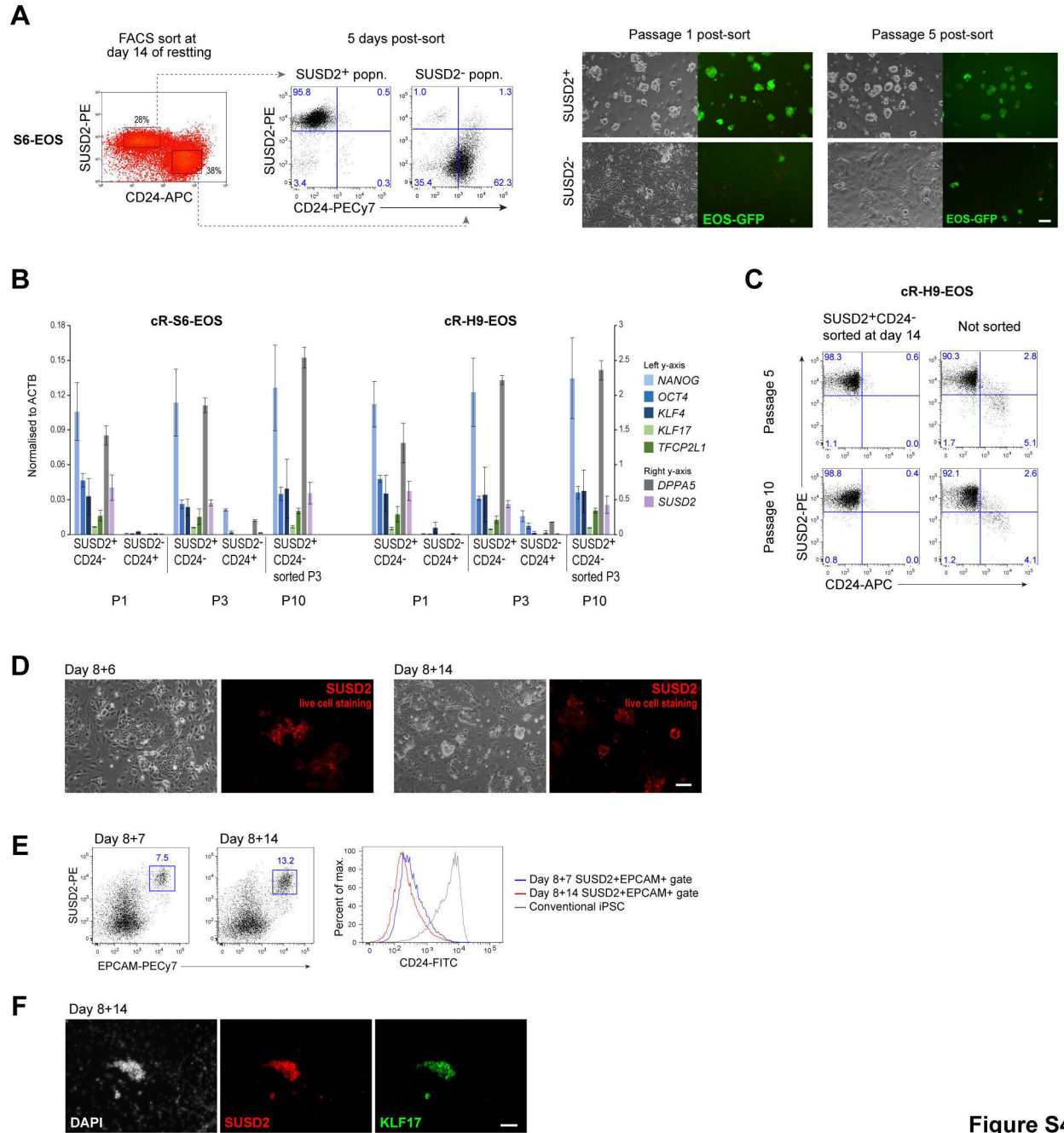DAPI   SUSD2   KLF17

Figure S4

Figure S1 Surface marker expression by RNA-seq analysis, Relates to Figure 1

(A) Human naïve cell surface marker (Collier et al., 2016; O'Brien et al., 2016) transcript levels in human embryos at different stages and lineages shown, and in various naïve (pale pink bars) and conventional (green bars) PSCs extracted from an integrated single cell RNA-sequence dataset (Stirparo et al., 2018). cMOR, compacted morula; eICM, early inner cells mass; TE, trophectoderm; Epi, epiblast; PrE, primitive endoderm. (B) Epiblast to primitive endoderm average FPKM ratio for published cell surface markers and SUSD2.

Figure S2 Surface marker expression in human embryos and naïve PSC, Relates to Figure 1

(A) Immunofluorescence staining for SUSD2, KLF17 and GATA4 in an E7 human blastocyst. KLF17 marks the ICM. Scale bar, 50µm. (B) RT-qPCR analysis of general and naïve pluripotency markers in HNES1 cells in t2iLGö and PXGL medium. (C) Top panel, flow cytometry analysis showing cell surface marker expression in conventional (SHEF6, S6), naïve chemically reset PSCs (cR-S6) and naïve embryo derived PSCs (HNES1). Bottom panel, images show bright field and immunofluorescence staining using conjugated antibodies as shown, on live naïve chemically reset PSCs (cR-H9) and naïve embryo derived PSCs (HNES1). Exposure time: CD7-PE, CD75-FITC, CD77-FITC, CD130-PE, 15s; SUSD2-FITC, 12s; SUSD2-PE, 3s. Scale bar, 100µm

Figure S3 Capacitation of human naïve PSC, Relates to Figure 1

(A) Bright field images show colony morphology changes during capacitation of cR-S6 and HNES1 cells in N2B27-XAV medium. Scale bars, 50µm. (B) Alkaline phosphatase staining of cR-S6 and HNES1 cells following 10 days capacitation and replating and culture in PXGL for 7 days.

Figure S4 Application of SUSD2 in identification and purification of human naïve PSC during resetting and reprogramming, Relates to Figure 2 and Figure 4

(A) Flow cytometry analysis showing SUSD2 and CD24 expression of cell populations five days after sorting. SUSD2$^+$, SUSD2$^+$CD24$^-$; SUSD2$^-$, SUSD2$^-$CD24$^+$. Bright field and GFP images of cell populations at passage one and five after sorting. Scale bar, 50µm. (B) RT-qPCR analysis of general and naïve pluripotency markers in sorted SUSD2$^+$CD24$^-$ and SUSD2$^-$CD24$^+$ cells at P1, P3 and P10 after resetting. Error bars indicate s.d. of three individual experiments. (C) Flow cytometry analysis showing SUSD2 and CD24 expression in cR-H9-EOS cells at passage 5 and 10. Left column shows unsorted reset cultures, bottom row shows cultures that were sorted for SUSD2$^+$CD24$^-$ at day 14 of resetting. (D) Images show bright field and immunofluorescence staining for SUSD2 on live reprogramming cultures by episomal pCXLE vectors. Scale bar, 50µm. (E) Flow cytometry analysis showing SUSD2, EPCAM and CD24 expression on reprogramming cultures. (F) Immunofluorescence staining for SUSD2 and KLF17 on live reprogramming cultures.