 doi:10.20944/preprints202103.0517.v1

Article

Stability of imprinting and differentiation capacity in naïve human cells induced by chemical inhibition of CDK8 and CDK19

Raquel Bernad [1], Cian J. Lynch [1], Rocio G. Urdinguio [2], Camille Stephan-Otto Attolini [3], Mario F. Fraga [2] and Manuel Serrano [1,4,*]

[1] Cellular Plasticity and Disease Group
Institute for Research in Biomedicine (IRB Barcelona)
Barcelona Institute of Science and Technology (BIST)
Barcelona 08028, Spain
[2] Cancer Epigenetics and Nanomedicine Laboratory,
Nanomaterials and Nanotechnology Research Center (CINN-CSIC), Health Research Institute of Asturias (ISPA). Institute of Oncology of Asturias (IUOPA), University of Oviedo, and Rare Diseases CIBER (CIBERER), Oviedo, Spain
[3] Bioinformatics-Biostatistics Unit,
Institute for Research in Biomedicine (IRB Barcelona)
Barcelona Institute of Science and Technology (BIST)
Barcelona 08028, Spain
[4] Catalan Institution for Research and Advanced Studies (ICREA)
Barcelona 08010, Spain
* Correspondence: manuel.serrano@irbbarcelona.org

Abstract: Pluripotent stem cells can be stabilized in vitro at different developmental states by the use of chemicals, factors and supplements. In mice, naïve pluripotent stem cells (PSCs) correspond to the early pre-implantation blastocyst and constitute the optimal starting state for subsequent developmental applications. However, in vitro stabilization of human naïve PSCs remains challenging because of their karyotypic abnormalities, aberrant DNA methylation patterns and severely compromised developmental potency. We have recently developed a novel culture method to induce and stabilize naïve human PSCs that consists in the simple addition of a chemical inhibitor for the closely related CDK8 and CDK19 kinases (CDK8/19i). Long-term cultured CDK8/19i-naïve human PSCs display normal karyotype and do not show widespread DNA demethylation. Here, we investigate the stability of allele-specific methylation at imprinted loci and the differentiation potency of CDK8/19i-naïve human PSCs. We report that CDK8/19i-naïve human PSCs retain their imprinting profile after long-term culture and also after teratoma formation. We have also tested the capacity of CDK8/19i-naïve human PSCs to differentiate into primordial germ cell (PGC)-like cells (PGCLCs) and trophoblast stem cells (TSCs), two cell types that are accessible from the naïve state. Interestingly, CDK8/19i-naïve human PSCs efficiently differentiated into PGCLCs and TSCs. We conclude that the inhibition of CDK8/19 stabilizes human PSCs in a functional naïve state that preserves imprinting and potency.

Keywords: pluripotency; naïve; stem; imprinting; primordial germ cells; trophoblast; CDK8; Mediator; enhancers.

1. Introduction

Mammalian pluripotency spans a continuum of interconvertible states, each with a distinct set of molecular and functional attributes. Naïve pluripotency is functionally and transcriptionally comparable to the cells of the pre-implantation epiblast [1, 2], while primed pluripotency resembles post-implantation epiblast cells [3]. The naïve pluripotent state is efficiently induced and stabilized in vitro by adding two kinase inhibitors for MEK and GSK3 (abbreviated 2i) to the culture medium [4]. Building on the success of 2i to stabilize naïve mouse pluripotent stem cells (PSCs), several groups have reported chemical cocktails that include 2i and that also induce naïve-features in human PSCs [5–11]. Remarkably, while all 2i-based cocktails for the induction of naïve human pluripotency

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2021 doi:10.20944/preprints202103.0517.v1

display an improvement in culture homogeneity, long-term expansion of human PSCs in 2i-based naïve media results in karyotypic instability, global genomic DNA demethylation, gradual loss of imprinting, and deficits in developmental potency [12–16]. These detrimental effects have been attributed to the chronic inhibition of MEK, a kinase whose multiple functions include the promotion of DNA methylation by DNMT1 [17], [18]. In support of this, reduction of the amount of MEK inhibitor partially alleviates the genomic instability of 2i-based naïve human PSCs [19].

Recently, we reported that the naïve state can be induced and stabilized by stimulating the Mediator complex at enhancers through the inhibition of its kinase repression module, composed of the highly similar kinases CDK8 and CDK19 [20, 21]. This study showed that the simple addition of one chemical, a CDK8 and CDK19 inhibitor (CDK8/19i), is sufficient to induce the primed to naïve transition in mouse and human PSCs. Interestingly, the inhibition of CDK8 and CDK19 does not affect genomic DNA methylation and, thereby, does not result in karyotypic abnormalities in human PSCs [20]. Moreover, CDK8/19i-naïve human PSCs efficiently form embryoid bodies and teratomas in mice [20], a property that is generally compromised in 2i-naïve human PSCs [12].

Here, we assess the stability of genomic imprints after long-term culture of human PSCs in the presence of CDK8/19i, and the capacity of CDK8/19i-treated human PSCs to differentiate into primordial germ cell (PGC)-like cells (PGCLCs) and trophoblast stem cells (TSCs).

2. Materials and Methods

2.1. Human PSC resources

WIBR3 and OCT4-ΔPE ESCs were shared by the laboratory of Jacob Hanna (Weizmann Institute of Science). H1 ESCs, and CB5, D2#2, and D2#4 human iPSCs were shared by the laboratory of Nuria Montserrat (IBEC, Institute for Bioengineering).

2.2. Human PSC culture conditions

Human PSCs (hPSCs) were maintained in conventional primed conditions by culture on growth factor-reduced phenol red-free Matrigel (Corning #356231) or human ESC-qualified Matrigel (Corning #354277) with mTeSR1 media (Stem Cell Technologies). Cultures were passaged every 5-7 days manually using 0.5 μM EDTA/1xPBS (Gibco). To reset hPSCs to CDK8/19i-naïve state, cells were maintained on Matrigel (Corning #356231 or #354277) using mTeSR1 (Stem Cell Technologies); the basal media was supplemented with 20 ng/ml of rhLIF (Peprotech), plus 0.4 μM of CNIO-CDK8/19 inhibitor, as reported [20]. Differentiation of OCT4-ΔPE hPSCs as control for the FACS experiments was induced with 10 μM retinoic acid (RA, Sigma) for 5 days [7].

2.3. Derivation of Primordial Germ Cells (PGCs)

Pre-exposure of hPSCs for 4 days in naïve conditions was performed prior to PGC induction. Naïve cocktails used were CDK8/19i [0.4 μM of CNIO-CDK8/19 inhibitor and 20 ng/ml hrLIF (Peprotech)], as reported [20], or the 2i-based PXGL [1 μM PD0325901 (Axon Medchem), 0.3 μM CHIR99021 (Axon Medchem), 2 μM Gö6983 (Selleckchem), 2 μM XAV939 (Selleckchem) and 20 ng/ml hrLIF (Peprotech)] [8]. PGCs-like derivation was performed as previously reported [22]. Briefly, hPSCs were induced to iMeLCs and then into PGC-like cells. For the induction of iMeLCs, hiPSCs were plated at a density of 5x104 cells/cm2 onto a fibronectin-coated (Millipore, FC010) 6-well plate in the GK15 medium [GMEM (Thermo Fisher Scientific) with 15% KSR, 0.1 mM NEAA, 2 mM L-glutamine, 1 mM sodium pyruvate, and 0.1 mM 2-mercaptoethanol] supplemented with 50 ng/ml activin A (R&D), 3 μM CHIR99021 (Axon Medchem) and 10 μM Y-27632 (Selleckchem). After 48 h, the cells (iMeLCs) were dissociated into single cells with TrypLE Select (Thermo Fisher Scientific) and were aggregated in a low-cell-binding V-bottom 96-well plate (Thermo Fisher Scientific, 81100574) at 3000 cells per well in the GK15 medium

 doi:10.20944/preprints202103.0517.v1

supplemented with 200 ng/ml BMP4 (R&D Systems), 100 ng/ml SCF (R&D Systems), 50 ng/ml EGF (R&D Systems), 20 ng/ml hrLIF (Peprotech), and 10 mM Y-27632 to be induced into hPGCLCs.

2.4. Derivation of Trophoblast Stem Cells (TSCs)

hPSCs were pre-treated for 4 days with naïve conditions: CDK8/19i cocktail [0.4 μM of CNIO-CDK8/19 inhibitor and 20 ng/ml hrLIF (Peprotech)], as reported [20], or the 2i-based cocktail PXGL [1 μM PD0325901 (Axon Medchem), 0.3 μM CHIR99021 (Axon Medchem), 2 μM Gö6983 (Selleckchem), 2 μM XAV939 (Selleckchem) and 20 ng/ml hrLIF (Peprotech)] [8]. TSCs derivation was performed as previously reported by Kojima and colleagues [23]. Naïve and primed hPSCs were single-cell dissociated by TrypLE Express, and 0.5 x 106 cells were seeded in a 6-well plated pre-coated with 5 mg/ml Collagen IV and cultured in TS medium [24].

2.5. Teratoma formation assay

Mice were housed at the specific pathogen-free (SPF) barrier area of the Institute for Research in Biomedicine (IRB Barcelona) in Barcelona. Two injections of human H1 and D2#2 PSCs were performed per mouse in testes of SCID beige mice (8 weeks of age, ENVIGO). 2x106 cells in 30μl of mTeSR media were injected per testis. Teratomas were extracted when reaching a size of about 1cm.

2.6. RNA isolation and qRT-PCR

Total RNA from cells was extracted on column by RNeasy kit with DNA digestion following provider's recommendations (Qiagen #74104, #79254) or with Trizol (Invitrogen) according to manufacturer's recommendations. Up to 1 μg of total RNA was retrotranscribed into cDNA using iScriptTM cDNA Synthesis kit (BioRad #170-8891) following manufacturer's protocol. Quantitative real time-PCR (qPCR) was performed using GoTaq® qPCR Master Mix (Promega A6002) in a QuantStudio 6 Flex thermocycler (Applied Biosystem). Input normalization of all the quantitative real time-PCR (qRT-PCR) data was by the ΔΔCt method using the housekeeping gene GAPDH as indicated in each figure. Primer sequences used for mRNA analyses are listed in Table S1.

2.7. RNA-Seq transcriptomic analyses

Samples of total RNA from long-term cultured (>15 passages) human PSCs (D2#2, D2#4, H1, CB5 and WIBR3) in the absence or presence of CDK8/19i. RNA preparations with RIN numbers in the range 9.0 to 10 (Agilent 2100 Bioanalyzer) were used. For library construction 10 ng of total RNA samples were processed with the SMART-Seq v4 Ultra Low Input RNA Kit (Clontech), following manufacturer instructions. Resulting cDNA was sheared on a S220 Focused-ultrasonicator (Covaris) and subsequently processed with the "NEBNext Ultra II DNA Library Prep Kit for Illumina" (NEB #E7645). Briefly, oligo(dT)-primed reverse transcription was performed in presence of a template switching oligonucleotide, double stranded cDNA was produced by 11 cycles of PCR and submitted to acoustic shearing. Fragments were processed through subsequent enzymatic treatments of end-repair, dA-tailing, and ligation to Illumina adapters. Adapter-ligated libraries were completed by limited-cycle PCR (8 cycles). The resulting directional cDNA libraries were sequenced for 50 bases in a single-read format, instrument: Illumina HiSeq2500 Primary data processing: Image analysis, per-cycle base-calling and quality score assignment was performed with Illumina Real Time Analysis software. Conversion of Illumina BCL files to bam format was performed with the Illumina2bam tool (Wellcome Trust Sanger Institute - NPG). The complete set of reads has been deposited in the GEO repository (accession numbers: GSE127186 and GSE152378). Paired-end reads were aligned to the hg19 human genome using STAR [25] with default parameters. Gene counts were computed using the Rsubread package [26] with the inbuilt annotation for the hg19 genome. Differential expression was performed using DESeq2 [27] using the cell line as

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2021 doi:10.20944/preprints202103.0517.v1

covariate. Genes were sorted by the shrunk fold change as computed by the "lfcShrink" function.

2.8. Differential gene expression comparison of published studies

Gene set enrichment analysis was performed using the pre-ranked GSEA function as implemented by the Broad Institute [28]. First, we generated two gene signatures of differentially expressed genes (DEGs), either up-regulated or down-regulated, in short-term cultured human PSCs (D2#2, D2#4, HERVH and H1) (raw data reported in [20] and accessible at GSE127186). We defined gene signatures as DEGs with adjusted FDR q-value lower than 0.1 and log2 fold changes larger than 1 (for up regulated genes) or lower than -0.5 (for downregulated genes). These short-term up and down signatures for CDK8/19i-naïve cells are in Table S2. These signatures were used to perform GSEA against a ranked list for DEGs in the long-term cultured PSCs in the absence or presence of CDK8/19i (n=5 cell lines, see above RNA-seq and Table S3). As a control for other naïve cells obtained with 2i-based cocktails, we used ranked lists of DEGs from [6], [7], [9], [10], [15], [29], [30], also detailed in [20]. For the data generated in this study, genes were ranked by the shrunk fold change as computed by the "lfcShrink" function. Gene set enrichment was run with standard GSEA settings. Results were summarized through the Normalized Enrichment score (NES). Data with $P<0.05$ and FDR $q<0.05$ were considered significant and marked with an asterisk (*) in the heatmap of GSEA NES scores: * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Heatmaps were plotted using the ggplot2 package [31] in R [32].

2.9. DNA methylation

DNA methylation analyses were performed by the laboratory of Mario Fraga (CINN, Oviedo). DNA was purified and CpG methylation status at individual CpG sites of DMR (Differentially Methylated Regions) regions, at imprinted genes, was assessed by DNA bisulphite-conversion and pyrosequencing, as described [33]. DNA was purified following the standard phenol:chloroform:isoamyl alcohol (25:24:1) (Sigma) extraction protocol. Bisulfite modification of DNA was performed with the EZ DNA Methylation-Gold kit (#D5006, Zymo Research) following the manufacturer's instructions. The set of primers for PCR amplification and sequencing were designed using the specific software Pyrosequencing Assay Design (version 2.0.01.15) from Biotage AB. Primer sequences were designed to hybridize with CpG free sites to ensure methylation-independent amplification, one containing a Biotin mark (Btn) at the 5' end (Table S1). After PCR amplification and purification (based on the biotin mark) of the region of interest with the specific primers, pyrosequencing was performed using PyroMark Q24 reagents, Vacuum Prep Workstation (Biotage AB) and specific-designed sequencing primer (Table S1). Sequencing primers were designed covering 4 or 5 CpGs per DMR. Results were analysed with PyroMark Q24 2.0.6 (Qiagen) software and equipment, obtaining percentage of methylation for each specific CpG site studied.

2.10. Exome sequencing

Genomic DNA (gDNA) was purified following manufacturer's instructions (Qiagen #69504). gDNA samples were purified using AMPure DNA Magnetic Beads and a RNASe treatment was performed on the samples. The purified gDNA was quantified using DNA Hs Qubit Assay and fragmented in 250-300bp size using COVARIS System. Exome capture Hybridization was performed according to manufacturer's instructions (AGILENT SureSelect XT Kit), followed by Exome capture Library preparation (NEBNext® Ultra™ DNA Library Prep Kit for Illumina). Next Generation Sequencing was performed at HiSeq 2500 Sequencing System Illumina, Sequencing Type 125 nt Pair End. The Picard suite (Picard tools http://broadinstitute.github.io/picard/) was used to trim Illumina adapter. Paired end reads were aligned to the hg19 human genome using bwa [34] followed by removal of duplicate reads using sambamba [35]. Following the best practices recommended by the authors of GATK [36] read quality was recalibrated (GATK version

 doi:10.20944/preprints202103.0517.v1

4.0.6.0). GATK's Haplotype caller [37] was used to discover single nucleotide and short polymorphisms. SNPs were filtered using the following thresholds: FS > 60.0, QD < 2.0, MQRankSum < -12.5, ReadPosRankSum < -8.0 and MQ < 40.0.

2.11. Imprinting score

A list of putative imprinted genes was compiled from [12, 14, 38]. For each gene all SNPs passing filters were selected. For RNAseq samples corresponding to each Exome seq sample, reads aligning to selected SNPs were counted and the percentage of the alternative allele was computed. We filtered-out positions with less than 10 reads. Following the thresholds in [38] we defined three categories of expression depending on the alternative allele fraction: Monoallelic 0 < AAF <= 0.14; Partially-monoallelic 0.15 < AFF <= 0.29; Biallelic 0.3 < AFF.

2.12. Statistical Analysis

For differential gene expression by RNAseq, a threshold of FDR q-value of q<0.05, or q<0.01 was applied, as indicated in each case. In GSEA analysis, the standard threshold for significance was applied, where P <0.05 and FDR q-value < 0.05.

3. Results

3.1. CDK8/19i stabilizes naïve human PSCs over long-term passaging

We have previously reported that human PSCs cultured in the presence of a CDK8/19 inhibitor for about 5 to 10 passages phenocopy molecular features of naïve pluripotency (**Figure 1A**) at the transcriptional, proteome and phospho-proteome levels [20]. Remarkably, all media cocktails based on 2i for the induction of naïve pluripotency result in karyotypic instability, gradual erasure of DNA methylation on imprinted genes and deficits in developmental potency [12–16]. Published literature considers stable long-term culture those that exceed 10 passages, which is the time required for the observation of karyotypic abnormalities [12, 19]. Shorter-term passages may not be enough to manifest this problem. We wondered if the CDK8/19i-naïve state can be preserved for long periods of culture. For this, we used a reporter of the naïve state that we had not tested previously, in particular, a reporter based on the *OCT4* (also known as *POU5F1*) enhancer (**Figure 1B**). It is well described that *OCT4* expression is primarily dependent on the proximal enhancer (PE) in primed human PSCs, while naïve human PSCs are characterized by the predominant utilization of the distal enhancer (DE) [7, 10]. After 15 passages (~90 days in culture) in the presence of CDK8/19i, human WIBR3 PSCs carrying the *OCT4*-ΔPE-GFP reporter presented a substantial increase in GFP fluorescence compared to the same cells in the absence of CDK8/19i (**Figure 1C**). Flow cytometry analyses, confirmed a substantial activation of *OCT4*-ΔPE-GFP in CDK8/19i-cultured cells compared to their counterparts in the absence of inhibitor (primed conditions), and compared to additional negative controls consisting of retinoic acid-induced (RA) differentiated cells and primed PSCs not carrying the reporter (**Figure 1D**).

To further support the stability of the naïve state in the presence of CDK8/19i, we performed whole-transcriptome profiling on long-term (>10 passages) adapted CDK8/19i-naïve and primed human PSCs from 5 different cell lines, namely, 3 iPSCs (CB5, D2#2, D2#4) and 2 ESCs (WIBR3, and H1) (see Methods for accession number in public databases). We asked if signatures for upregulated and downregulated genes in short-term cultured CDK8/19i-naïve vs primed cells (reported in [20]; see Table S2) were enriched in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2021 doi:10.20944/preprints202103.0517.v1

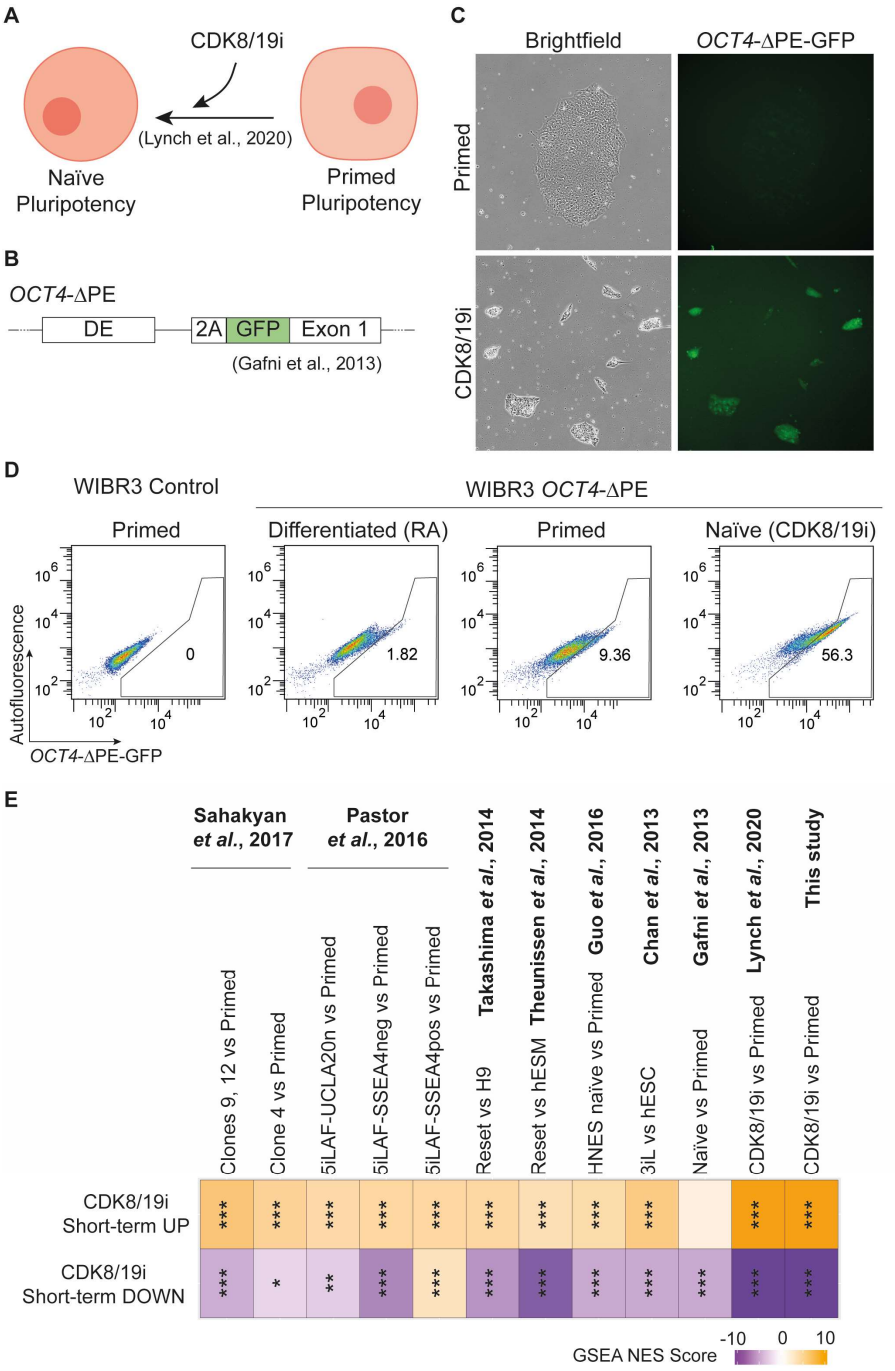


Figure 1. Stability of hPSC-naïve pluripotency after long-term culture with CDK8/19i. (A) Experimental scheme for inducing naïve conditions using the previously reported CDK8/19i chemical approach [20]**. (B)** Schematic diagram of *OCT4*-ΔPE-GFP reporter [7] for naïve human pluripotency based on the *OCT4* distal enhancer (DE). **(C)** Brightfield images showing colony morphology (left panels) or *OCT4*-ΔPE-GFP expression (right panels), both in primed (upper images) or CDK8/19i treated (bottom images) WIBR3 reporter hESCs. **(D)** Comparative fluorescent cytometry analyses of WIBR3 reporter hESCs after differentiation with retinoic acid (RA), primed conditions, or CDK8/19i-naïve conditions, as well as the parental line WIBR3 (primed) as negative control (left panel). Numbers indicate the % of cells in the GFP-positive gate. **(E)** Heatmap of the normalized enrichment scores (NES) of the gene-set enrichment analyses (GSEA). As signatures, we used the differentially up- or down- regulated genes in short-term CDK8/19i-naïve (relative to primed). These signatures were tested on the ranked lists of gene expression changes in previously published naïve conditions (see references) and in our long-term adapted CDK8/19i hPSCs (this study). For the last comparison, we used a total of 5 hPSCs: D2#2, D2#4, H1, CB5 and WIBR3.

 doi:10.20944/preprints202103.0517.v1

the differentially expressed genes of our newly generated transcriptomes of long-term cultured cells (n=5) in the presence or absence of CDK8/19i (see Table S3). Interestingly, the short-term CDK8/19i naïve signatures were highly enriched in the long-term CDK8/19i-naïve cells (**Figure 1E**). The short-term CDK8/19i-naïve signatures were also compared with previously published naïve data sets [6, 7, 9, 10, 15, 29, 30] and in all cases showed high enrichment scores similar to the ones in our long-term CDK8/19i-naïve cells (**Figure 1E**). Of note, both short- and long-term CDK8/19i altered genes show a significant overlap to 2i-based naïve cocktails, as we previously reported [20]. Therefore, we conclude that naïve state achieved by inhibition of the Mediator kinases CDK8/19 is stable over long-term passages.

3.2. Long-term preservation of genomic imprints in CDK8/19i-naïve human PSCs

Uncontrolled global DNA demethylation and loss of imprints have been reported for all tested 2i-based naïve human PSCs [12–16]. Recent studies have shown that the inhibition of MEK/ERK signaling is directly responsible for these undesirable effects [19, 39]. Considering that CDK8/19 inhibition does not affect MEK signaling [20], we hypothesized that long-term culture of human PSCs with CDK8/19i might preserve genomic imprinting.

First, we followed a strategy previous used based on the combination of exome and RNA sequencing [14, 38]. In particular, we performed exome sequencing of 5 human PSCs long-term cultured (>15 passages) in the presence of CDK8/19i, and this was used to identify SNPs in a total of 296 imprinted genes (obtained from [12, 14]). Then, RNA-seq data were used to discriminate the levels of expression of each allele. Following previous criteria [38], sites displaying 0-14% of minor allele contribution to total gene expression were defined as monoallelic, 15-29% as partially monoallelic and 30-50% as biallelic. Only a small fraction of imprinted genes (from 9 to 13) were informative (presence of SNPs and >10 reads, see Methods) and expressed in a monoallelic or partially monoallelic manner in human primed PSCs. Interestingly, most (77% to 100%) of the primed imprinted genes remained monoallelically expressed in long-term cultured CDK8/19i-naïve human PSCs (**Figure 2A**). This includes the *H19/IGF2* locus which is known to be particularly susceptible to loss of imprinted expression in PSCs [14, 15, 38].

We next assessed by pyrosequencing the methylation status of specific CpG positions within the differentially-methylated regions (DMRs) of some maternal (*SNRPN*) and paternal (*MEST*, *PEG10*) imprinted loci. As expected, all the CpG positions analyzed (4 or 5 positions per DMR) were hemi-methylated in primed PSCs (**Figure 2B**). Interestingly, hemi-methylation was preserved in all the long-term cultured CDK8/19i-naïve human PSCs (**Figure 2B**). As a further challenge to the stability of the imprints, we injected the long-term cultured human PSCs, both primed and naïve, into SCID mice until the formation of teratomas. Genomic DNA from the teratomas was analyzed as before to determine the methylation status of DMR. The teratomas obtained with CDK8/19i-naïve H1 cells preserved hemi-methylation at all the sites analyzed (**Figure 2C**). In the case of CDK8/19i-naïve D2#2 cells, two DMRs lost their imprints (*MEST* and *PEG10*) and one DMR preserved hemi-methylation (SNRP) (**Figure 2C**).

All together, we conclude that prolonged culture of PSCs in the presence of CDK8/19i does not erase DNA methylation imprints. However, we detected some loss of imprinting after teratoma formation. The high stability of imprinting in CDK8/19i-naïve human PSCs is in contrast with the recent findings that >70% of imprinted DMRs were erased in 2i-based naïve human PSCs [16].

3.3. Differentiation potential of CDK8/19i-naïve human PSCs into primordial germ cells

As a first test of the differentiation potential of CDK8/19i-naïve hPSCs, we examined their capacity to generate primordial germ cell (PGC)-like cells (PGCLCs). PGCLCs have been recently generated from human PSCs, through the formation of embryoid bodies (EBs) in the presence of specific media supplements [22, 40]. These studies have shown

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2021 doi:10.20944/preprints202103.0517.v1

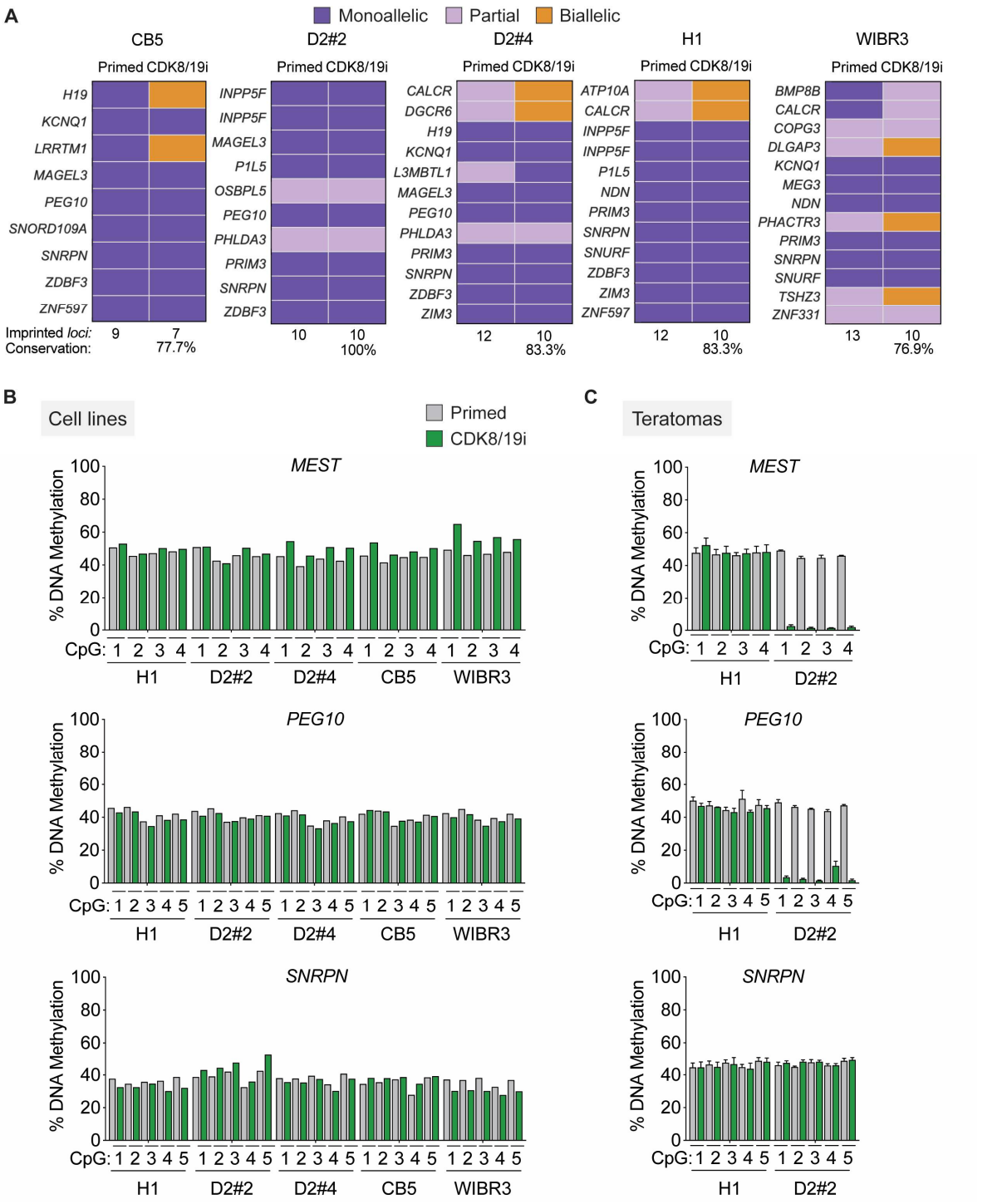


Figure 2: Long-term cultured CDK8/19i-naïve hPSCs retain monoallelic expression and hemi-methylation at imprinted *loci*. (A) Heatmaps of allele specific expression data for the indicated informative imprinted genes in primed *versus* CDK8/19i long-term cultured hPSCs. The lower part of each panel indicates the number and percentage of genes expressed in a monoalleic or partially-monoallelic manner after long-term culture in the presence of CDK8/19i. **(B)** CpG methylation status of differentially-methylated regions (DMRs) at the indicated imprinted *loci* in five hPSC lines adapted to CDK8/19i for >10 passages. For each DMR, a total of 4 or 5 individual CpGs were analyzed by pyrosequencing. **(C)** CpG methylation status of DMRs in teratomas obtained from primed hPSCs or long-term adapted to CDK8/19i.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2021 doi:10.20944/preprints202103.0517.v1

robust production of hPGCLCs from primed PSCs or PSCs exposed to a naïve pre-treatment for 4 days with a 2i-based cocktail known as NHSM [7].

Here, we have compared the effect of a naïve pre-treatment with CDK8/19i or with another 2i-based cocktail known as PXGL [8] on primed H1 ESCs (**Figure 3A**). These naïve pre-treatments for 4 days were sufficient to increase the mRNA levels of *NANOG*, thereby supporting the induction of naïve features by both methods (**Figure 3B**). After the naïve pre-treatment, PGCLCs were induced according to the protocol of Kojima et al. [22] (**Figure 3A**). Notably, primed and CDK8/19i-exposed cells efficiently formed EBs (**Figure 3C**) and upregulated specific PGCLCs markers, such as *CD38*, *PRDM1*, and *NANOS3* (**Figure 3D**). Of note, naïve pre-treatment with PXGL was detrimental for the formation of EBs and we could not assess PGCLC differentiation. These results indicate that short treatment of primed PSCs with CDK8/19i is sufficient to induce naïve features without affecting their capacity to form embryoid bodies and differentiate into primordial germ cells.

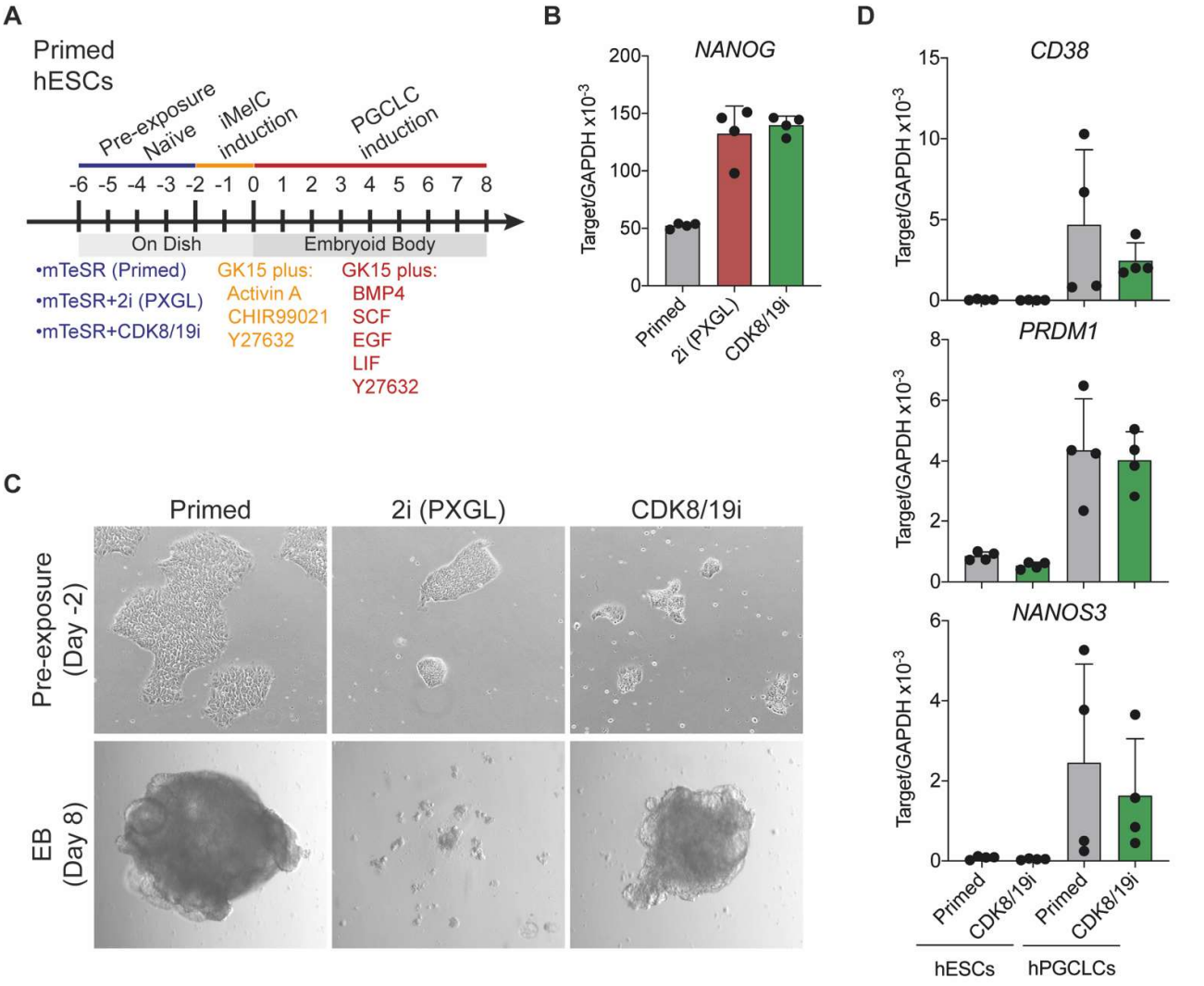
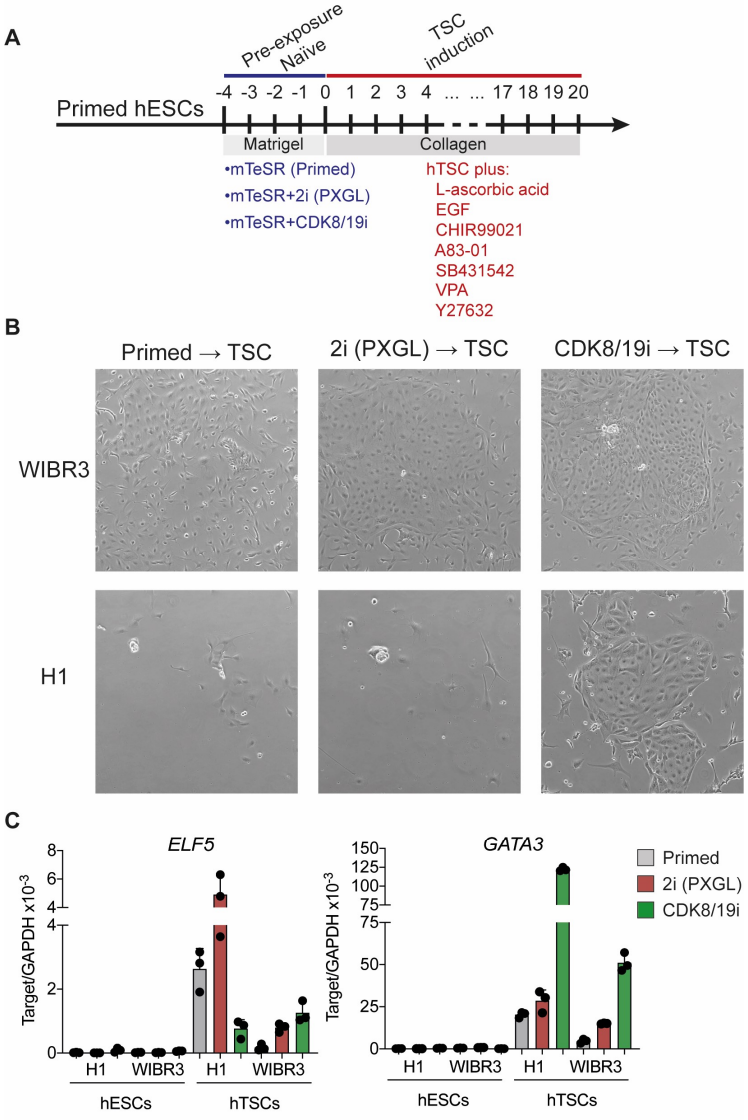


Figure 3: CDK8/19i-naïve hPSCs can differentiate into primordial germ cell-like cells (hPGCLCs). (A) Experimental scheme for assessing PGC differentiation of primed, 2i (PXGL) and CDK8/19i–naïve hESCs. **(B)** mRNA expression level of *NANOG* assessed by qRT-PCR. Upregulation of *NANOG* is a feature of the naïve state. Data represent mean ± Std Dev from 4 technical replicates. **(C)** Brightfield images showing H1 hESCs before and after exposure to naïve conditions and embryoid bodies (EBs) in hPGCLCs induction conditions at day 8. **(D)** mRNA expression level of selected hPGCLC markers in hESCs and EBs differentiated into hPGCLCs assessed by qRT-PCR. Data of hESC are mean ± Std Dev from 4 technical replicates and data of hPGCLCs are mean ± Std Dev from 4 biological replicates.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 March 2021 doi:10.20944/preprints202103.0517.v1

3.4. Differentiation potential of CDK8/19i-naïve human PSCs into trophoblast stem cells

Previous studies have reported that naïve human PSCs have some capacity to acquire extra-embryonic fates, and this is a distinctive property of the naïve state absent in primed PSCs [16, 23]. Here, following the same strategy used above to induce PGCLC differentiation, primed H1 and WIBR3 ESCs were pre-treated for 4 days with CDK8/19i or with 2i-based PXGL to induce naïve features (see above **Figure 3B**). Cells were then cultured for 20 days in human TSC medium, as previously reported [24] (**Figure 4A**). In the case of WIBR3 cells, both PXGL-naïve and CDK8/19i-naïve cells formed colonies with a typical human TSC morphology (**Figure 4B**). In contrast, primed WIBR3 did not form colonies with TSC morphology (**Figure 4B**). In the case of H1 cells, TSC colonies were only observed when using CDK8/19i-naïve cells (**Figure 4B**). The expression of trophoblast marker *GATA3* correlated with the formation of TSC colonies, thus being maximally expressed in cultures derived from CDK8/19-naïve H1 and WIBR3 cells (**Figure 4C**). In the case of trophoblast marker *ELF5*, expression was higher under TSC culture conditions but their levels did not correlate with TSC colony formation (**Figure 4C**). Together, these findings indicate that the CDK8/19i-naïve PSCs can efficiently differentiate into trophoblast stem cells.



 doi:10.20944/preprints202103.0517.v1

Figure 4: CDK8/19i-naïve hPSCs can differentiate into trophoblast stem cells (hTSCs). (A) Experimental scheme for assessing hTSCs differentiation of primed, 2i (PXGL) and CDK8/19i-naïve hESCs. **(B)** Brightfield images showing hTSC-like cells derived from hESCs (WIBR3 and H1) after exposure to naïve conditions. **(C)** mRNA expression level of selected trophoblast markers assessed by qRT-PCR. Data are mean ± Std Dev of 3 technical replicates.

4. Discussion

This study reports three main findings regarding the ability of CDK8/19i to induce naïve features in human PSCs. First, long-term culture (>15 passages) in the presence of CDK8/19i does not affect the expression of naïve features, as indicated by the OCT4-ΔPE-GFP reporter. Second, long-term CDK8/19i-naïve hPSCs preserve intact their genomic imprints, including hemi-methylation of differentially methylated regions and monoallelic expression. And, third, the naïve state induced by CDK8/19i in hPSCs allows subsequent differentiation into primordial germ cells or into trophoblast stem cells, the latter is considered a property characteristic of naïve hPSCs.

Naïve hPSCs have been previously achieved using a variety of chemical cocktails [12], all of them based on the mouse 2-inhibitor cocktail (2i) [4]. However, 2i-based-naïve hPSCs are, in general, unstable compared to primed hPSCs. This instability includes karyotypic abnormalities and loss of imprinted marks [10, 12, 15, 30]. The inhibition of MEK, a key feature of all the 2i-based cocktails, is considered the key inducer of this genomic instability. In support of this, reduction of the amount of MEK inhibitor in naïve cocktails reduces the accumulation of karyotypic abnormalities [19]. The reason why MEK inhibition results in genomic instability and loss of imprinting is not completely understood, although multiple lines of evidence implicate the profound DNA demethylation caused by the 2i-based cocktails. During pre-implantation embryonic stages, genomic hypomethylation is transient and genomic imprinting remains protected [41], however in contrast, DNA hypomethylation in 2i-naïve hPSCs is not transient and does not replicate the embryonic demethylation patterns [15]. The role of MEK in DNA methylation could be mediated by its stabilizing effect on UHFR1, a key factor for the recruitment of the maintenance DNA methyltransferase DNMT1 necessary for genomic stability and maintenance of imprinting [18, 42–44].

In contrast to the above, CDK8/19 inhibition does not affect MEK activity and does not result in global DNA hypomethylation [20]. Given the loss of imprinting caused by 2i-based cocktails, here we considered it of high importance to determine the stability of imprinting in long-term cultured CDK8/19i-naïve hPSCs. Interestingly, we found that CDK8/19i-naïve hPSCs (five different cell lines, including three iPSCs and two ESCs) retain monoallelic expression and hemi-methylation of maternal and paternal imprinted genes. In the case of mouse 2i-naïve PSCs, it has been reported that female cells are more susceptible to undergo loss-of-imprinting [39, 45]. In this regard, it is worth mentioning that our five tested hPSC lines include two female cell lines, WIBR3 and CB5. Moreover, teratomas generated from long-term CDK8/19i-naïve hPSCs also preserved most of their tested genetic imprints. We conclude that CDK8/19i-naïve hPSCs have a high degree of imprinting stability.

Naïve and primed hPSCs exhibit different levels of differentiation potential. Several studies have tried to model in culture the derivation of primordial germ cell-like cells (PGCLCs) from human PSCs [22, 40, 46, 47]. One of these studies compared the ability to differentiate into PGCLCs of 2i-naïve versus primed hPSCs, observing that 2i-naïve hPSCs are more efficient in this differentiation [40]. Here, we have used an alternative differentiation protocol [22] and we have observed that CDK8/19i-naïve cells are as efficient as primed cells. We conclude that the induction to CDK8/19i-naïve features in human PSCs does not impair their capacity to generate PGCLCs.

It has been reported that 2i-naïve human PSCs present some intrinsic plasticity to express genes characteristic of the trophectoderm [48, 49]. The study of human trophoblast stem cells TSCs has been greatly accelerated thanks to the identification of culture conditions for these cells [24]. Based on these sets of observations, it was found that 2i-naïve hPSCs, but not primed hPSCs, can be indeed converted into hTSCs [23, 50]. Given that

this is a differentiation assay uniquely accessible to naïve hPSCs, but not to primed hPSCs, we wondered if CDK8/19i-naïve hESCs can also differentiate into hTSCs. In fact, we observed with two hPSC lines that CDK8/19i pre-treatment was critical for their ability to become hTSCs.

Taken together, we conclude that CDK8/19i-naïve cells are stable after prolonged culture, preserving imprinting and potency, and capable of efficiently differentiating into hPGCLCs and hTSCs.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, **Table S1**: primers used in this study, **Table S2**: RNA-seq analysis in short-term human PSCs adapted to culture in Primed or CDK8/19i conditions, **Table S3**: RNA-seq analysis in long-term human PSCs adapted to culture in Primed or CDK8/19i conditions.

Author Contributions: R.B. designed and performed most of the experiments, contributed to bioinformatics data analysis and cowrote the manuscript. C.J.L. provided experimental support and contributed to bioinformatics data analysis. C.S.-O.A., performed the bioinformatics analyses. R.G.U. and M.F.F. performed the DNA methylation analysis. M.S. designed and supervised the study, secured funding, analysed the data and cowrote the manuscript. All of the authors discussed the results and commented on the manuscript.

Funding: Work in the laboratory of M.F.F. was funded by the Spanish Association Against Cancer (PROYE18061FERN to M.F.F.), the Asturias Government (PCTI) co-funding 2018-2022/FEDER (IDI/2018/146 to M.F.F.), Fundación General CSIC (0348_CIE_6_E to M.F.F.) and the Health Institute Carlos III (Plan Nacional de I+D+I) co-funding FEDER (PI18/01527 to M.F.F). R.G.U. is supported by Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER). We also acknowledge support from the IUOPA-ISPA-FINBA (the IUOPA is supported by the Obra Social Cajastur-Liberbank, Spain). Work in the laboratory of M.S. was funded by the IRB and by grants from Spanish Ministry of Economy co-funded by the European Regional Development Fund (SAF2017-82613-R), ERC (ERC-2014-AdG/669622), Banco Santander (Santander Universities Global Division), laCaixa Foundation, and Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement of Catalonia (Grup de Recerca consolidat 2017 SGR 282). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Institutional Review Board Statement: Studies with hPSCs were ethically approved at the Institute for Research in Biomedicine (IRB Barcelona) in Barcelona by the Ethics Committee of the CMRB, by the National Advisory Committee for Human Tissue and Cell Donation and Use of the ISCIII and the Ministry of Health from the Government of Catalonia (project numbers: 0336S/2473/2017; 0336/747/2018).

Data Availability Statement: RNA-seq data are available from the GEO database under accession numbers GSE127186 and GSE152378. Exome-seq data are available from the NCBI (SRA) repository under accession number PRJNA629655.

Acknowledgments: We thank Jacob Hanna and Nuria Montserrat for generously providing reagents, as well as the IRB core facilities.

Conflicts of Interest: The authors declare no competing interests.

References

1. Boroviak, T., Loos, R., Bertone, P., Smith, A. & Nichols, J. The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. Nat. Cell Biol. 16, 513–525 (2014).
2. Marks, H. et al. The transcriptional and epigenomic foundations of ground state pluripotency. Cell 149, 590–604 (2012).
3. Nichols, J. & Smith, A. Naive and primed pluripotent states. Cell Stem Cell 4, 487–92 (2009).
4. Ying, Q. L. et al. The ground state of embryonic stem cell self-renewal. Nature 453, 519–523 (2008).
5. Hanna, J. et al. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc. Natl. Acad. Sci. U. S. A. 107, 9222–9227 (2010).
6. Chan, Y.-S. et al. Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. Cell Stem Cell 13, 663–675 (2013).
7. Gafni, O. et al. Derivation of novel human ground state naive pluripotent stem cells. Nature 504, 282–6 (2013).
8. Guo, G. et al. Epigenetic resetting of human pluripotency. Development 144, 2748–2763 (2017).
9. Takashima, Y. et al. Resetting Transcription Factor Control Circuitry toward Ground-State Pluripotency in Human. Cell 158, 1254–1269 (2014).

10. Theunissen, T. W. et al. Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell 15, 471–487 (2014).
11. Ware, C. B. et al. Derivation of naïve human embryonic stem cells. Proc. Natl. Acad. Sci. 111, 4484–4489 (2014).
12. Liu, X. et al. Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. Nat. Methods 14, 1055–1062 (2017).
13. Avior, Y., Eggan, K. & Benvenisty, N. Cancer-related mutations identified in primed and naive human pluripotent stem cells. Cell Stem Cell 25, 456–461 (2019).
14. Bar, S., Schachter, M., Eldar-Geva, T. & Benvenisty, N. Large-scale analysis of loss of imprinting in human pluripotent stem cells. Cell Rep. 19, 957–968 (2017).
15. Pastor, W. A. et al. Naive human pluripotent cells feature a methylation landscape devoid of blastocyst or germline memory. Cell Stem Cell 18, 323–329 (2016).
16. Theunissen, T. W. et al. Molecular criteria for defining the naive human pluripotent state. Cell Stem Cell 19, 1–14 (2016).
17. Von Meyenn, F. et al. Impairment of DNA methylation maintenance is the main cause of global demethylation in naive embryonic stem cells. Mol. Cell 62, 1–14 (2016).
18. Li, J. et al. Activated MEK/ERK pathway drives widespread and coordinated overexpression of UHRF1 and DNMT1 in cancer cells. Sci. Rep. 9, (2019).
19. Di Stefano, B. et al. Reduced MEK inhibition preserves genomic stability in naive human embryonic stem cells. Nat. Methods 15, 732–740 (2018).
20. Lynch, C. J. et al. Global hyperactivation of enhancers stabilizes human and mouse naive pluripotency through inhibition of CDK8/19 Mediator kinases. Nat. Cell Biol. 22, 1223–1238 (2020).
21. Lynch, C. J., Bernad, R., Calvo, I. & Serrano, M. Manipulating the Mediator complex to induce naïve pluripotency. Exp. Cell Res. 395, (2020).
22. Kojima, Y. et al. Evolutionarily distinctive transcriptional and signaling programs drive human germ cell lineage specification from pluripotent stem cells. Cell Stem Cell 21, 517–532 (2017).
23. Dong, C. et al. Derivation of trophoblast stem cells from naïve human pluripotent stem cells. Elife 9, (2020).
24. Okae, H. et al. Derivation of human trophoblast stem cells. Cell Stem Cell 22, 50–63 (2018).
25. Dobin, A. et al. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013).
26. Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res. 41, (2013).
27. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 1–21 (2014).
28. Subramaniana, A. et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A. 102, 15545–15550 (2005).
29. Sahakyan, A. et al. Human naive pluripotent stem cells model X chromosome dampening and X inactivation. Cell Stem Cell 20, 1–15 (2017).
30. Guo, G. et al. Naive pluripotent stem cells derived directly from isolated cells of the human inner cell mass. Stem Cell Reports 6, 1–10 (2016).
31. Wickham, H. ggplot2: Elegant Graphics for Data Analysis Hadley. vol. 35 (2009).
32. The R Core Team. R: a language and environment for statistical computing. http://www.R-project.org/ vol. 2 (2008).
33. Wang, R. Y. H., Gehrke, C. W. & Ehrlich, M. Comparison of bisulfite modification of 5-methyldeoxycytidine and deoxycytidine residues. Nucleic Acids Res. 8, 4777–4790 (1980).
34. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760 (2009).
35. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. Bioinformatics 31, 2032–2034 (2015).
36. McKenna, A. et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303 (2010).
37. Poplin, R. et al. Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv (2017) doi:10.1101/201178.
38. Rugg-Gunn, P. J., Ferguson-Smith, A. C. & Pedersen, R. A. Status of genomic imprinting in human embryonic stem cells as revealed by a large cohort of independently derived and maintained lines. Hum. Mol. Genet. 16, 243–251 (2007).
39. Choi, J. et al. Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells. Nature 548, 219–223 (2017).
40. Mitsunaga, S. et al. Relevance of iPSC-derived human PGC-like cells at the surface of embryoid bodies to prechemotaxis migrating PGCs. Proc. Natl. Acad. Sci. U. S. A. 114, E9913–E9922 (2017).
41. Guo, H. et al. The DNA methylation landscape of human early embryos. Nature 511, 606–610 (2014).
42. Branco, M. R., Oda, M. & Reik, W. Safeguarding parental identity: Dnmt1 maintains imprints during epigenetic reprogramming in early embryogenesis. Genes Dev. 22, 1567–1571 (2008).
43. Hirasawa, R. et al. Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. Genes Dev. 22, 1607–1616 (2008).
44. von Meyenn, F. et al. Impairment of DNA methylation maintenance is the main cause of global demethylation in naive embryonic stem cells. Mol. Cell 62, 848–861 (2016).

45. Yagi, M. et al. Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation. Nature 548, 224–227 (2017).
46. Irie, N. et al. SOX17 is a critical specifier of human primordial germ cell fate. Cell 160, 1–16 (2015).
47. Sasaki, K. et al. Robust in vitro induction of human germ cell fate from pluripotent stem cells. Cell Stem Cell 17, 178–194 (2015).
48. Pastor, W. A. et al. TFAP2C regulates transcription in human naive pluripotency by opening enhancers. Nat. Cell Biol. 20, 553–564 (2018).
49. Qin, H. et al. YAP induces human naive pluripotency. Cell Rep. 14, 2301–2312 (2016).
50. Cinkornpumin, J. K. et al. Naive human embryonic stem cells can give rise to cells with a trophoblast-like transcriptome and methylome. Stem Cell Reports 15, 198–213 (2020).