nature cell biology

TECHNICAL REPORT

https://doi.org/10.1038/s41556-018-0254-5

Direct generation of human naive induced pluripotent stem cells from somatic cells in microfluidics

Stefano Giulitti[1,2,3,10], Marco Pellegrini[3,10], Irene Zorzan[3], Paolo Martini[4], Onelia Gagliano[1,2], Margherita Mutarelli[5], Michael Johannes Ziller[6], Davide Cacchiarelli[5,7], Chiara Romualdi[4], Nicola Elvassore[1,2,8,9*] and Graziano Martello[3*]

Induced pluripotent stem cells (iPSCs) are generated via the expression of the transcription factors *OCT4* (also known as *POU5F1*), *SOX2*, *KLF4* and *cMYC* (OSKM) in somatic cells. In contrast to murine naive iPSCs, conventional human iPSCs are in a more developmentally advanced state called primed pluripotency. Here, we report that human naive iPSCs (niPSCs) can be generated directly from fewer than 1,000 primary human somatic cells, without requiring stable genetic manipulation, via the delivery of modified messenger RNAs using microfluidics. Expression of the OSKM factors in combination with *NANOG* for 12 days generates niPSCs that are free of transgenes, karyotypically normal and display transcriptional, epigenetic and metabolic features indicative of the naive state. Importantly, niPSCs efficiently differentiate into all three germ layers. While niPSCs can be generated at low frequency under conventional conditions, our microfluidics approach enables the robust and cost-effective production of patient-specific niPSCs for regenerative medicine applications, including disease modelling and drug screening.

Pluripotent stem cells (PSCs) were derived from early embryos as embryonic stem cells (ESCs)[1,2] and subsequently from somatic cells via transcription factor-mediated reprogramming[3]. Murine PSCs are thought to represent a naive pluripotent state, characterized by the expression of the transcription factors *Oct4*, *Sox2*, *Nanog*, *Klf4* and *Tfcp2l1* in response to the cytokine LIF and the inhibition of GSK3 and MEK kinases[4,5]. Murine naive PSCs display low levels of repressive epigenetic modification, such as trimethylation of lysine 9 on histone 3 (H3K9me3) or cytosine methylation, and are metabolically sustained by high levels of oxidative phosphorylation[4,5]. Importantly, murine naive PSCs efficiently generate cells of all three germ layers.

Conversely, conventional human PSCs, either derived from early embryos[2] or via the reprogramming of somatic cells using OSKM[3], resemble a distinct, more advanced developmental stage called primed pluripotency. Primed PSCs express OCT4, SOX2 and NANOG in response to fibroblast growth factor (FGF) and transforming growth factor-β signals, display higher levels of repressive epigenetic modifications and are mostly glycolytic[4]. Different primed PSC lines display a differentiation bias towards some germ layers[6].

Recently, human naive PSCs have been generated through either the expression of transgenes together with genomic reporter constructs or directly from human embryos[5,7]. The use of human embryos has ethical limitations and obviously does not allow the generation of patient-specific naive PSCs. Meanwhile, the conversion of somatic cells to naive pluripotency using available protocols requires one or more rounds of stable genetic manipulations, which are time-consuming, inefficient and potentially mutagenic. To overcome such limitations, we derived a strategy for the efficient generation of transgene-free niPSCs directly from somatic cells.

Results

Devising a strategy for the transgene-free generation of niPSC colonies. We have previously shown that the delivery of modified messenger RNAs (mmRNAs) encoding for OSKM, *LIN28* and *NANOG* using microfluidics results in the rapid and efficient generation of primed iPSCs[8]. Here, we compared the global gene expression profile of primary iPSC colonies generated via microfluidics or under conventional culture conditions (CCC) and observed that several transcription factors associated with naive pluripotency were upregulated (Fig. 1a,b). This result suggests that microfluidics could favour activation of the naive network. Therefore, we determined whether naive PSCs could be derived directly from primary somatic cells using a microfluidics approach. To test this hypothesis, we transfected BJ human fibroblasts with mmRNAs via microfluidics using our previously described setup[8] (Supplementary Fig. 1a), leaving out the primed pluripotency factor LIN28 (that is, using OSKMN)[9]. *DPPA5*, a gene highly induced in human naive cells[10,11], became detectable from day 14 (Supplementary Fig. 1b), despite the use of FGF2-containing medium for generating primed PSCs (that is, primed reprogramming medium (PRM)). The generated colonies showed the morphology of primed PSCs and lost *DPPA5* expression after passaging. We then tried to sustain the expression

[1]Department of Industrial Engineering, University of Padova, Padua, Italy. [2]Venetian Institute of Molecular Medicine, Padua, Italy. [3]Department of Molecular Medicine, Medical School, University of Padova, Padua, Italy. [4]Department of Biology, University of Padova, Padua, Italy. [5]Telethon Institute of Genetics and Medicine (TIGEM), Pozzuoli, Italy. [6]Department of Translational Psychiatry, Max Planck Institute of Psychiatry, Munich, Germany. [7]Department of Translational Medicine, University of Naples "Federico II", Naples, Italy. [8]Stem Cell and Regenerative Medicine Department, DBC Program, UCL Great Ormond Street Institute of Child Health, University College London, London, UK. [9]Shanghai Institute for Advanced Immunochemical Studies (SIAIS), ShanghaiTech University, Shanghai, China. [10]These authors contributed equally: Stefano Giulitti, Marco Pellegrini. *e-mail: n.elvassore@ucl.ac.uk; graziano.martello@unipd.it

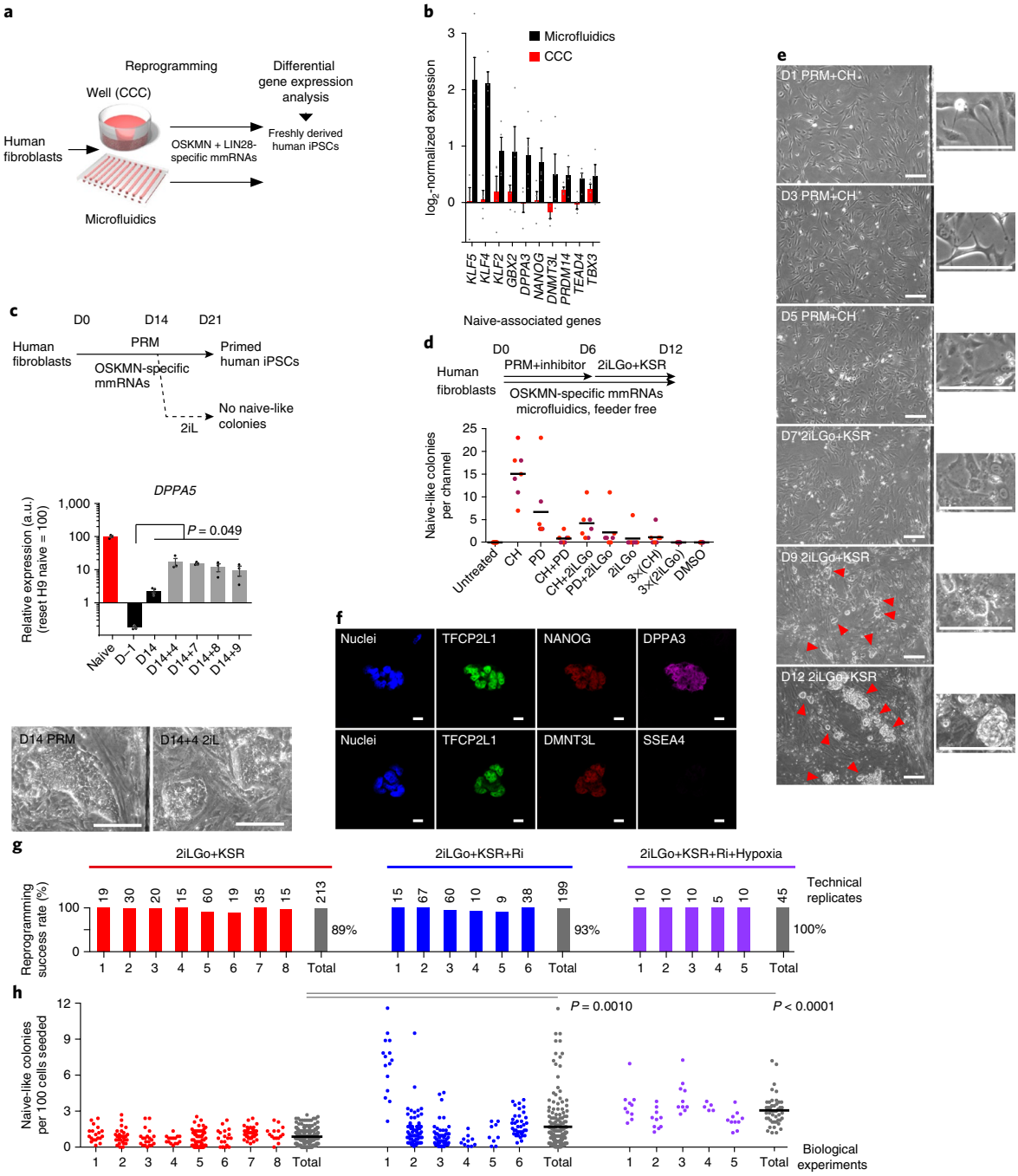

Fig. 1 | Devising a strategy for the transgene-free generation of naive-like iPSCs. a, Primary colonies of primed iPSCs generated in wells (CCC) or via microfluidics and mmRNAs underwent differential expression analysis[8]. **b**, The analysis revealed an increased expression of naive-associated markers in microfluidic conditions compared to CCC. Mean ± s.e.m., $n = 4$ biological replicates, shown as circles. **c**, OSKMN-specific mmRNA transfection in the presence of PRM led to the formation of primed human iPSC colonies at day 14 (D14; bottom, left panel), expressing *DPPA5* (black bar of middle panel; see also Supplementary Fig. 1b). Exposure of primary colonies to 2iL transiently upregulated *DPPA5* expression (grey bars) but did not lead to the formation of naive-like iPSCs. GAPDH served as the loading control. Bars indicate the mean ± s.e.m. of $n = 3$ biologically independent experiments, shown as circles. Unpaired two-tailed Mann–Whitney *U*-test. **d**, Evaluation of the effect of different inhibitors during the first 6 days in PRM. CHIR99021 (CH) promotes colony formation and facilitates efficient transfection (Supplementary Fig. 1d). Black bars indicate the means of technical replicates ($n = 9$ for untreated, $n = 7$ for all other conditions) pooled from 2 independent experiments, shown in two different colours. PD, PD0325901. **e**, Fibroblast reprogramming using PRM+CHIR99021 for 6 days followed by 2iLGo+KSR. Fibroblasts convert from a spindle to an epithelial-like morphology (D5). Cells become progressively compact and small colonies emerge (arrowheads). Representative pictures of eight independent experiments. **f**, Immunostaining of primary naive-like colonies at D15 reveals expression of naive markers and an absence of SSEA4. Representative pictures of three independent experiments. **g**,**h**, Addition of a ROCK inhibitor (Ri) and hypoxia increased both the success rate (**g**) and efficiency (**h**) of reprogramming. For **g**, each bar indicates the success rate of each independent experiment, calculated as the percentage of technical replicates (numbers above bars) containing naive-like colonies. For **h**, circles indicate technical replicates, whose number is shown in panel **g**. Black bars indicate the means of technical replicates pooled from independent experiments. Kruskal–Wallis test. See also Supplementary Fig. 1f. Scale bars 200 µm (**c**, **e**), 10 µm (**f**). Images from additional repeats for panels **e** and **f** have been deposited in Figshare under DOI: 10.6084/m9.figshare.7151777. See Supplementary Table 5 for all individual numerical values.

of *DPPA5* by providing the correct signalling environment. The cytokine LIF, the GSK3 inhibitor CHIR99021 and the MEK inhibitor PD0325901 in N2B27 medium (2iL) are present in several media for murine and human naive PSC culture[5,7]. Switching the media to 2iL at day 14 boosted and maintained *DPPA5* expression for the subsequent 8 days (Fig. 1c). However, they were morphologically indistinguishable from colonies kept in PRM (Fig. 1c, bottom panels), and rapidly collapsed upon passaging in 2iL. We reasoned that by day 14 in PRM, iPSCs have already lost the competence to reach the naive state, so we exposed cells to 2iL from the beginning of the reprogramming protocol (Supplementary Fig. 1c). However, no iPSC colonies formed, and we observed fibroblast senescence, reduced transfection efficiency and failure to execute the mesenchymal-to-epithelial transition (MET)[12], which normally occurs by day 6 (Supplementary Fig. 1c,d).

Based on these observations, we devised a two-step protocol. We used PRM for the first 6 days to avoid senescence and allow MET and subsequently applied 2iL supplemented with the PKC inhibitor Gö6983 (2iLGo) and 1% knockout serum replacement (KSR), which has previously been used to enhance and stabilize human naive pluripotency[10,11,13] (Fig. 1d, top panel). During the first 6 days, we also tested the effect of adding various inhibitors; we observed the formation of several naive-like colonies under different combinations, with the most efficient condition being PRM+CHIR99021 (Fig. 1d, bottom panel). Under these conditions, mmRNA transfection was robust (Supplementary Fig. 1e), MET occurred by day 5 and was followed by the appearance of naive-like colonies from day 9 (Fig. 1e, red arrowheads). Such colonies expressed the naive markers TFCP2L1, DPPA3 and DNMT3L (Fig. 1f), the core factor NANOG and were negative for the primed surface maker SSEA4[14]. The protocol appeared robust given that 89% of 213 individual microfluidic channels presented naive-like colonies (Fig. 1g,h, red), representing an efficiency of 0.87% (that is, 0.87 colonies generated for 100 fibroblasts seeded). Inhibition of Rho-associated protein kinase (ROCK) and hypoxic conditions[15,16] resulted in the generation of colonies in 100% of the microfluidic channels, representing an increased efficiency of 3.02% (Fig. 1g,h, purple). We conclude that delivering OSKMN-specific mmRNAs via microfluidics facilitates the rapid and robust generation of naive PSC colonies from human fibroblasts.

Long-term expansion and characterization of naive iPSCs. Different media and extracellular matrix proteins have been used for the culture of human naive cells[5,7]. Cells transfected with OSKMN were cultured via microfluidics under conditions of hypoxia for 6 days in PRM+CHIR99021+ROCK inhibition (Ri), followed by 6 days in different naive media, such as 4iLA[17] or the commercially available RSeT (Fig. 2a). Naive colonies, expressing *OCT4* and *KLF17*, formed robustly in all types of media (Fig. 2b; Supplementary Fig. 1g), with an increase in colony size in RSeT medium (Fig. 2c,d). Coating the microfluidic channels with fibronectin, Matrigel or laminin facilitated the generation of naive colonies with comparable efficiency (Supplementary Fig. 1g,h), indicating that our protocol can be easily adapted to different culture conditions. Next, we tested whether primary colonies could be expanded long term. To do so, we transferred them from microfluidic conditions to CCC on feeders in the respective media and under hypoxia and we succeeded in expanding them over multiple passages (Fig. 2e,f,h). RSeT again appeared to be the most robust among the three types of media, so it was used for all further experiments.

niPSCs, either freshly derived (passage 0) or after several passages (up to passage 21), displayed robust expression of naive pluripotency markers. Moreover, they exhibited a strongly reduced expression of the differentiation marker T and of the primed pluripotency markers SSEA4, OTX2 and ZIC2, indicating that a stable naive phenotype was achieved (Fig. 2e–h; Supplementary Fig. 2a). We also obtained niPSCs from four other somatic cells: human male foreskin fibroblasts (HFFs), female lung fibroblasts (WI-38 and IMR-90) and primary skin fibroblasts of an 80-year-old female donor (Fig. 2i; Supplementary Fig. 2b–d; Supplementary Table 1). All these niPSC lines expressed naive markers and expanded robustly over multiple passages, retaining naive morphology and high clonogenicity (Fig. 2e,f,j,k; Supplementary Fig. 3a,b).

Some human naive PSC lines, derived from embryos or from primed PSCs, display chromosomal instability after extensive culture[11,14,18]. Using DNA staining followed by flow cytometry, we observed that all lines were diploid for the first 8 passages. A stable diploid karyotype was observed in 6 out of 8 lines for up to 42 passages. Only 2 lines (HPD01 and HPD07) showed a fraction of tetraploid cells after 12 and 23 passages, respectively (Supplementary Fig. 3c). We isolated the diploid population via fluorescence activated cell sorting (FACS), and these cells then maintained a correct karyotype. A correct karyotype was also confirmed by Q-banding in five independent niPSC lines (Fig. 2k; Supplementary Fig. 3d). We conclude that reprogramming via microfluidics does not induce chromosomal abnormalities per se, which are instead probably due to extended culture.

A naive state transcriptome is rapidly acquired during niPSC generation. We performed transcriptome profiling of niPSCs, primed iPSCs generated via microfluidics (Supplementary Fig. 1c) and human fibroblasts using RNA-sequencing (RNA-seq). Unsupervised clustering based on specific markers grouped niPSCs together with other previously described human naive cells[10,11,17,18], clearly separating them from primed PSCs (Fig. 3a). Similar results were obtained by analysing the whole transcriptome (Fig. 3c; Supplementary Fig. 4a). Naive iPSCs have been recently generated from human fibroblasts via Sendai virus-mediated delivery of OSKM without a microfluidics approach[19,20]. Depending on the media used, some of those lines grouped together with embryo-derived naive PSCs and with niPSCs (Fig. 3a,c; Supplementary Figs. 4a–c and 8; see also Discussion).

The presence of a unique set of transposable elements expressed by cleavage-stage embryos characterize the human naive pluripotent state[17]. The transposable element profile of niPSCs was clearly distinct from those of primed cells and comparable to previously described naive PSCs (Fig. 3b). Notably, primary colonies ("niPSCs p0" in Fig. 3a–c) collected at day 15 were transcriptionally indistinguishable from established naive PSCs. We conclude that our reprogramming strategy enables the rapid and full acquisition of a human naive transcriptome.

Next, we determined whether our system would allow the study of the trajectories of iPSC generation from fibroblasts. As proof of principle, we transfected fibroblasts with OSKMN-specific mmRNAs either under the optimized protocol (described in Fig. 2a) to generate niPSCs or in PRM to generate primed iPSCs. A gene expression analysis of a panel of naive- and primed-state markers revealed that the two reprogramming trajectories diverged significantly after day 8 (Fig. 3d,e; Supplementary Fig. 4d). Between days 5 and 8, cells undergo dramatic morphological changes indicative of a MET (Fig. 1e). We measured a reduction in mesenchymal markers (Fig. 3f, grey) and a gradual activation of epithelial markers (Fig. 3f, orange) for both reprogramming protocols, indicating that MET is a common early molecular event. We conclude that our reprograming method allows the study of distinct trajectories that lead to either primed or naive pluripotency.

Global reduction in repressive epigenetic modifications in niPSCs. Naive pluripotency is associated with a reduction in repressive epigenetic modification, such as H3K9me3 or cytosine methylation[7,10,14,17,18]. Immunostaining experiments showed that both markers were reduced in multiple niPSCs compared to primed PSCs

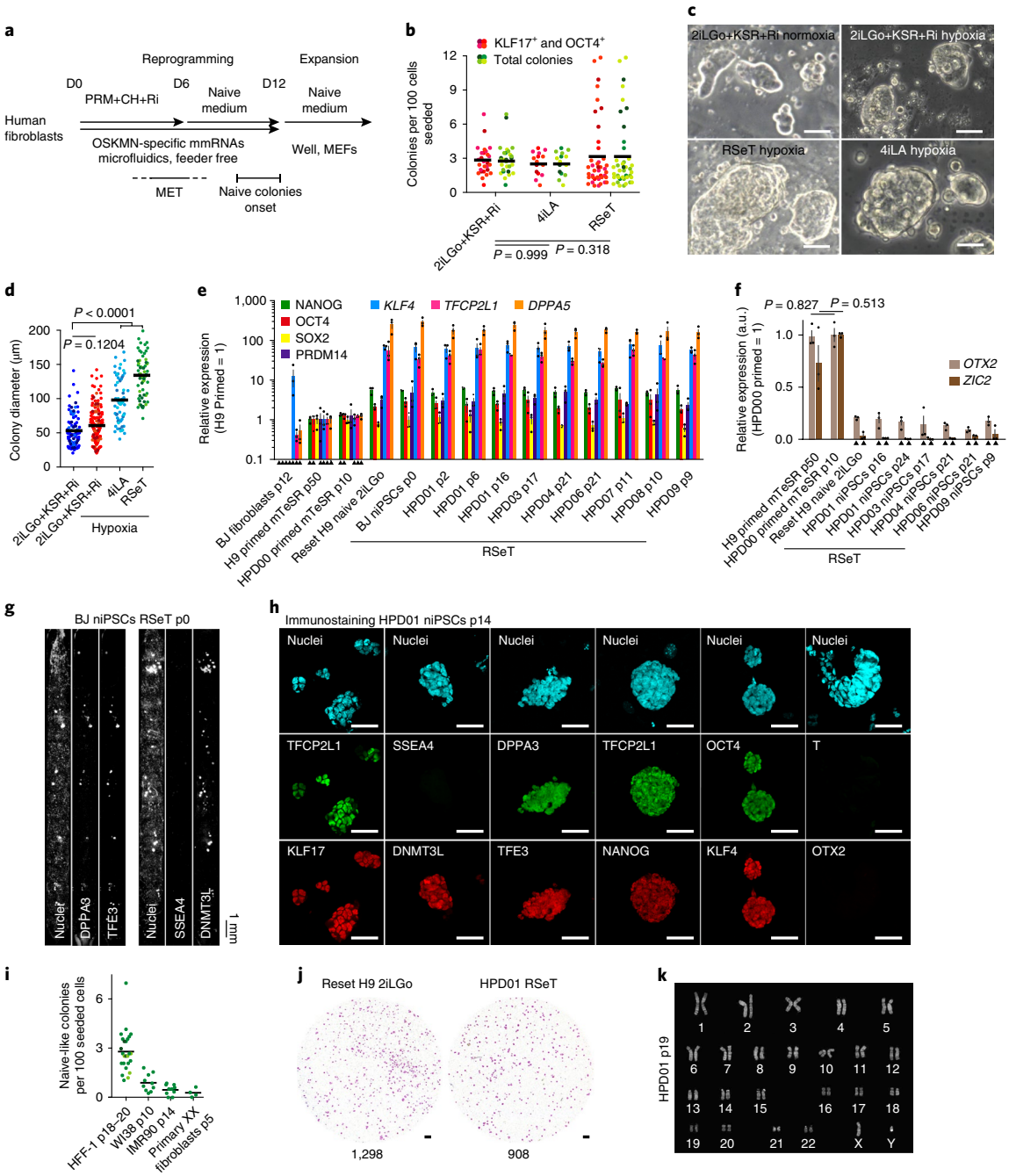

Fig. 2 | Optimized generation and stabilization of the naive phenotype. a, Optimized naive reprogramming setup. **b**, Effect of different naive media applied from D6. Colonies were quantified at D15 either by their compact morphology (green circles) or by the coexpression of *KLF17* and *OCT4* (red circles). Bars indicate the means of technical replicates (circles, $n = 28$, 15 and 40 replicates for 2iLGo+KSR+ROCK inhibition, 4iLA and RSeT, respectively) pooled from 4 independent experiments, shown in different shades of colours. Kruskal–Wallis test. **c**, Morphology of representative colonies. **d**, Quantification of colony size as equivalent diameters (Methods). Bars indicate the means of technical replicates (circles, $n = 98$, 163, 67 and 65 replicates for 2iLGo+KSR+ROCK inhibition, 2iLGo+KSR+ROCK inhibition in hypoxia, 4iLA in hypoxia and RSeT in hypoxia, respectively) obtained in 3 independent experiments, shown in different colours. Kruskal–Wallis test. **e**, Gene expression analysis by qPCR of niPSCs, fibroblasts, primed PSCs and reset H9 cells at different passage (p) numbers. A significant reduction in naive markers relative to reset H9 cells was observed only in fibroblasts and primed PSCs. **f**, Reduction in primed pluripotency markers in niPSCs. Expression relative to HPD00 primed iPSCs. In both **e** and **f**, GAPDH served as the loading control, arrowheads indicate samples with statistical significance (unpaired two-tailed Mann–Whitney *U*-test), bars indicate mean ± s.e.m. of $n = 3$ independent experiments, shown as circles. **g**, Immunostaining of all the microfluidic channels at D15. See also Supplementary Fig. 2a. **h**, Immunostaining for pluripotency and differentiation markers. Representative images from 1 out of 2 independent experiments showing consistent results. **i**, Generation of niPSCs from different somatic cells. Bars indicate the means of technical replicates (circles, $n = 25$, 10, 10 and 4 for HFF-1, WI38, IMR90 and primary XX fibroblasts, respectively) pooled from 3 independent experiments for HFF-1 and 1 experiment for the other 3 lines. **j**, Clonal assay of HPD01 niPSCs and reset H9 cells. See Supplementary Fig. 3b for results obtained in five other niPSCs. Representative images of 1 out of 3 independent experiments showing consistent results. **k**, Q-banding showing a normal karyotype in HPD01. Similar results were obtained in 7 out of 7 metaphases. See also Supplementary Fig. 3d. Scale bars, 50 µm (**c**,**h**), 1 mm (**g**,**j**). Images from additional repeats for panels **c**,**h** and **j** have been deposited in Figshare under DOI: 10.6084/m9.figshare.7151801. See Supplementary Table 5 for all numerical values and *P* values.

(Fig. 4a,b; Supplementary Fig. 5a). Consistent with the reduced levels of 5-methylcytosine (5mC), we observed decreased levels of the DNA methyltransferases *DNMT3A* and *DNMT3B*, a dramatic increase of the catalytically inactive *DNMT3L*, and of the 5mC oxidases *TET1* and *TET2* (Fig. 4c). We then determined the pattern of genome methylation using reduced representation bisulfite sequencing (RRBS) in five niPSC lines, three fibroblasts from which they were derived and an isogenic primed iPSC line (Supplementary Table 1). The global methylation levels (Fig. 4d) and distributions (Supplementary Fig. 5b) of niPSCs were comparable to those of human blastocysts and other naive PSCs[10,14,17]. Unsupervised clustering of the methylation pattern clearly separated naive PSCs from somatic cells and primed iPSCs (Fig. 4e). A set of promoters hypomethylated in human blastocysts compared to conventional primed human ESCs[21,22] were found to be highly methylated in somatic and primed pluripotent cells, but hypomethylated in niPSCs, which is in line with their naive identity (Fig. 4f).

Analysis of imprinted and X-linked genes status. DNA methylation is crucial for genomic imprinting, the phenomenon whereby either the maternal or paternal copy of a gene is expressed. A previous study[14] reported that DNA methylation is lost at imprinted loci in embryo-derived naive human ESCs; therefore, we tested whether niPSCs would have the same defects. We analysed 67 bona fide imprinted loci, divided into the following 3 classes (Fig. 5a): placental, maternal and paternal. Placental loci, methylated in human oocytes and blastocysts, but not in somatic cells (blood)[23], were hypomethylated in fibroblasts and remained so after reprogramming. Maternally imprinted loci lost methylation only in niPSCs, as previously reported[14], while a fraction (5 out of 8) of paternally imprinted loci surprisingly maintained methylation in all niPSCs (Fig. 5a,b). Loss of DNA methylation at the imprinted regions should result in the aberrant biallelic expression of the cognate transcripts. We interrogated RNA-seq data, taking advantage of annotated single nucleotide polymorphisms (SNPs) in *MEG3* mRNA, and found monoallelic expression in somatic cells and biallelic expression in all naive PSCs, in agreement with its reduced methylation (Fig. 5c).

Female naive pluripotent cells contain two active X chromosomes[24], characterized by low levels of DNA methylation at CG-rich regions called CpG islands. CpG island methylation on the X chromosome in female niPSCs was dramatically reduced (Fig. 5d) relative to female fibroblasts, and were comparable to levels observed in male cells. An analysis of SNPs on X-linked genes showed biallelic expression of several transcripts in female niPSCs (Fig. 5e). Finally, high *XIST* expression was detected in female niPSCs compared to male niPSCs and somatic cells (Fig. 5f), as previously reported for female naive pluripotent cells in vitro and in vivo[24]. These results indicate a reactivation of both X chromosomes in female niPSCs.

Naive iPSCs display high mitochondrial activity. Mitochondrial activity has been reported to be higher in both murine and human naive cells compared to their primed counterparts[5,7]. We measured the mitochondrial membrane potential by tetramethylrhodamine methyl ester (TMRM) staining and it was barely detectable in primed cells (Fig. 6a). Conversely, niPSCs displayed robust TMRM staining. Both niPSCs and reset H9 cells[10] displayed increased mitochondrial transcripts compared to primed PSCs (Fig. 6b) and fibroblasts (Fig. 6c), as previously reported in murine PSCs[25]. We conclude that niPSCs display epigenetic and metabolic features consistent with a naive pluripotent state.

Characterization of the differentiation potential of niPSCs. Recent reports have shown that some human naive PSC lines are lineage biased and fail to differentiate towards mature cell types[26,27], thus limiting their use for disease modelling and developmental studies. We tested the differentiation potential of three isogenic niPSC lines (HPD01, HPD03 and HPD04) to gauge the technical variability of our reprogramming system. A primed iPSC line (HPD00) was included as a control. First, we applied three different monolayer differentiation protocols and found that all three niPSC lines expressed mesoderm and endoderm markers after 6 days of differentiation, while ectoderm markers were robustly expressed around day 12 (Fig. 7a). As expected, primed iPSCs differentiated faster (Supplementary Fig. 6a, top). Second, we performed embryoid body differentiation and detected the expression of multiple markers of the three germ layers at day 22 in all 3 niPSC lines (Fig. 7b; Supplementary Fig. 6b). Third, after 50 days of embryoid body differentiation (Fig. 7c; Supplementary Fig. 6c), we detected comparable levels of markers of mature cell types in all iPSC lines. Fourth, we directly tested the ability of niPSCs to form mature cell types, adapting protocols previously defined for primed PSCs. After 15 days of a hepatic differentiation protocol[28], we observed polygonal-shaped hepatocyte-like cells expressing the mature markers CYP3A and HNF4A, while a neuronal differentiation protocol[29] successfully generated cells positive for MAP2, TUJ1 and NeuN (Fig. 7d,e; Supplementary Fig. 7a,b). Finally, all differentiation assays were repeated with an independent niPSC line (HPD06) derived from HFFs (Supplementary Figs. 6 and 7). In summary, these results demonstrate that niPSCs are pluripotent, respond to differentiation cues and are able to form mature cell types.

The confined microfluidic environment promotes establishment of naive pluripotency. Initial attempts to generate niPSCs using mmRNA under CCC failed. We reinvestigated this approach with our optimized protocol (Fig. 8a), comparing microfluidic conditions and CCC. The generation of niPSCs via microfluidics was robust (100% of channels contained colonies, with an efficiency of $3.7 \pm 1.0\%$ and $3.1 \pm 0.4\%$ for BJ and HFF-1 cells, respectively). Conversely, in CCC, only 25% of wells contained naive-like colonies, with efficiency values of 0.2% and 0.04% for BJ and HFF-1 cells, respectively. Interestingly, a mixture of colonies displaying primed and naive morphology emerged in CCC (Fig. 8b, blue and red circles), a phenomenon never observed in the microfluidic culture conditions. These results strongly indicate that the confined microfluidic microenvironment promotes the generation of naive iPSCs over primed iPSCs, and that the culture conditions used (that is, RSeT medium with feeder cells) facilitate the propagation of different pluripotent cell types (Supplementary Fig. 8). By adding the FGF receptor inhibitor PD173074, we selectively eliminated primed-like iPSCs (Fig. 8c, right) and obtained a stable niPSC line (HPD05) transcriptionally comparable to other naive iPSCs and clearly distinct from primed PSCs (Figs. 3a–c and 8d,e).

To directly test whether the confined microenvironment promotes reprogramming to the naive state, we generated microfluidic chips with channels of different heights, inside which we then reprogrammed BJ fibroblasts. At the height of 200 µm, used in all other experiments, we obtained colonies at high efficiency, and all colonies expressed both KLF17 and OCT4 (Fig. 8f; Supplementary Fig. 7c). Increasing the height caused a reduction in both the number of colonies and the percentage of double-positive colonies, indicating the presence of either primed or partially reprogrammed colonies. Finally, decreasing the channel height to 100 µm caused a reduction in colony number, which is probably due to rapid exhaustion of the culture medium. We conclude that reprogramming to naive pluripotency is enhanced with an optimal spatial confinement.

Finally, we determined whether microfluidics could also promote the conversion of primed PSCs to naive pluripotency. We generated primed H9 human ESCs expressing *NANOG* and *KLF2* (NK2) in a doxycycline-inducible manner, as previously reported[10,11]. Exposure to 2iL+doxycycline led to the formation of naive colonies, with a fivefold increase when using microfluidic conditions compared to CCC (Fig. 8g). In contrast, no naive colonies were observed under

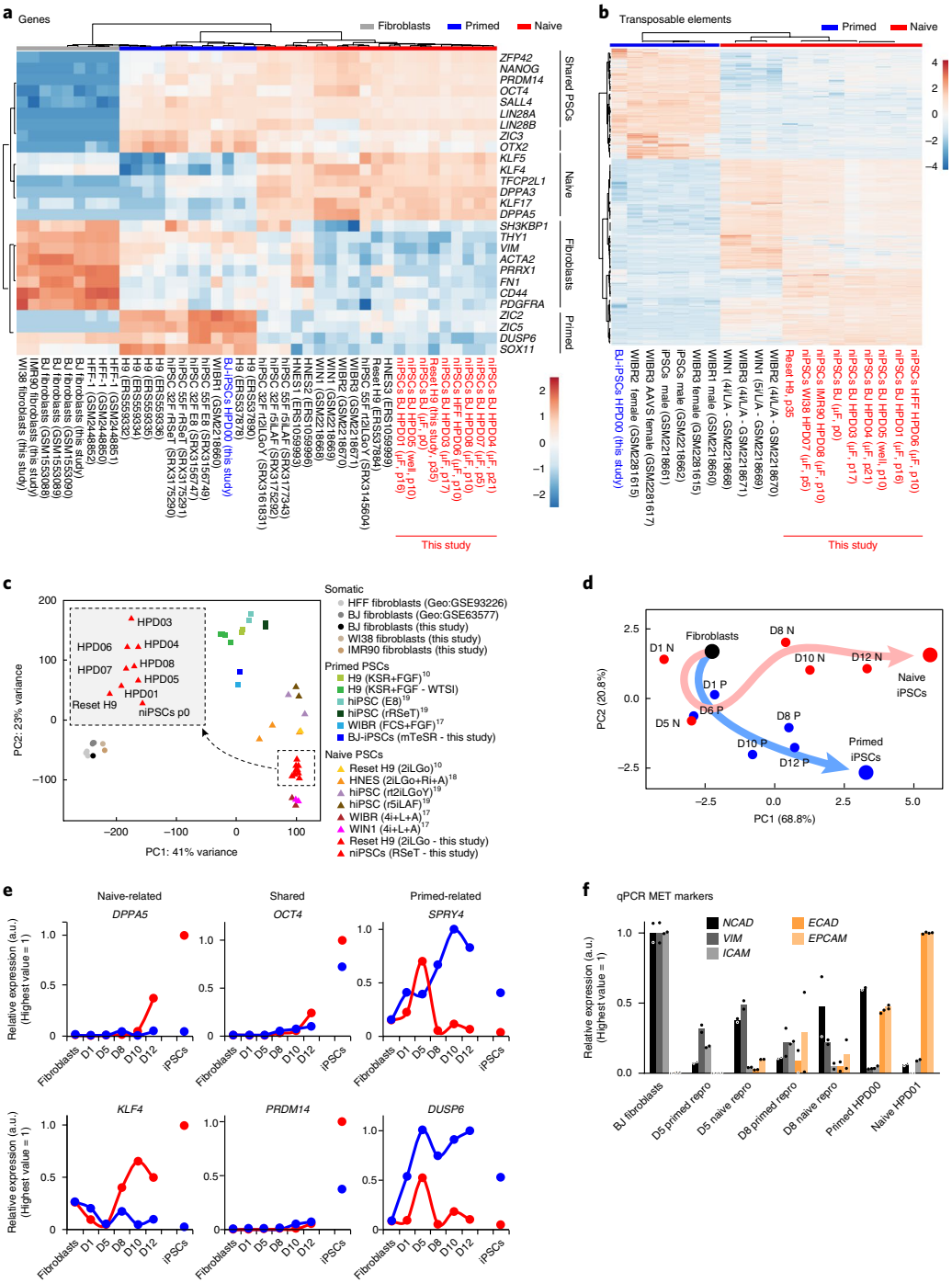

Fig. 3 | A naive-state transcriptome is rapidly acquired during niPSC generation. a, Heatmap of unsupervised hierarchical clustering based on somatic- and pluripotency-associated markers highly expressed in human naive ESCs and human embryos (for example *TFCP2L1*, *KLF17*, *DPPA5* and *KLF4*)[10,11,18,34], or specifically expressed in primed PSCs (for example, *ZIC2*, *ZIC5* and *SOX11*)[11] or in fibroblasts (for example *VIM*, *FN1* and *CD44*)[35]. The naive and primed human iPSCs (hiPSCs) generated in this study, shown in red and blue text, respectively, cluster together with previously characterized PSCs. Values displayed correspond to the gene expression level (normalized $\log_2$ pseudocounts) scaled by the row mean. μF, microfluidics. **b**, Heatmap of unsupervised clustering based on the expression of transposable elements[17] in primed and naive PSCs. Values displayed are transposon expression level (normalized $\log_2$ pseudocounts) scaled by row mean. **c**, PCA of RNA-seq samples. Inset shows naive PSC lines derived in this study. As an internal control, we sequenced the transcriptome from reset H9[10] cells cultured in parallel to niPSCs. See also Supplementary Fig. 4a for a heatmap of unsupervised clustering based on the transcriptome. **d**–**f**, Gene expression analysis by qPCR of reprogramming trajectories of cells reprogrammed via microfluidics. Fibroblasts were plated at the same density and transfected with OSKMN either in PRM (blue) to generate primed iPSCs or in conditions facilitating niPSC formation (red), as in Fig. 2a. Samples were collected on D1, D5, D8, D10 and D12, and the expression of pluripotency markers was analysed. $n=2$ biologically independent experiments. See also Supplementary Fig. 4d. PCA analysis (**d**) shows that reprogramming towards primed (P) and naive (N) pluripotency clearly diverge at D8. Expression of the indicated markers during reprogramming to either naive or primed pluripotency (red and blue, respectively; **e**). Reduced expression of the mesenchymal markers (grey) and gradual increase in epithelial (orange) markers at D5 and D8 during reprogramming (repro; **f**). For **e** and **f**, The expression relative to the highest value was calculated, and GAPDH served as the loading control. Bars indicate the means of two independent experiments, shown as circles. See Supplementary Table 5 for all individual numerical values.

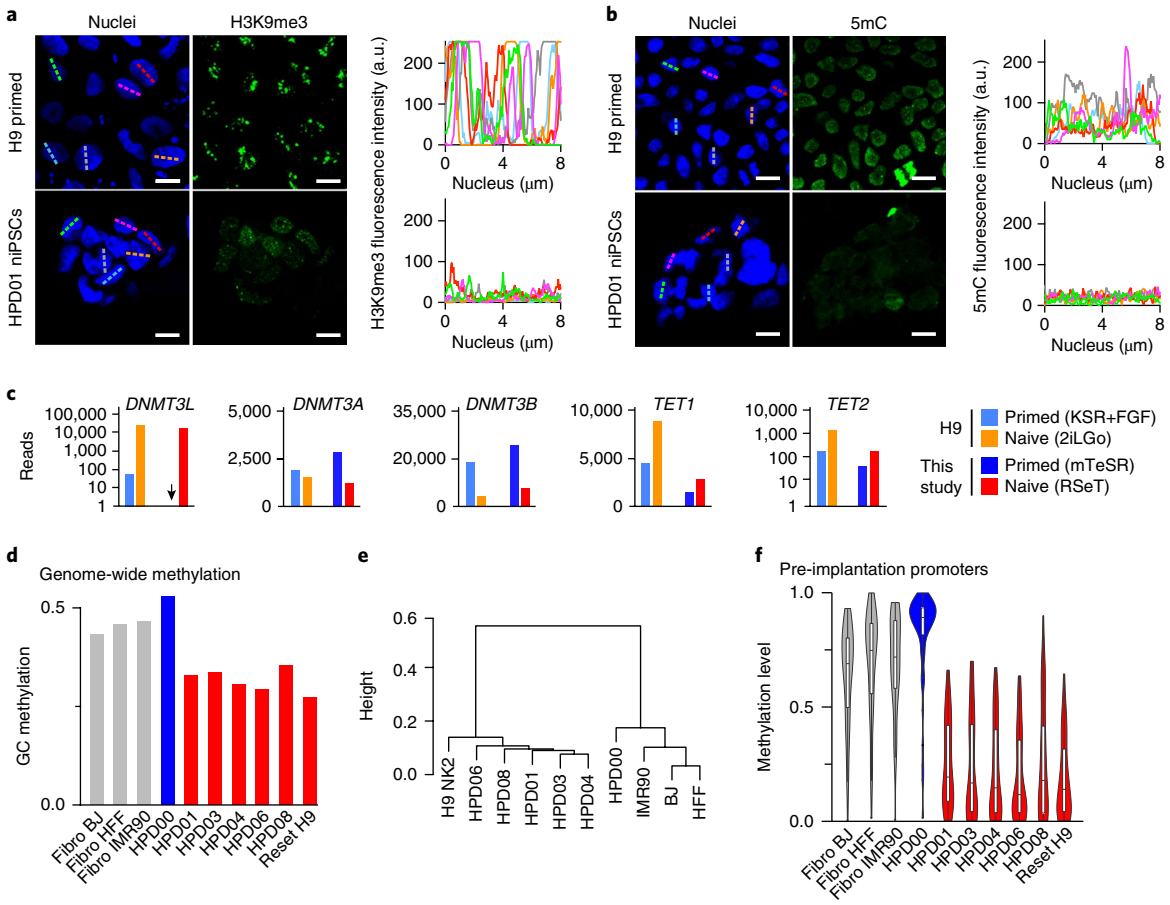

Fig. 4 | Global reduction in repressive epigenetic modifications in niPSCs. a,**b**, niPSCs show significantly lower histone H3K9me3 (**a**) and DNA (**b**) methylation compared to primed PSCs as determined by immunofluorescence staining. Left: representative images. Right: staining intensity quantification via image analysis of six randomly selected nuclei. See Supplementary Fig. 5 for similar results obtained in five independent niPSC lines. **c**, Expression of known regulators of DNA methylation in primed and naive PSCs measured by RNA-seq. **d**, Average genome-wide CG methylation levels measured by RRBS in somatic (grey), primed (blue) and naive (red) iPSCs. Reset H9 cells[10] were also included as a control. Fibro, fibroblasts. See also Supplementary Fig. 5b. **e**, Unsupervised hierarchical clustering of the genome-wide methylation pattern separates niPSCs from somatic and primed cells. **f**, Violin plots showing DNA methylation levels on 52 promoters identified[10,21] as hypomethylated in pre-implantation blastocysts relative to primed human ESCs. Box plots show medians and first and third quartiles. Violins span the interval between the maximal and minimal value of each sample. See also Supplementary Fig. 5b. Scale bars, 10 μm (**a**,**b**). See Supplementary Table 5 for all individual numerical values.

conditions promoting primed pluripotency (with FGF and KSR), even under microfluidic conditions. We conclude that a confined environment promotes the activation of the naive pluripotency network, regardless of the starting cell type, but only in conditions supporting naive pluripotency.

Discussion

Microfluidics enables the manipulation of cells in a confined environment, whereby autocrine and paracrine factors are more concentrated[8]. The use of microfluidics facilitates the robust and efficient generation of naive iPSCs, reducing by several orders of magnitude the costs and number of somatic cells required. The efficiency of reprogramming in our system (3–5%) is substantially higher than those of other non-integrating methods for generating conventional human iPSCs[30]. We speculate that cells undergoing reprogramming release extracellular factors that support survival or cell fate transition. Identification of such factors could help improve the efficiency of reprogramming protocols under CCC.

Previous studies have reported the conversion of human somatic cells to naive pluripotency[10,11,13], but they relied on secondary systems, viral vectors or required two rounds of reprogramming (OSKM followed by NK2) via an intermediate primed state. Our results show that it is possible to directly convert somatic cells to naive pluripotency through the transient transfection of OSKMN-specific mmRNAs. While *KLF2* has previously been used to convert primed PSCs to naive pluripotency[10,11,17], it was not included in our reprogramming cocktail because of the functional redundancy between *KLF4* and *KLF2*, as observed in murine cells[31]. Correspondingly, *KLF4* expression has recently been used to convert human primed PSCs to naive pluripotency[19].

While this manuscript was under revision, two studies were published reporting the generation of naive iPSCs from somatic cells via the expression of OSKM using Sendai viruses[19,20]. In those studies, iPSCs generated in t2iLGöY (titrated 2iL plus the protein kinase C inhibitor Gö6983 and the ROCK inhibitor Y27632) or 5iLAF (a cocktail of 5 inhibitors targeting GSK3, MAPK/Erk, BRaf, Src and ROCK, plus the cytokines LIF, Activin A and FGF2) media consistently displayed a transcriptome similar to embryo-derived naive PSCs and niPSCs (Fig. 3a,c; Supplementary Fig. 4a–c), with robust expression of naive markers such as KLF17, TFCP2L1 and DNMT3L. Conversely, the use of RSeT medium led to variable results: in one study[19], iPSCs generated in RSeT were transcriptionally similar to primed iPSCs (Fig. 3a,c), while in the other study[20], iPSC lines displayed either a naive or an intermediate phenotype,

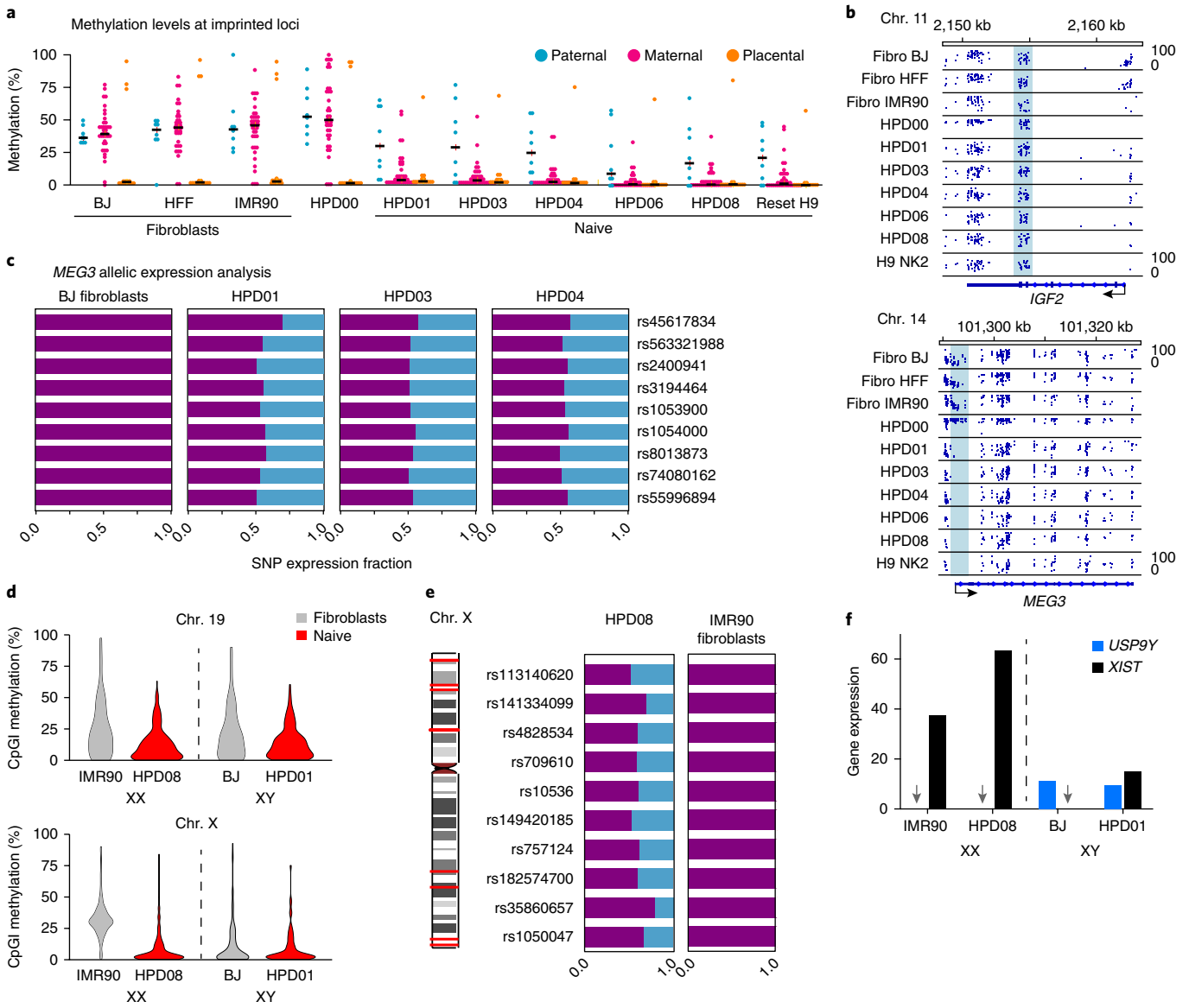

Fig. 5 | Analysis of imprinted and X-linked genes status. a, Methylation levels at imprinted loci. Maternal or paternal loci are those in which the maternal or the paternal alleles are methylated[23]. Placental loci are characterized by the transient methylation of the maternal allele in oocytes and blastocysts that is then lost in somatic cells. Each circle indicates the average methylation level at an imprinted locus. Black bars indicate the mean methylation levels of each class of loci ($n = 8$, 39 and 15 for paternal, maternal and placental, respectively). Somatic and primed cells show robust methylation at maternal and paternal loci, while niPSCs retain only partial methylation at paternal loci. See also Supplementary Table 5. **b**, Methylation patterns at the human *IGF2* and *MEG3* loci. The *y* axis indicates the level of methylation. Imprinted regions are highlighted in light blue. **c**, SNP analysis of the *MEG3* transcript reveals monoallelic expression in BJ fibroblasts and biallelic expression in three isogenic niPSC lines. **d**, Violin plots show methylation levels at CpG islands (CpGI) on the X chromosome and at chromosome (Chr.) 19. Methylation on chromosome 19 is reduced to the same extent in male and female niPSCs compared to fibroblasts. Methylation on chromosome X is very low in both male cells and female niPSCs compared to female fibroblasts. Violins span the interval between the maximal and minimal value of each sample. **e**, SNP analysis of X-linked genes. Right: histograms showing the fraction of SNP expression. Female fibroblasts show monoallelic, while female niPSCs show biallelic expression. Left: vertical bars indicate the position of the SNPs analysed on the X chromosome. **f**, Expression levels measured by RNA-seq of *XIST*. Female niPSCs express high levels of *XIST*. The Y-linked gene *USP9Y* is shown as a control.

with heterogeneous expression of KLF17 (Supplementary Fig. 4b,c). In contrast, we used OSKMN-specific mmRNAs, microfluidic conditions and RSeT and obtained niPSCs homogeneously expressing several naive markers (Fig. 2g,h; Supplementary Fig. 2a,b) and displaying transcriptomes similar to other embryo-derived naive PSCs (Fig. 3a–c; Supplementary Fig. 4a). These disparate outcomes with RSeT medium may be due to other methodological parameters, such as the use of Sendai viruses versus mmRNAs, OKSM versus OKSMN, and CCC versus microfluidic conditions.

We compared RSeT, 2iLGo and 4iLA media during naive iPSC generation via microfluidics (Fig. 2b; Supplementary Fig. 1g) and observed similar efficiencies regarding the formation of iPSC colonies positive for both KLF17 and OCT4. This result suggests that the confined environment promotes the acquisition of naive pluripotency and that this effect is dominant over media composition. Indeed, all three types of media facilitated the robust expansion of established niPSCs (Supplementary Fig. 8). To directly test the effect of the confined environment, we increased the height of the

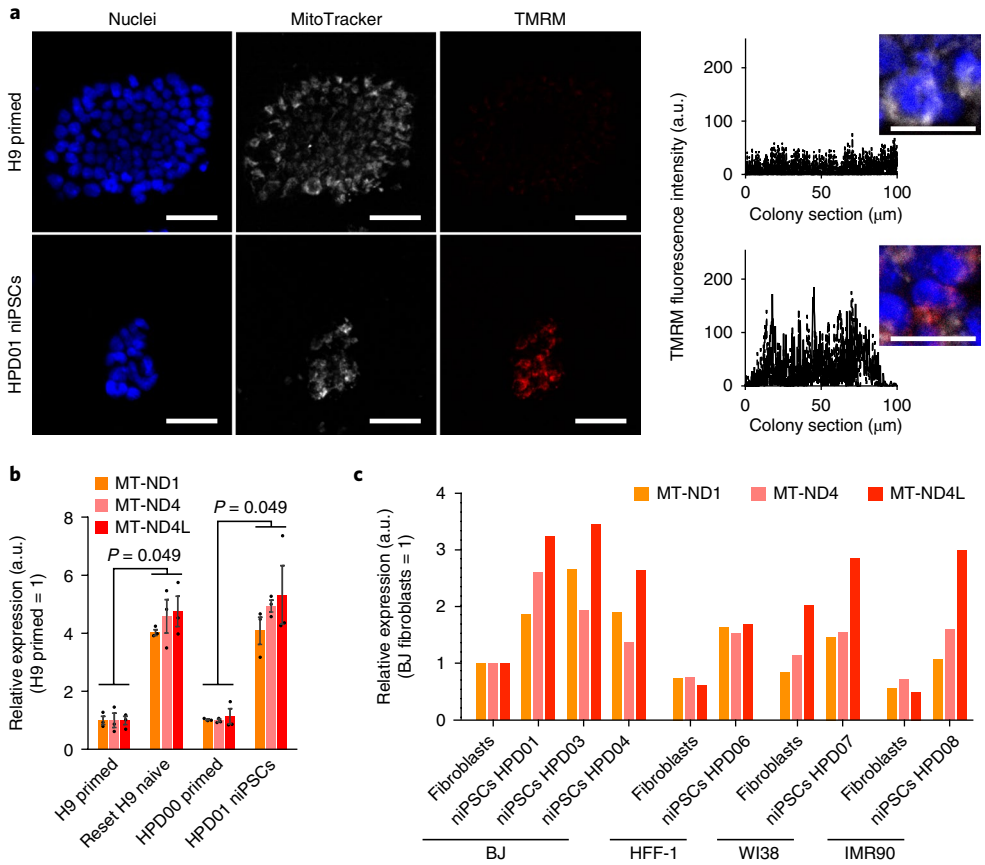

Fig. 6 | Enhanced mitochondrial activity in niPSCs. a, TMRM fluorescent accumulation in niPSCs shows a significant increase in mitochondrial activity compared to primed PSCs. Left: representative images of PSC colonies. MitoTracker stains all mitochondria, while the TMRM signal is dependent on the mitochondrial membrane potential. Right: quantification of TMRM intensity across PSC colonies. $n = 4$ and 5 technical replicates for H9-primed and HPD01 niPSCs, respectively, from 1 biological experiment. Images from two additional independent experiments, showing similar results have been deposited in Figshare under DOI: 10.6084/m9.figshare.7151816. **b**, Genes encoded by the mitochondrial DNA are upregulated in naive PSCs compared to primed PSCs. Gene expression analysis by qPCR of the indicated PSC lines for three transcripts of the mitochondrial genome. Expression is relative to primed H9 ESCs was calculated. GAPDH served as the loading control. Bars indicate the mean ± s.e.m. of three independent experiments, shown as circles. *P* values calculated using unpaired two-tailed Mann–Whitney *U*-test. **c**, Analysis by RNA-seq of the indicated mitochondrial transcripts in fibroblasts and niPSCs. Expression was normalized to BJ fibroblasts. Scale bars, 50 μm (**a**, left), 25 μm (**a**, right). See Supplementary Table 5 for source data.

microfluidic channels (Fig. 8f; Supplementary Fig. 7c) and performed reprogramming under CCC rather than microfluidics. We observed that only in an optimally confined environment could we obtain a homogenous population of naive colonies at high efficiency (Fig. 8a–c). Interestingly, using both a higher channel height and CCC, we obtained a mixed population of primed- and naive-like colonies in RSeT (Fig. 8b), resembling the results reported in a previous study[20]. Thus, we argue that the confined microfluidic environment is instructive for the induction of naive pluripotency, and that RSeT is a more permissive medium, which allows the formation and expansion of a range of different pluripotent phenotypes under CCC.

A recent study[32] investigated the acquisition of naive pluripotency with a secondary system of in vitro-differentiated fibroblasts bearing a doxycycline-inducible *OSKM* cassette. A naive transcriptome was observed after 20 days of reprogramming, and transgene-independent iPSCs were obtained at day 24. In contrast, our method appears more rapid, given that by day 15 naive iPSC colonies showed a naive transcriptome (Fig. 3a–c; Supplementary Fig. 4a) and a homogeneous protein expression profile of naive factors (Figs. 1f and 2g; Supplementary Fig. 2a,b), and could be readily expanded without further mmRNA delivery. Since the generation of primed iPSCs from fibroblasts with mmRNA was also faster than with Sendai virus[12,30], we suggest that the rapid and efficient generation of naive iPSCs achieved in this study is attributable to the combination of mmRNA and microfluidic technologies. It will be interesting to exploit our system to study at high resolution the trajectories followed by somatic cells towards naive pluripotency. As a proof of principle, we performed bulk analyses on fibroblasts reprogrammed to either naive or primed pluripotency and observed a common initial phase followed by two divergent trajectories (Figs. 3d,e and 8h).

Chromosomal abnormalities in some naive PSCs have been reported[14,17,18]. Only two out of eight lines displayed a tetraploid fraction of cells after long-term expansion, which could be eliminated by FACS. We conclude that our method allows the generation of karyotypically normal naive PSCs, possibly due to the use of mmRNAs, which have been shown to cause the lowest number of chromosomal abnormalities compared to other non-integrating methods[30]. We argue that the prolonged expansion in culture, together with some yet unknown factors, might negatively influence genomic stability.

In conclusion, we showed that the confined microfluidic environment strongly enhances the direct conversion of somatic cells

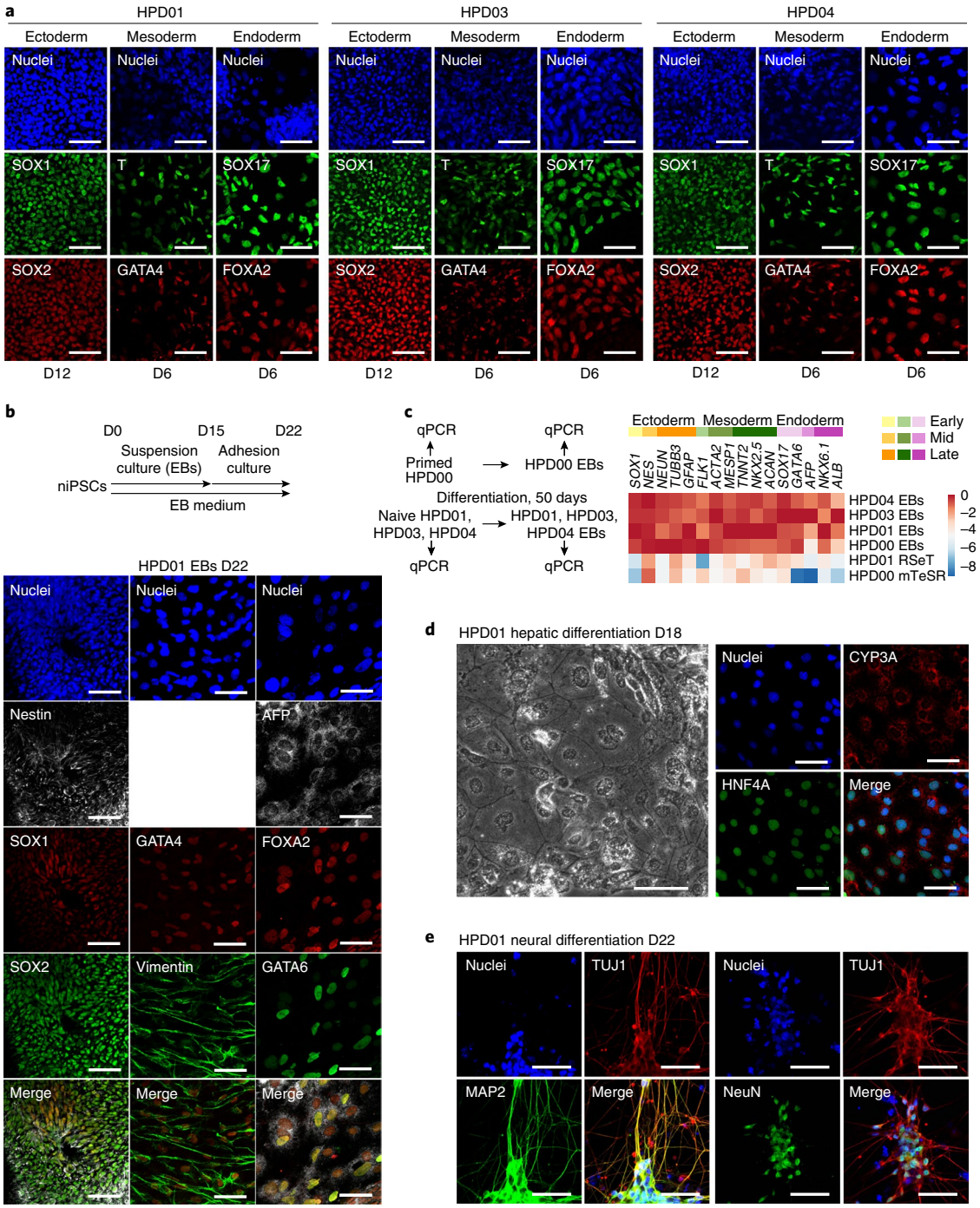

Fig. 7 | Characterization of the differentiation potential of niPSCs. a, Monolayer differentiation of three isogenic niPSC lines (HPD01, HPD03 and HPD04). Mesoderm and endoderm markers were detected in all lines at D6 of differentiation, while neurectoderm markers were detected at D12. $n=4$ biologically independent experiments. See also Supplementary Fig. 6a. **b**, niPSCs form embryoid bodies (EBs). After 15 days of culture in suspension, EBs were plated on Matrigel to allow the spreading of differentiated cells for 7 days. Immunostaining confirmed the expression of markers of the three germ layers. $n=4$ biologically independent experiments. See also Supplementary Fig. 6b. **c**, Starting from BJ human fibroblasts, we generated isogenic primed (HPD00) or naive PSCs (HPD01, HPD03 and HPD04) using mmRNAs via microfluidics. Expression of markers of the three germ layers was measured by qPCR before and after 50 days of EB differentiation. Early, mid and late markers are shown in different shades of colours. Robust and comparable induction of differentiation markers was observed in niPSCs and primed iPSCs. For each preparation, we pooled at least 15 EBs that were >0.5 mm in size. See also Supplementary Fig. 6c. Data are expressed as the $\log_2$ fold-change relative to the highest value. GAPDH served as the loading control. $n=2$ independent experiments. **d**, In vitro hepatic differentiation of niPSCs generated mature polygonal hepatocyte-like cells expressing the specific markers HNF4A and CYP3A. $n=2$ biologically independent experiments conducted in 2 niPSC lines. See also Supplementary Fig. 7 for similar results obtained with an independent niPSC line. **e**, In vitro differentiation of niPSCs in mature neurons. At D22 of differentiation, we observed cells positive for the pan-neuronal markers TUJ1, NeuN and MAP2. We should stress that the efficiency of hepatic and neural differentiation was higher in HPD01 and HPD06 lines, indicating the need for protocols optimized for naive PSCs. $n=2$ biologically independent experiments. See also Supplementary Fig. 7 for similar results obtained with an independent niPSC line. Scale bars, 50 μm (**a**,**b**,**e** and immunofluorescence in **d**), 100 μm (bright-field in **d**).

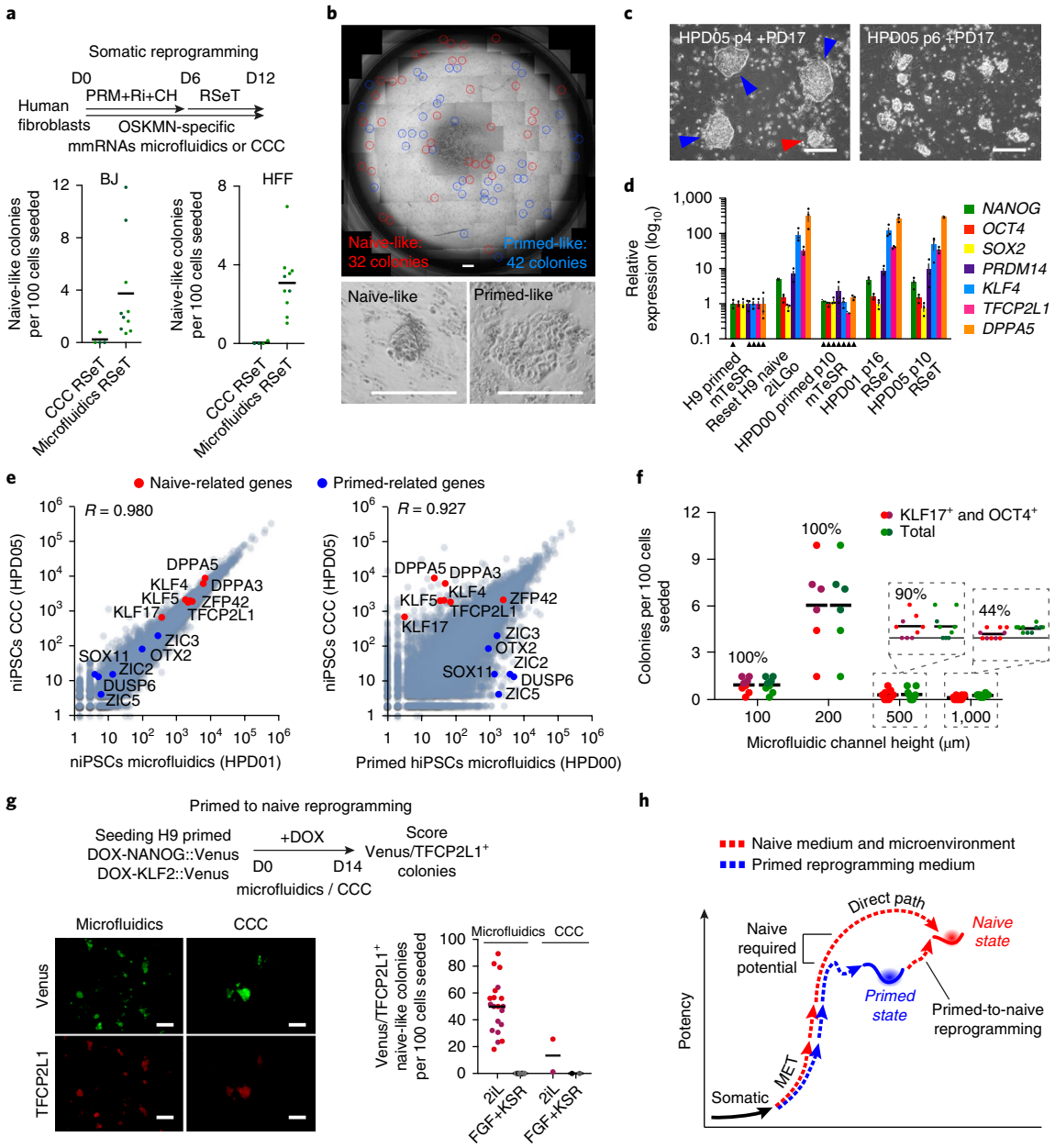

Fig. 8 | Comparison of naive iPSC generation under microfluidic conditions and CCC. a, Comparison of our optimized protocol in microfluidic conditions versus CCC. Bars indicate the means of technical replicates (circles, $n=10$ in microfluidics and $n=4$ in CCC) pooled from 2 independent experiments performed with BJ cells and 2 with HFF lines, shown in different shades of green. **b**, At D12 in CCC, a mixed population of naive- and primed-like colonies emerged. $n=2$ independent experiments. **c**, Clearance of primed-like colonies after culture in presence of RSeT+FGF receptor inhibitor PD173074 for 6 passages. Colonies with primed morphology (blue arrowheads) were still present at passage 4, together with naive-like colonies (red arrowheads). **d**, niPSCs generated in CCC (HPD05) show an expression profile comparable to other naive PSCs. Expression relative to primed H9 human ESCs measured by qPCR. GAPDH served as the loading control. Bars indicate the mean ± s.e.m. of $n=3$ independent experiments, shown as circles. Arrowheads indicate samples with statistical significance (unpaired two-tailed Mann–Whitney *U*-test). **e**, The transcriptomes of isogenic niPSCs generated via CCC and microfluidic conditions are highly similar (Fig. 3c) and distinct from primed iPSCs. Naive- and primed-specific genes are indicated. **f**, Effect of channel height on BJ reprogramming. Colonies with compact morphology (green circles) or coexpressing KLF17 and OCT4 (red circles) were quantified at D15. The percentages of double-positive colonies are indicated. Bars indicate the means of technical replicates (circles, $n=8$, 6, 9 and 10 replicates for 100, 200 500 and 1,000 μm, respectively) pooled from 2 independent experiments shown in different shades of colours. See also Supplementary Fig. 7c. **g**, Resetting of primed H9 in both microfluidic conditions and CCC using doxycycline (DOX)-inducible NK2 transgenes in the presence of 2iL[10], or conditions supporting primed pluripotency (FGF+KSR). TFCP2L1-positive colonies scored at D14. Bars indicate the means of technical replicates (circles, $n=20$ in microfluidic conditions and $n=2$ in CCC) pooled from 2 independent experiments, shown in different colours. **h**, Model of human somatic cell reprogramming routes. Scale bars, 100 μm (**b**–**c**,**g**), 1 mm (**b**, tiling). Images from additional repeats for panels **b** and **g** have been deposited in Figshare under DOI: 10.6084/m9.figshare.7151828. See Supplementary Table 5 for all individual numerical values and *P* values.

to naive pluripotency. With our technology, transgene-free niPSCs could be efficiently generated from patients, and their high clonogenic capacity will allow efficient gene editing. Conventional PSC lines display strong differentiation biases[6] and retain some somatic epigenetic features[33]. We provided evidence of prompt differentiation capacity; however, it is still an open question as to whether naive PSCs could represent a better system for the generation of differentiated cell types for therapeutic purposes or for disease modelling. Furthermore, naive PSCs represent a good model system to study early development, as exemplified by a recent study[24] of X-chromosome inactivation dynamics. Future studies comparing the epigenetic profile and differentiation potential of multiple patient-specific niPSCs and conventional primed iPSCs will reveal the full potential of human niPSCs in regenerative medicine, disease modelling and developmental studies.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41556-018-0254-5.

Received: 3 October 2018; Accepted: 15 November 2018;
Published online: 31 December 2018

References

1. Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
2. Thomson, J. A. et al. Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145–1147 (1998).
3. Takahashi, K. et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
4. Hackett, J. A. & Surani, M. A. Regulatory principles of pluripotency: from the ground state up. *Cell Stem Cell* **15**, 416–430 (2014).
5. Davidson, K. C., Mason, E. A. & Pera, M. F. The pluripotent state in mouse and human. *Development* **142**, 3090–3099 (2015).
6. Osafune, K. et al. Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat. Biotechnol.* **26**, 313–315 (2008).
7. Weinberger, L., Ayyash, M., Novershtern, N. & Hanna, J. H. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat. Rev. Mol. Cell Biol.* **17**, 155–169 (2016).
8. Luni, C. et al. High-efficiency cellular reprogramming with microfluidics. *Nat. Methods* **13**, 446–452 (2016).
9. Zhang, J. et al. LIN28 regulates stem cell metabolism and conversion to primed pluripotency. *Cell Stem Cell* **19**, 66–80 (2016).
10. Takashima, Y. et al. Resetting transcription factor control circuitry toward ground-state pluripotency in. *Hum. Cell* **158**, 1254–1269 (2014).
11. 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).
12. Warren, L. et al. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* **7**, 618–630 (2010).
13. Gafni, O. et al. Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282–286 (2013).
14. 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).
15. Yoshida, Y., Takahashi, K., Okita, K., Ichisaka, T. & Yamanaka, S. Hypoxia enhances the generation of induced pluripotent stem cells. *Cell Stem Cell* **5**, 237–241 (2009).
16. Watanabe, K. et al. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat Biotechnol.* **6**, 681–686 (2007).
17. Theunissen, T. W. et al. Molecular criteria for defining the naive human pluripotent state. *Cell Stem Cell* **19**, 502–515 (2016).
18. Guo, G. et al. Naive pluripotent stem cells derived directly from isolated cells of the human inner cell mass. *Stem Cell Rep.* **6**, 437–446 (2016).
19. Liu, X. et al. Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. *Nat. Methods* **14**, 1055–1062 (2017).
20. Kilens, S. et al. Parallel derivation of isogenic human primed and naive induced pluripotent stem cells. *Nat. Commun.* **9**, 360 (2018).
21. Cacchiarelli, D. et al. Integrative analyses of human reprogramming reveal dynamic nature of induced pluripotency. *Cell* **162**, 412–424 (2015).
22. Smith, Z. D. et al. DNA methylation dynamics of the human preimplantation embryo. *Nature* **511**, 611 (2014).
23. Okae, H. et al. Genome-wide analysis of DNA methylation dynamics during early human development. *PLoS Genet.* **10**, e1004868 (2014).
24. Sahakyan, A. et al. Human naive pluripotent stem cells model X chromosome dampening and X inactivation. *Cell Stem Cell* **20**, 87–101 (2017).
25. Carbognin, E., Betto, R. M., Soriano, M. E., Smith, A. G. & Martello, G. Stat3 promotes mitochondrial transcription and oxidative respiration during maintenance and induction of naive pluripotency. *EMBO J.* **35**, 618–634 (2016).
26. Lee, J.-H. et al. Lineage-specific differentiation is influenced by state of human pluripotency. *Cell Rep.* **19**, 20–35 (2017).
27. Warrier, S. et al. Direct comparison of distinct naive pluripotent states in human embryonic stem cells. *Nat. Commun.* **8**, 15055 (2017).
28. Hay, D. C. et al. Efficient differentiation of hepatocytes from human embryonic stem cells exhibiting markers recapitulating liver development in vivo. *Stem Cells* **26**, 894–902 (2008).
29. Errichelli, L. et al. FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons. *Nat. Commun.* **8**, 14741 (2017).
30. Schlaeger, T. M. et al. A comparison of non-integrating reprogramming methods. *Nat. Biotechnol.* **33**, 58–63 (2014).
31. Nakagawa, M. et al. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol.* **26**, 101–106 (2007).
32. Wang, Y. et al. Unique molecular events during reprogramming of human somatic cells to induced pluripotent stem cells (iPSCs) at naïve state. *eLife* **7**, e29518 (2018).
33. Urbach, A., Bar-Nur, O., Daley, G. Q. & Benvenisty, N. Differential modeling of fragile X syndrome by human embryonic stem cells and induced pluripotent stem cells. *Cell Stem Cell* **6**, 407–411 (2010).
34. Blakeley, P. et al. Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Dev. Camb. Engl.* **142**, 3151–3165 (2015).
35. Quintanilla, R. H. Jr, Asprer, J. S. T., Vaz, C., Tanavde, V. & Lakshmipathy, U. CD44 is a negative cell surface marker for pluripotent stem cell identification during human fibroblast reprogramming. *PLoS ONE* **9**, e85419 (2014).

Acknowledgements

The authors thank M. Montagner and S. Dupont for critical reading of the manuscript, and the Martello Laboratory for discussions and suggestions. Thanks are also given to the following: the Smith Laboratory for sharing reset H9 cells and plasmids; Miltenyi Biotec for providing mmRNAs; A. Rosa and R. De Santis for their help with the neuronal differentiation of niPSCs; and A. Manfredi, TIGEM NGS and the Bioinformatics Core for their technical support on library generation and data processing. The authors are indebted to P. Brun for providing primary skin fibroblasts. G.M.'s Laboratory is supported by grants from the Giovanni Armenise–Harvard Foundation, the Telethon Foundation (TCP13013) and an ERC Starting Grant (MetEpiStem). D.C.'s Laboratory is supported by grants from the Giovanni Armenise–Harvard Foundation, the Telethon Foundation, the Rita Levi Montalcini programme from MIUR and an ERC Starting Grant (CellKarma). C.R.'s Laboratory is supported by the Italian Association for Cancer Research (IG17185). M.J.Z. is supported by a BMBF eMed grant (01ZX1504) and the Max Planck Society. N.E.'s Laboratory is supported by grants from the University of Padova (TRANSAC and PRAT), the CaRiPaRo Foundation, the Telethon Foundation (GGP15275), an Oak Foundation Award (W1095/OCAY-14-19) and the NIHR GOSH BRC. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Author contributions

S.G. and M.P. performed reprogramming, isolation, characterization and differentiation of niPSCs. I.Z. helped in establishing conditions for the expansion of naive PSCs and characterizing niPSCs. O.G. helped with the microfluidic experiments. P.M. and C.R. performed the bioinformatic analyses. D.C., M.J.Z. and G.M. designed the RRBS experiment, and M.M. and M.J.Z. analysed the RRBS data. S.G., M.P. and G.M. designed the experiments. N.E. and G.M. supervised the study and wrote the manuscript. S.G. and M.P. edited the manuscript.

Competing interests

S.G., M.P., N.E and G.M. are co-inventors on a patent filing describing the generation of human naive iPSCs from somatic cells. All other authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41556-018-0254-5.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to N.E. or G.M.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Cell culture. Naive human ESCs (reset H9; described previously[10], generated via the transient expression of NK2 and provided by A. Smith's Lab) were cultured on mitotically inactivated mouse embryonic fibroblasts (MEFs; DR4, American Type Culture Collection) in 2iLGo medium prepared as follows: 1% N2 and 2% B27 supplements in DMEM-F12:Neurobasal medium (N2B27, all Thermo scientific) were supplemented with 1 µM PD0325901 (Axon Medchem), 1 µM CHIR99021 (Axon Medchem), 10 ng ml^{-1} human LIF (produced in-house) and 1–2 µM Gö6983 (Axon Medchem). It was critical to titrate the concentration of each batch of Gö6983 to minimize cell stress due to accumulation of the inhibitor and to maximize the expression of naive markers. niPSCs were also generated and cultured on MEFs in the naive-supporting media 4iLA and RSeT. 4iLA medium was prepared as follows: N2B27 supplemented with 1 µM PD0325901, 1 µM WH-4-023 (Axon Medchem), 0.5 µM SB590885 (Axon Medchem), 10 µM ROCK inhibitor, 10 ng ml^{-1} LIF and 20 ng ml^{-1} activin-A. The FGF-free medium RSeT (05970, StemCell Technologies) was prepared following the manufacturer's instructions. niPSCs were passaged every 4–6 days as follows: cells were washed with phosphate buffer without Ca^{2+}/Mg^{2+} (PBS) and incubated with 300 µl TripLE select (Thermo Scientific) per 12-well plate for 3 min at room temperature. N2B27 medium (700 µl) was added to inhibit dissociation. Clusters (3–5 cells) were obtained by pipetting the entire volume twice. Cell suspensions were centrifuged at 300 × *g* for 4 min, resuspended in naive medium with 10 µM of the ROCK inhibitor Y27632 (Axon Medchem) and seeded at ~300 MEFs per mm^2. The ROCK inhibitor was used only for 24 h after passaging. The HFF cell lines BJ (passage 12) and HFF-1 (passage 18) (American Type Culture Collection), the female lung fibroblast cell lines WI-38 and IMR-90 (American Type Culture Collection), and primary skin fibroblasts of an 80-year-old female donor were cultured in DMEM with 10% fetal bovine serum (FBS; Sigma-Aldrich) before reprogramming. Somatic cells were cultured in normoxia (21% O_2, 5% CO_2, 37 °C), whereas pluripotent stem cells were cultured in hypoxia (5% O_2, 5% CO_2, 37 °C); the medium was changed daily. All cell lines were mycoplasma-negative (Mycoalert, Lonza). Low passage and karyotipically normal primed human ESCs (H9; also known as WA09) were obtained from and used under authorization from WiCell Research Institute. Primed human ESCs and iPSCs (HPD00) were cultured on Matrigel-coated (Corning) conventional plates in mTeSR1 medium (StemCell Technologies).

Microfluidic chips production. Microfluidic chips were produced as described at Nature Protocol Exchange (https://doi.org/10.1038/protex.2018.083). The standard height of each channel is 200 µm and the area is 27 mm^2. Each channel holds 10 µl of medium when considering the inlet, the outlet and the culture channel itself (Supplementary Fig. 1a).

Reprogramming. A detailed protocol for the generation of niPSCs via microfluidics has been shared at Nature Protocol Exchange (https://doi.org/10.1038/protex.2018.083). All reprogramming experiments were performed in hypoxia conditions unless stated otherwise. Transfection of mmRNA leads to the transient production of the proteins of interest, with a peak between 12 and 24 h, and complete clearance at 42 h[12]. Thus, to guarantee stable protein expression, we transfected cells daily. Moreover, for analyses of primary colonies by immunofluorescence, quantitative PCR (qPCR) or RNA-seq, we waited at least 3 days since the last round of transfection to allow full clearance of the mmRNAs.

Microfluidic channels were coated with 25 µg ml^{-1} fibronectin (Sigma-Aldrich) for 30 min at room temperature and fibroblasts were seeded at day −1 at 25 cells per mm^2 in DMEM/10% FBS. Of note, a density of 10 cells per mm^2, optimal for the generation of primed iPSCs[8], did not lead to the generation of niPSCs. On day 0, 2 h before the first mmRNA transfection, we applied a medium optimized for mRNA transfection and previously used for the derivation of primed PSCs (PRM, StemMACS ReproBREW XF, Miltenyi Biotec). Where indicated, we added the ROCK inhibitor (5 µM) and CHIR99021 (1 µM) to the PRM. Cells were transfected daily at 9:00, and fresh PRM was given daily at 18:00. From day 6, we used a naive-supporting medium with the same regimen of transfection and fresh medium change. From days 0 to 12, media were supplemented with 200 ng ml^{-1} B18R (eBioscience) to suppress single-strand-RNA-induced immune responses. The transfection mix was prepared according to the StemMACS mRNA transfection kit (Miltenyi Biotec): 100 ng µl^{-1} mmRNA mix of OSKMN and nuclear green fluorescent protein (with stoichiometry 3:1:1:1:1:1) was diluted in transfection buffer by mixing 10 µl of the mmRNA mix with 30 µl of the buffer. Transfection reagent was diluted separately in transfection buffer by mixing 3 µl of the transfection reagent with 37 µl of the transfection buffer. The two solutions were mixed (final volume of 80 µl) and incubated for 20 min. For each microfluidic channel, we diluted 1.2 µl of the transfection solution in 8.8 µl of medium (either PRM or naive medium) and added it to the cells (corresponding to 0.28 ng mm^{-2}). During the first 6 days, the dose of mmRNAs was gradually increased to facilitate expansion of fibroblasts.

For reprogramming in wells (CCC), we used the conditions described for microfluidics but increased the amount of mmRNA. For each well of a 12-well plate, we used 46.3 µl of transfection solution in 0.8 ml of medium, which corresponds to 1.5 ng mm^{-2}.

Reprogramming of fibroblasts towards the primed state via microfluidics was achieved using the protocol described above, but using PRM until day 12. To allow direct comparison of the reprogramming trajectories towards primed and naive pluripotency (Fig. 3d–f), we used the same seeding density of 25 cells per mm^2 and the same OSKMN mmRNA cocktail.

For testing the effect of different microfluidic channel heights (Fig. 8f), the same mmRNA amount per cell was delivered daily, independently of the channel heights.

Resetting of primed H9 cells to naive was performed as described previously[10]. In brief, primed H9 cells were transfected with doxycycline-inducible NK2 piggyBac constructs on MEF feeders in DMEM-F12 supplemented with 20% KSR (Gibco) and 10 ng ml^{-1} bFGF (produced in-house). After selection, cells were dissociated into single cells with TrypLE and replated on day −2 in the presence of the ROCK inhibitor for 24 h in microfluidic conditions and CCC at a density of 4 cells per mm^2. Medium was changed daily for CCC and twice a day for microfluidic conditions. On day −1, cells were incubated with doxycycline (1 µg ml^{-1}; Thermo Scientific). From day 0, the medium was switched to 2iL where indicated. On day 14, cells were fixed and stained with an anti-TFCP2L1 antibody.

In vitro differentiation. For germ layer differentiation, niPSCs were seeded on 1% Matrigel-coated coverslips with 100 MEFs per mm^2 and cultured in RSeT medium for 2 days. Germ layer-specific medium was used thereafter, with daily medium changes. Ectoderm medium was used for 12 days, mesoderm medium was used for 6 days and endoderm medium was used for 6 days. Ectoderm medium consisted of knockout DMEM (Thermo Scientific), 15% KSR, 1% non-essential amino acids (NEAA), 1% L-glutamine, 0.1 µM LDN193189 (Miltenyi Biotec) and 20 ng ml^{-1} hFGF2 (PeproTech); after the first 3 days, hFGF2 was replaced with 10 µM SB431542 (Miltenyi Biotec) for 9 days. Mesoderm medium consisted of RPMI with 2% B27 (Thermo Scientific), 20 ng ml^{-1} hFGF2, 50 ng ml^{-1} hBMP4 (R&D) and 3 µM CHIR99021 (only for the first 2 days). Endoderm medium consisted of RPMI with 2% B27, 100 ng ml^{-1} activin-A (PeproTech) and 3 µM CHIR99021 (only for the first 2 days). For embryoid body culture, niPSC colonies were mechanically scratched with a tip and transferred into ultra-low adhesive wells (Corning) in the presence of DMEM, 20% FBS, 200 mM L-glutamine, 1% NEAA and 0.1 mM 2-mercaptoethanol[8]. Medium was changed every other day for 15 days before plating the embryoid bodies onto 1% Matrigel-coated glass plates (Labtek). After 5 days, adherent and spread cells were fixed for immunostaining.

Neuronal differentiation. The neuronal differentiation protocol was adapted from a previously published protocol[29]. Naive iPSCs were seeded as single cells at high density (530 cells per mm^2) on Matrigel-coated plates. Cells were cultured for 2 days in RSeT. On day 0, cells were cultured in N2B27, 1% NEAA, 200 ng ml^{-1} L-ascorbic acid (neural medium) supplemented with 20 ng ml^{-1} bFGF and 0.1 µM LDN193189. Medium was refreshed daily. On day 3, medium was switched to neural medium supplemented with 0.1 µM LDN193189 and 10 µM SB431542. On days 4–9, the same medium was supplemented with 1 µM all-*trans* retinoic acid (Sigma-Aldrich) and 1 µM SAG (Calbiochem) and was refreshed daily. On days 10–15, medium was switched to neural medium supplemented with 5 µM DAPT (Sigma-Aldrich) and 4 µM SU-5402 (Sigma-Aldrich), 1 µM all-*trans* retinoic acid and 1 µM SAG (Smoothened Agonist, Calbiochem), and was refreshed daily. On day 16, cells were dissociated with TrypLE for 10 min, and 2 volumes of neural medium was added to inhibit dissociation. Cells were seeded on Matrigel-coated glass coverslips in wells and cultured in maturation medium, which is based on neural medium supplemented with 20 ng ml^{-1} BDNF, 10 ng ml^{-1} GDNF, 10 ng ml^{-1} CNTF (PeproTech) and 10 µM of the ROCK inhibitor for the first 24 h. Fresh maturation medium was provided daily up to day 22 when cells were fixed for immunostaining.

Hepatic differentiation. The hepatic differentiation protocol was adapted from a previously published protocol[28]. Briefly, niPSCs were seeded as single cells (20 cells per mm^2) on Matrigel-coated plates with sparse MEF coverage (50 cells per mm^2). Cells were cultured for 2 days in RSeT. On day 0, cells were cultured in RPMI, 2% B27 and 3 µM CHIR99021. The same medium with the addition of 100 ng ml^{-1} activin-A was refreshed on days 1–2. On days 3–8, the medium was switched to knockout DMEM with 20% KSR, 2 mM L-glutamine, 1% NEAA, 0.1 mM 2-mercaptoethanol and 1% dimethylsulfoxide (DMSO), and changed every other day. On days 9–15, cells were cultured in maturation medium, which is based on L15 basal medium, 8% FBS, 8% tryptose phosphate broth (Sigma), 10 µM hydrocortisone (Sigma), 1 µM insulin, 2 mM L-glutamine, 10 ng ml^{-1} HGF (PeproTech) and 20 ng ml^{-1} oncostatin-M (PeproTech). Fresh maturation medium was provided every other day.

Immunofluorescence and stainings. Immunofluorescence analysis was performed either on 1% Matrigel-coated glass coverslips in wells or in situ in microfluidic channels with the same protocol. Cells were fixed in 4% (w/v) paraformaldehyde (Sigma-Aldrich) in PBS for 10 min and blocked in 5% horse serum with 0.3% (v/v) Triton-X-100 (Sigma-Aldrich) for 1 h. Blocking buffer:PBS (1:2) was used to dilute primary antibodies (Supplementary Table 3). For staining for 5mC, fixed samples were permeabilized with 0.5% Triton X-100 for 1 h and treated with 2 N HCl for 30 min at room temperature to denature DNA. Samples were neutralized with PBS before blocking and antibody incubation. Secondary antibodies (Supplementary

Table 3) were incubated at room temperature for 45 min at 1:500. Nuclei were stained with either Hoechst 33342 (Thermo Scientific) or 4,6-diamidino-2-phenylindole (DAPI; Sigma-Aldrich). For mitochondrial staining, cells were incubated with 1:50,000 TMRM and 1:20,000 MitoTracker (Thermo Scientific, T668, M7514) for 30 min in culture medium at 37 °C and washed twice in PBS before image acquisition. For EdU staining, cells were exposed to an EdU pulse of 1 h before fixation in formaldehyde for 15 min. Samples were processed using a Click-iT EdU Alexa Fluor 488 Imaging kit and counterstained with Hoechst nuclear dye (all Thermo Scientific). Fluorescence images were acquired using a Leica SP5 II confocal system or a Leica 6000B epifluorescence microscope. For alkaline phosphatase staining, cells were fixed with a citrate–acetone–formaldehyde solution and stained using an alkaline phosphatase kit (Sigma-Aldrich). Plates were scanned using an Epson scanner and scored manually.

Flow cytometry and cell sorting. Single cells in suspension were obtained by incubating samples for 5 min with TrypLE. For DNA content analysis by flow cytometry, cell pellets fixed in cold 70% ethanol were resuspended in PBS, incubated with 50 μg ml^{-1} propidium iodide, and processed using a FACSCanto II (BD). For cell sorting based on DNA content, cells were resuspended in naive medium with 10 μM of the ROCK inhibitor and 5 μM Vybrant DyeCycle Ruby live stain (Thermo Scientific) for 30 min before sorting. Sorted samples were cultured on new MEFs with 10 μM of the ROCK inhibitor for 24 h. See Supplementary Fig. 3c for the gating strategy.

Karyotyping. Cells were incubated with 0.06 μg ml^{-1} KaryoMAX (Thermo Scientific) in culture medium for 6 h at 37 °C. niPSCs were isolated with ReLeSR (StemCell Technologies) and resuspended in PBS before centrifugation. The pellet was resuspended in pre-warmed 75 mM KCl for 10 min at 37 °C. After centrifugation, the pellet was gently resuspended in 1 ml of freshly prepared fixative (3:1 methanol:acetic acid). This step was repeated twice. Q-banded fixed samples were analysed by Research & Innovation.

qPCR. Total RNA was isolated using a Total RNA Purification kit (Norgen Biotek), and complementary DNA was made from 500 ng using M-MLV Reverse Transcriptase (Invitrogen) and dN6 primers. For real-time PCR, SYBR Green Master mix (Bioline, BIO-94020) was used. Primers are detailed in Supplementary Table 4. Three technical replicates were performed for all qPCR experiments. GAPDH was used as the endogenous control to normalize expression.

RNA-seq. niPSCs were isolated from MEFs using ReLeSR. Briefly, cells were washed with PBS and then incubated with ReLeSR for 60 s. ReLeSR was then removed and cells were left with a thin film of liquid for 7 min. N2B27 medium was added and pipetted 4–5 times to detach colonies. Total RNA was isolated as above and sequenced using an Illumina NexSeq500, in 75-base pair pair-end format. Libraries were prepared following standard protocols from Illumina using a TruSeq Stranded mRNA Library kit. Expression levels for all the genes from ENSEMBL 87 were quantified using RSEM v.1.3.0[36] with STAR v.2.5.2b[37] (human genome GRCh38.p7). We built the genome index for STAR alignment using "rsem-prepare-reference" with options '–star-sjdboverhang' set to mean read length minus 1 according to STAR guidelines. Alignment and quantification were performed with default parameters using the stranded pair-end mode. The gene expression level quantification of 21 samples collected from publicly available datasets was performed as described for the in-house samples. The RSEM parameters were set according to the library design (mean reads length, stranded/non-stranded, paired/single end; see Supplementary Table 2). We generated the final expression matrix excluding the genes that had less than 10 raw counts in at least 3 out of 34 samples. After applying this filter, we obtained the expression of 20,936 genes.

All RNA-seq statistical analyses were carried out in R environment (v.3.4.3) with Bioconductor 3.6. We computed the differentially expressed genes (DEGs) among the three groups (primed, naive, fibroblast) using edgeR[38] (function call "fit = glmQLFit(counts, design=~group); glmTreat(fit, coef=n, lfc=2)" where 'n' is one of the contrasts). A gene was considered a DEG when the absolute value of $\log_2$-fold change was >2 and the adjusted *P* value was <0.05 (p.adjust function, Benjamini–Hochberg method). Principal component analysis (PCA) was performed using $\log_2$-normalized pseudocounts (defined as count plus 1) with the prcomp function and default parameters using DEGs. Heatmaps were made using the $\log_2$-normalized pseudocounts (unless stated otherwise) with the pheatmap function from pheatmap R package (v.1.0.8, distance used 'euclidean', 'ward' linkage, scale='row') on DEGs or selected markers. The raw counts were normalized using the betweenLaneNormalization function with the upper quantile method (EDAseq R package)[39].

For the comparison between our RNA-seq dataset and the DGE-seq dataset[20] presented in Supplementary Fig. 4c, we first obtained expression values in UMI per million from Supplementary Data 1 of the other study[20], while our dataset was transformed into transcripts per million. Both datasets were quantile-normalized and *z*-scored by row separately, as described in the other study[20]. We merged the two datasets via the gene identification number to obtain a single dataset containing 14,944 genes. PCA was computed using the 4,606 genes obtained from intersecting the merged dataset, and the DEGs used for the RNA-seq comparison in the PCA are shown in Fig. 3c.

Transposon analysis. Transposon coordinates were downloaded from UCSC repeat masker track (hg38). We filtered out those transposons that overlap any genomic features annotated in ENSMBL 87. We aligned the raw reads using bowtie2 sensitive and end-to-end mode on the human genome GRCh38.p7. Transposon expression was quantified using bamtools multicov (v.2.26.0).

We analysed transposons with at least 20 counts per million in at least one sample. Differentially expressed transposons were identified using the edgeR R package. A transposon was considered differentially expressed when the *P* value was ≤0.05 and the $\log_2$-fold change was ≥2. Heatmaps were made using the $\log_2$-normalized pseudocounts scaled by row means with the pheatmap R package as describe for the RNA-seq data.

SNP analysis for imprinting and X reactivation. From the alignments, we extracted the reads with a minimum quality of 30 (samtools view –q 30 align.bam align-f.bam). Using GATK haplotypeCaller (genotyping_mode: DISCOVERY, minReadsPerAlignmentStart: 5, max_alternate_alleles: 1, stand_conf_call: 1 and filter_reads_with_N_cigar), we generated allele counts over the SNPs of dbSNP human v.b149. Non-biallelic variants as well as insertions and deletions were filtered out. Moreover, we considered only SNPs with at least ten reads in at least one sample. A SNP was considered heterozygous when the ratio of the counts between minor and major alleles was greater than 0.2 and the minor allele had at least 5 reads. A minor allele is the allele with fewer counts.

The loss of imprinting of a gene was quantified as the number of SNPs that showed heterozygosity in each gene. We called a loss of imprinting when at least two SNPs showed heterozygosity in a gene. X reactivation was quantified by selecting the heterozygous SNP from X-chromosome genes and by plotting the minor/total allele count ratios of heterozygous SNPs (defined as above). We excluded the pseudo-autosomal regions (PAR1: chrX:10,000–2,781,479 and chrY:10,000–2,781,479; PAR2: chrX:155,701,383–156,030,895 and chrY:56,887,902–57,217,415) from the analysis.

Methylation analysis. RRBS libraries were produced using an Ovation RRBS Methyl-Seq System (NuGEN) starting with 100 ng of genomic DNA extracted with a Quick-DNA Plus kit (Zymo) according to the manufacturer's specifications. Libraries were sequenced on a NextSeq 500 (Illumina) using a single-end 75-cycle high-output flow cell. Sequence reads were first trimmed using Trim Galore software (v.0.4.1, Babraham Bioinformatics, available at www.bioinformatics.babraham.ac.uk/projects/trim_galore/) to remove adapter sequences and low-quality end bases and then trimmed with a custom python script provided by NuGEN Technical Support (v.1.11) to remove any read that did not contain an MspI site signature (YGG) at the 5′-end. Alignment of the reads onto the hg19 reference sequence and methylation calling were performed using Bismark[40](v.0.18.1). Coordinates of differentially methylated regions were obtained from the literature[21,23]. Subsequently, average DNA methylation levels and total coverage for each differentially methylated region were determined for all Bismarck-processed RRBS data files using R (v.3.3.3) and the methylkit package (v.0.9.5)[41]. To that end, RRBS data files were processed with the methRead function and the bismarkCoverage parameter set. Next, the regionCounts function was used to determine the number of methylated and unmethylated C residues in each of the differentially methylated regions in each sample. These values were then combined using the unite() function with the min.per.group paramater set to 1. Finally, we only retained those regions that were covered by at least five reads and computed a coverage weighted average methylation level for each region across all CpG islands that were covered. These values are plotted in Fig. 4d–f and listed in Supplemenary Table 5 along with the genomic coordinates of each genomic region on hg19.

Image analyses. Fiji 1.0 (ImageJ2) was used for image analyses. The size of niPSC colonies (Fig. 2d) was measured by delimiting a colony area and calculating the equivalent diameter as $(\text{area}/\pi)^{0.5}$ (using the Shape filter plugin function). Fluorescence intensity (H3K9me3 or 5mC) across nuclei (Fig. 4a,b; Supplementary Fig. 5a) was measured using the Plot Profile function. For each cell line, we selected representative fields and six nuclei were randomly chosen as described previously[10].

Statistics and reproducibility. For statistical analyses, normality distribution was tested using D'Agostino–Pearson normality tests in Prism (Graphpad). The non-parametric Kruskal–Wallis test was used with non-normal datasets, using Dunn's correction for multiple comparisons. For datasets with $n = 3$ normality distribution and for which variance could not be estimated, we used the non-parametric two-tailed Wilcoxon Mann–Whitney *U*-test, for which the asymptotic *P* values were calculated in R to avoid continuity correction. Tests and exact *P* values are reported in Supplementary Table 5. The numbers of replicates and independent experiments are indicated in the figure legends and in Supplementary Table 5, which reports the numerical values for all technical replicates of each dataset in this study. All error bars indicate the standard error of the mean (s.e.m.). All qPCR experiments were repeated at least two times independently, with three technical replicates for each experiment. All key experiments were repeated independently using different cells lines (for example, reprogramming of five different somatic cell types, molecular

characterization of five niPSC lines and differentiation of four different niPSC lines). For reprogramming experiments under CCC, we performed at least two independent experiments, while for microfluidic experiments, we performed up to eight independent experiments, with 5–60 technical replicates. All differentiation experiments were performed in at least two biologically independent experiments, using between two and four different niPSCs lines. For RNA-seq and RRBS, we profiled at least five niPSC lines together with relevant controls.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

RNA-seq and RRBS data of this study have been deposited in the Sequence Read Archive (SRA) under BioProject number PRJNA381757 and GEO under accession code GSE110377. Accession numbers of other published datasets are reported in each figure plotting RNA-seq data and in Supplementary Table 2. Source data of all repeats of all experiments are provided in Supplementary Table 5. For figure panels showing representative images of morphologies or immunostainings, images from additional independent repeats are available at Figshare (https://doi.org/10.6084/m9.figshare.c.4250195). All other data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

36. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
37. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* **29**, 15–21 (2013).
38. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma. Oxf. Engl.* **26**, 139–140 (2010).
39. Risso, D., Schwartz, K., Sherlock, G. & Dudoit, S. GC-content normalization for RNA-Seq data. *BMC Bioinformatics* **12**, 480 (2011).
40. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinforma. Oxf. Engl.* **27**, 1571–1572 (2011).
41. Akalin, A. et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* **13**, R87 (2012).

nature research

Corresponding author(s): Graziano Martello

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	For qPCR data collection the Applied Biosystems™ QuantStudio™ 6 Flex Real-Time PCR System version 1.3 was used. For immunostaining data collection the Leica TCS SP5 LAS AF version 2.7.3.9723 was used, and Leica 6000B LAS AF version 2.7 was used for tiling of entire microfluidics channels. For flow citometry and cell sorting data collection the BD FACSCanto™ II FACSDiva version 6.1.3 and the Bio-Rad S3e™ Cell Sorter ProSort™ version 1.4.0.20 were used.
Data analysis	For statistical analysis we used Prism 6 (Graphpad, version 6.05). Fiji 1.0 (ImageJ2) was used for image analysis. For RNA-Seq data analysis, expression levels for all the genes from ENSEMBL 87 were quantified using RSEM 1.3.0 (RNA-Seq by Expectation-Maximization) with STAR 2.5.2b (RNA-Seq aligner). All RNA-seq statistical analyses were carried out in R environment (version 3.4.3) with Bioconductor 3.6. We computed the differentially expressed genes (DEG) among the 3 groups (primed, naïve, fibroblast) using edgeR. Principal component analysis (PCA) was performed using log2-normalized pseudo-counts (defined as count plus 1) with prcomp function with default parameters using DEGs. Heatmaps were made using the log2-normalized pseudo-counts (unless stated otherwise) with pheatmap function from pheatmap R package (version 1.0.8), distance used 'euclidean', 'ward' linkage, scale='row') on DEGs or selected markers. The raw counts were normalized using betweenLaneNormalization function with upper quantile method (EDAseq R package). We provide in Supplementary Table 5 (Statistic Source Data) all the values used to build heatmaps and PCAs.

SNP analysis was performed using GATK haplotypeCaller version 3.8-0-ge9d806836 (genotyping_mode: DISCOVERY, minReadsPerAlignmentStart: 5, max_alternate_alleles: 1, stand_conf_call: 1 and filter_reads_with_N_cigar).

Libraries generated in RRBS were analysed as follow: sequence reads were first trimmed using Trim Galore software (Version 0.4.1) to remove adapter sequences and low-quality end bases and then trimmed with a custom python script provided by NuGEN Technical Support (Version 1.11) to remove any read that does not contain an MspI site signature (YGG) at the 5'-end. Reads alignment on hg19 reference sequence and methylation calling was then performed with Bismark (Version 0.18.1_dev). Coordinates of DMRs were obtained from literature. Subsequently, average DNA methylation levels and total coverage for each DMR region was determined for all Bismarck processed RRBS data files using R (Version 3.3.3) and the methylkit package (version 0.9.5). To that end, RRBS data files were processed with the methRead function and the bismarkCoverage parameter set. Next, the regionCounts function was used to determine the number of methylated and unmethylated C's in each of the DMRs in each sample. These values were then combined using the unite() function with the min.per.group paramater set to 1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All RNA-seq and RRBS data generated in this study can be found in SRA Accession number: PRJNA381757
Repository/DataBank Accession: GEO
Accession ID: GSE110377
Databank URL: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE110377
Additional images from indepedent repeats of morphological analyses and immunostainings on Figshare Collection "Generation of human naive induced pluripotent stem cells in microfluidics" (10.6084/m9.figshare.c.4250195).

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size calculation was not performed. Sample size was chosen based on standards in the field and is indicated in figure legends and in Supplementary Table 5, which reports all source data for each dataset in this study. All qPCR experiments have been repeated at least two times independently, with 3 technical replicates for each experiment. All key experiments were repeated independently using different cells lines (e.g. reprogramming of 5 different somatic cell types, molecular characterization of 5 niPSC lines and differentiation of 4 different niPSC lines). For reprogramming experiments under conventional condition we performed at least two independent experiments, while for microfluidic experiments we performed up to 8 independent experiments, with 5 to 60 technical replicates. All differentiation experiments were performed at least in two biologically independent experiments, using between 2 and 4 different niPSCs lines. For RNAseq and RRBS we profiled at least 5 niPSC lines together with relevant controls.
Data exclusions	We presented all data obtained, also in the case of negative results.
Replication	The main findings, such as the reprogramming of somatic cells to naive pluripotency, have been reproduced with n>400. For qPCR at least 2 independent experiments have been performed and confirmed by independent techniques (e.g. RNAseq in Fig. 3a, immunostaining in Fig. 2g-h and Supplementary Fig. 2 and qPCR for pluripotency markers in Fig. 2e-f). For RNA-seq and RRBS we analysed 5 independently generated niPSC lines, together with relevant controls.
Randomization	Randomization was not relevant. All cell lines or biological samples were analysed or treated in the same manner.
Blinding	Blinding was not applicable to reprogramming and differentiation assays, but comparable results were obtained by two distinct operators in two different laboratories. We applied blinding for RNA-seq processing and data analysis.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used: Please see Supplementary Table 3 for the list of antibodies, product codes, dilutions used for immunostaining and validation from literature.

Validation: Antibodies were used following manufacturers guidelines and were validated in our laboratory by using negative or positive controls. For instance when appropriate naive, primed or differentiated pluripotent stem cells were used. Finally, the antibodies selected for this work have been previously used in literature.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s): MEF DR4, BJ and HFF-1 human foreskin fibroblasts, WI-38 and IMR-90 female lung fibroblasts were obtained from ATCC.
H9 (WA09) human ES cells were obtained from WiCell.
The Reset H9 described in Takashima were kindly provided by Austin Smith's laboratory.
Primary Human Dermal Fibroblasts were kindly provided by Professor Vindigni at University of Padua.

Authentication: H9 human ES cells were authenticated by STR profiling by WiCell.
MEF DR4, BJ, HFF-1, WI-38 and IMR-90 were authenticated by STR profiling by ATCC.
Reset H9 cells were authenticated in this study by SNP profiling together with a panel of niPSCs and primed iPSCs generated in this study (HPD00, HPD01, HPD03, HPD04, HPD06, HPD08).

Mycoplasma contamination: Cells were routinely tested for Mycoplasma contamination. All cell lines were Mycoplasma negative.

Commonly misidentified lines (See ICLAC register): The cell lines used are not listed in ICLAC.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation: Single cells in suspension were obtained by incubating samples for 5 min with TrypLE. For DNA-content analysis by flow cytometry, cells pellets fixed in cold 70% ethanol were resuspended in PBS, incubated with 50 µg ml-1 propidium iodide, and processed with FACSCanto II (BD). For cell sorting based on DNA-content, cells were resuspended in naïve medium with 10 µM Ri and 5 µM Vybrant DyeCycle Ruby live stain (Thermo Scientific) for 30 min before sorting and sorted at S3e sorter (Bio-rad). Sorted samples were cultured on new MEFs with 10 µM Ri for 24 h.

Instrument: FACSCanto II (BD) and S3e sorter (Bio-rad)

Software: FACSDiva version 6.1.3 and ProSort™ version 1.4.0.20

Cell population abundance: We sorted 2000 diploid HPD01 cells in Supplementary Fig. 3c. Reanalysis after 6 passages after Propidium Iodide staining indicated a pure diploid population.

Gating strategy: We stained diploid primed H9 hES cells (WiCell) with Vybrant DyeCycle Ruby, in order to set the gate on the G1 fraction. Such

Gating strategy

gate was used to sort diploid HPD01 cells from a mixed population of diploid and tetraploid cells.
See Supplementary Fig. 3c bottom panel for exemplification of the gating strategy.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.