

ARTICLE

Received 26 Mar 2016 | Accepted 23 Feb 2017 | Published 21 Apr 2017

DOI: 10.1038/ncomms15055

OPEN

# Direct comparison of distinct naive pluripotent states in human embryonic stem cells

S. Warri<sup>1</sup>, M. Van der Jeught<sup>1</sup>, G. Duggal<sup>1,2</sup>, L. Tilleman<sup>3</sup>, E. Sutherland<sup>1</sup>, J. Taelman<sup>1</sup>, M. Popovic<sup>1</sup>, S. Lierman<sup>1</sup>, S. Chuva De Sousa Lopes<sup>1,4</sup>, A. Van Soom<sup>5</sup>, L. Peelman<sup>6</sup>, F. Van Nieuwerburgh<sup>3</sup>, D.I.M. De Coninck<sup>3</sup>, B. Menten<sup>7</sup>, P. Mestdagh<sup>7</sup>, J. Van de Sompele<sup>7</sup>, D. Deforce<sup>3</sup>, P. De Sutter<sup>1</sup> & B. Heindryckx<sup>1</sup>

Until recently, human embryonic stem cells (hESCs) were shown to exist in a state of primed pluripotency, while mouse embryonic stem cells (mESCs) display a naive or primed pluripotent state. Here we show the rapid conversion of in-house-derived primed hESCs on mouse embryonic feeder layer (MEF) to a naive state within 5–6 days in naive conversion media (NCM-MEF), 6–10 days in naive human stem cell media (NHSM-MEF) and 14–20 days using the reverse-toggle protocol (RT-MEF). We further observe enhanced unbiased lineage-specific differentiation potential of naive hESCs converted in NCM-MEF, however, all naive hESCs fail to differentiate towards functional cell types. RNA-seq analysis reveals a divergent role of PI3K/AKT/mTORC signalling, specifically of the mTORC2 subunit, in the different naive hESCs. Overall, we demonstrate a direct evaluation of several naive culture conditions performed in the same laboratory, thereby contributing to an unbiased, more in-depth understanding of different naive hESCs.

<sup>1</sup>Department for Reproductive Medicine, Ghent-Fertility and Stem cell Team (G-FaST), Ghent University Hospital, 9000 Ghent, Belgium.

<sup>2</sup>Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DS, UK. <sup>3</sup>Department of Pharmaceutics, Laboratory for Pharmaceutical Biotechnology, Faculty of Pharmaceutical Sciences, Ghent University, 9000 Ghent, Belgium. <sup>4</sup>Department of Anatomy and Embryology, Leiden University Medical Center, 2333 ZA Leiden, The Netherlands. <sup>5</sup>Department of Reproduction, Obstetrics and Herd Health, Faculty of Veterinary Medicine, Ghent University, 9820 Merelbeke, Belgium. <sup>6</sup>Department of Nutrition, Genetics and Ethology, Faculty of Veterinary Medicine, Ghent University, 9820 Merelbeke, Belgium. <sup>7</sup>Department of Pediatrics and Medical Genetics, Center for Medical Genetics, Ghent University Hospital, 9000 Ghent, Belgium. Correspondence and requests for materials should be addressed to B.H. (email: bjorn.heindryckx@ugent.be).

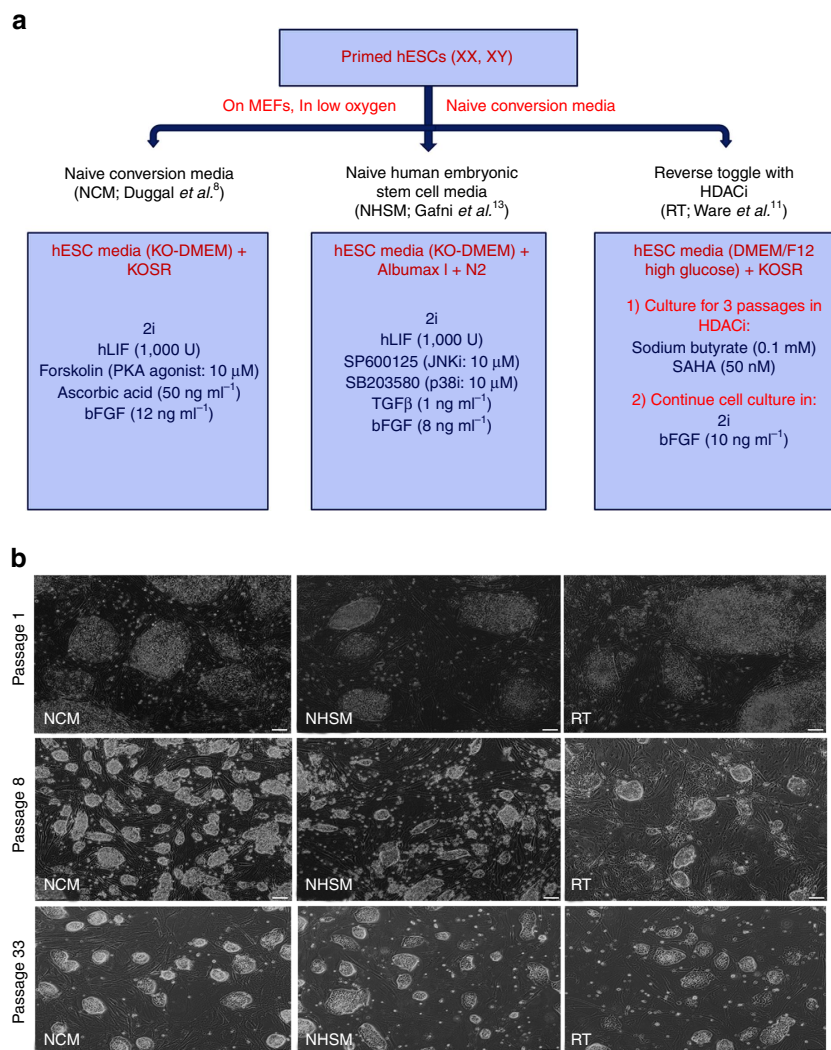
Although both originate from the inner cell mass of pre-implantation blastocysts<sup>1,2</sup>, mouse embryonic stem cells (mESCs) and human ESCs (hESCs) exhibit distinctive characteristics. Human ESCs display a primed state of pluripotency, similar to mouse epiblast stem cells (EpiSCs)<sup>3,4</sup> derived from the post-implantation epiblast. Primed hESCs display flat colony morphology, low single-cell clonogenicity, dependence on TGF $\beta$ /activin/nodal signalling and show inefficiency to contribute to chimeras. Conversely, mESCs reside in the so-called naive pluripotent state, characterized by domed colonies, increased single-cell survival, dependence on JAK/STAT signalling and efficient contribution to chimeras<sup>5–7</sup>. Their high single-cell clonogenicity facilitates bulk culture, deeming them more favourable for future practical applications. Moreover, naive mESCs are also more homogeneous, leading to unbiased and efficient directed differentiation towards germ layer derivatives<sup>8,9</sup>. Recently, several groups have formulated protocols to induce naive pluripotency in hESC, both from pre-existing primed hESCs and by direct derivation from the blastocyst stage. This has been achieved using naive human stem cell media (NHSM)<sup>10</sup>, the reverse-toggle (RT) protocol<sup>11</sup>, 5i/L/F/A medium<sup>12</sup> or by ‘resetting’ primed hESCs, via ectopic *NANOG* and *KLF2* expression, followed by exposure to naive culture medium without bFGF<sup>13</sup>. Likewise, our group recently formulated a novel naive conversion medium (NCM) that facilitates the induction of naive pluripotency from primed hESCs, as well as the derivation of naive mESCs from the blastocyst stage<sup>8</sup>. To date, efforts to compare naive hESC induced via different culture conditions have primarily been performed using published data sets from several groups, which may introduce bias due to the use of different stem cell lines in different laboratories and different transcriptomic platforms employed. Takashima *et al.*<sup>13</sup> hybridized their reset cells on a similar microarray platform as NHSM-naive hESCs and compared their biological results with published data sets for NHSM and 3iL-naive hESCs. Lineage-specific markers, such as *AFP* and *EOMES* were expressed in NHSM and 3iL-naive hESCs, while chromatin modifiers were downregulated compared to reset cells<sup>13</sup>. Whether these findings are truly biologically representative is speculative as only the reset cells were induced in culture, while the analysis of NHSM and 3iL-naive hESCs exclusively relied on published data sets from other studies. In addition, several differences were observed between the different lines converted in NHSM conditions<sup>13</sup>. Further, detailed comparative computational analysis of naive and primed states<sup>14</sup> also suggested that the resulting differentially expressed genes are taken as a proxy of the target protein, which may lead to false positive and false negative results<sup>14</sup>. Therefore, to achieve more comprehensive and reliable biological conclusions, performing experiments in parallel within the same laboratory and on identical hESC lines, rather than solely relying on data set analysis, is imperative. An alternative study performed weighted gene co-expression network analysis on various naive data sets<sup>15</sup> and suggested a potential overlap between reset cells and 5iL/F/A-naive hESCs with minor overlap with primed hESCs. However, the authors stress for caution in data interpretation, as the variations in culture conditions result in transcriptomic noise. Therefore, in order to accurately evaluate biological differences, culture conditions require standardization, while the transcriptomic sample numbers need to be increased<sup>15</sup>. Hence, we report, a detailed comparative analysis of naive hESCs converted in distinct naive media conditions in the same laboratory to provide a more unbiased approach to study the underlying variances between these differentially converted naive hESCs. Since all naive culture conditions to convert primed state hESCs towards a naive pluripotent state used mouse embryonic

fibroblast (MEFs) as feeder layer, we have denoted the three conditions as NCM-MEF, NHSM-MEF and RT-MEF below. We report efficient conversion of three in-house-derived primed hESCs towards a naive pluripotent state, specifically in NCM-MEF and NHSM-MEF media. We also demonstrate that the naive hESCs exhibit dependence on mTORC2 subunit of PI3K/AKT/mTOR pathway, are unique transcriptomically to their parental primed counterparts and can differentiate efficiently towards lineage-specific progenitors.

## Results

**Enhanced naive state displayed by NCM/NHSM-MEF naive hESCs.** To study the characteristics and properties of differentially induced naive hESCs in a controlled setting, we converted existing, in-house-derived three primed hESC lines, within the same laboratory, in different culture conditions. These included NHSM, RT and the NCM protocol. We specifically selected these protocols, as they were shown to generate karyotypically normal naive hESCs, without the need for gene transfection. We also selected conversion on MEF feeder layer considering only NHSM culture medium is thus far capable of supporting feeder-free culture of naive hESC. Hence, to avoid bias, we utilized standard naive culture conditions on MEF feeder layer for all three naive media for downstream comparative analysis, similar to primed hESC culture conditions<sup>10</sup>. Although, secondary optimization also revealed PKCi and ROCKi as optional boosters for NHSM-naive hESCs condition<sup>10</sup> and ROCKi for NCM<sup>8</sup>, we only selected the initial core components of each media for the comparative analysis on MEF feeder layer. The naive hESC lines induced in 5i/L/F/A medium were shown to be karyotypically abnormal and exhibited irreversible loss of imprinting<sup>12,16</sup>, which is concerning as their inability of methylation memory inheritance can influence differentiation capacity for future therapeutic applications. As well, the ‘reset cells’ require ectopic expression of *NANOG* and *KLF2*, prior to culture in naive culture medium, and thus, these naive culture conditions were excluded from this study<sup>13</sup>.

Upon exposing the in-house-derived primed hESC lines in NCM-MEF, NHSM-MEF and RT-MEF conditions (Fig. 1a), we observed successful conversion towards a more naive state in all three conditions, with the first evidence depicted by the appearance of domed-like colony morphology (Fig. 1b). These mouse-like naive colonies appeared within 5–6 days in NCM-MEF, 6–10 days in NHSM-MEF and 14–20 days in RT-MEF naive media. NCM-MEF cells manifested a faster doubling time (12–18 h) compared to NHSM-MEF (14–19 h) and RT-MEF (22–24 h; Fig. 2a, Supplementary Figs 1 and 2). Similarly, NCM-MEF and NHSM-MEF naive hESC lines demonstrated high single-cell survival (95%), whereas RT-MEF naive hESC lines showed a lower efficiency of 77%. Nonetheless, the doubling time and single-cell clonogenicity of the cells converted in all naive media continued to be significantly variable ( $P < 0.05$ ) compared to their primed counterparts, with primed hESC lines exhibiting increased doubling time of 47–52 h and decreased single-cell clonogenicity of only 25% (Fig. 2a, Supplementary Figs 1 and 2). Next, all primed hESCs and their NCM-MEF naive counterparts exhibited a normal chromosomal profile as verified by array comparative genome hybridization, however, in NHSM-MEF and RT-MEF naive hESC culture conditions, only UGent11-70 and UGent11-60/UGent11-70 naive hESCs were karyotypically normal, respectively (Fig. 2b, Supplementary Figs 1 and 2). Converted naive hESCs were positive for OCT4/*NANOG* (Fig. 2c, Supplementary Figs 1 and 2) and spontaneously differentiated into all the three germ layers (Fig. 2d, Supplementary Figs 1 and 2). Interestingly, embryoid



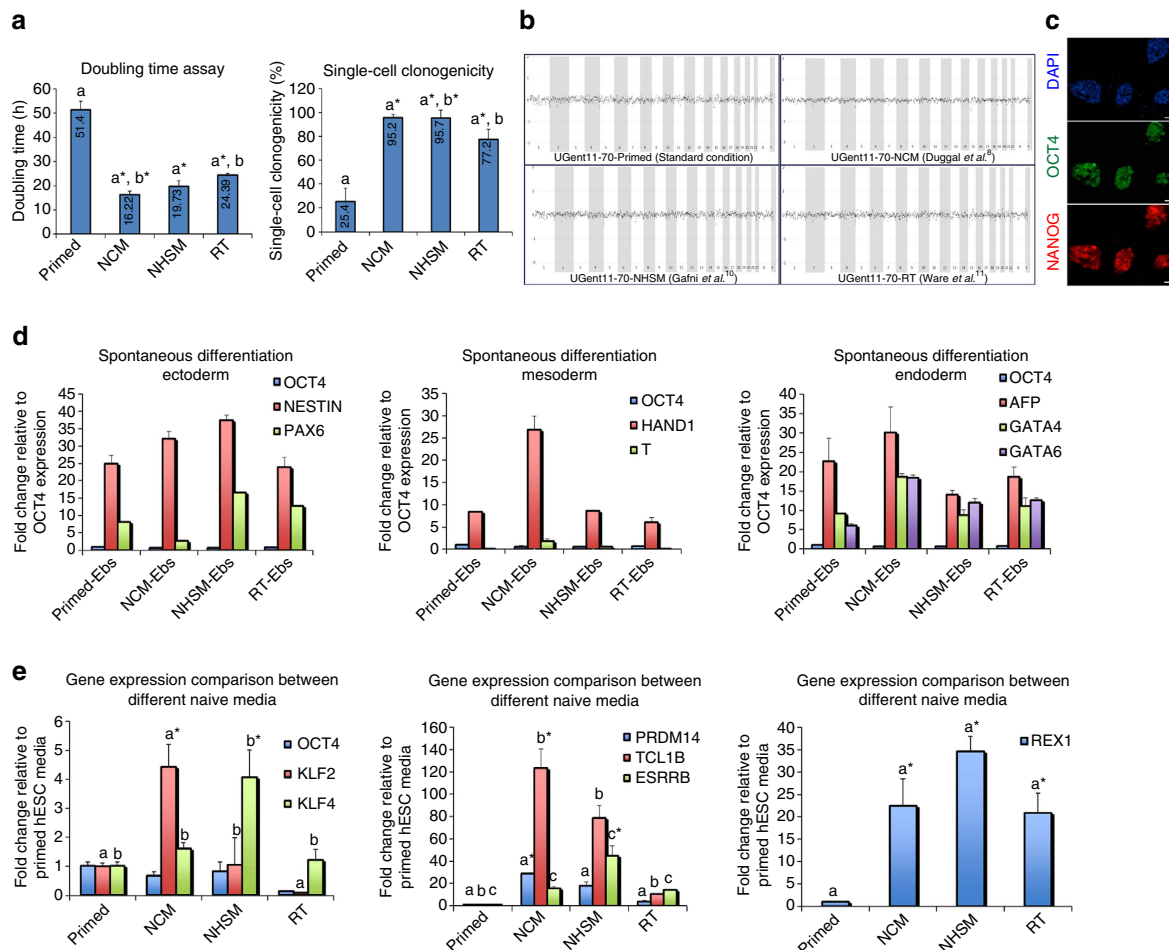
**Figure 1 | Conversion of primed hESCs towards a naive state in different naive media combinations.** (a) Overview of small molecules and growth factors supplemented in NCM, NHSM and RT naive media. (b) Morphological analysis of hESC cultures in different naive media and their transition from primed pluripotency-specific flat morphology towards dome-shaped undifferentiated naive colonies. NCM and NHSM media comprised of several domed colonies whereas their frequency was lower in RT media. Scale bars, 200  $\mu$ m.

bodies (EBs) derived from NCM-MEF naive hESCs showed a consistent differentiation efficiency towards ectoderm (*NESTIN*), mesoderm (*HAND1*) and endoderm (*AFP*, *GATA4*, *GATA6*). EBs from NHSM-MEF naive hESCs expressed lower levels of *HAND1* and those from RT-MEF naive hESCs showed decreased expression of both *NESTIN* and *HAND1* (Fig. 2d).

#### Naive hESCs rely on mTORC2 subunit to maintain pluripotency.

We further confirmed naive pluripotency by quantitative real-time PCR (qRT-PCR) gene expression analysis on the three naive hESC lines and their primed counterparts. Significantly increased *KLF2* ( $n = 3$ ,  $P < 0.05$ ) expression was observed in NCM-MEF naive hESCs, while *KLF4* ( $n = 3$ ,  $P < 0.05$ ) was expressed higher in NHSM-MEF naive hESCs. RT-MEF naive hESCs showed reduced expression of these markers, similar to primed hESCs. *OCT4* expression was consistent in all hESC lines (Fig. 2e). The naive marker *REX1* was highly expressed in all naive hESC lines compared to primed ( $n = 3$ ,  $P < 0.05$ ), however, lower expression observed in RT-MEF naive hESC lines compared to NCM- and NHSM-MEF naive hESC lines (Fig. 2e, Supplementary Figs 1 and 2).

We observed a significant upregulation of naive markers including *KLF2* (6–10-fold), *PRDM14* (5–40-fold) and *TCL1B* (60–80-fold) in hESCs converted in the NCM-MEF protocol compared to NHSM- and the RT-MEF protocol ( $n = 3$ ,  $P < 0.05$ , Fig. 2e, Supplementary Figs 1 and 2). This suggests that in NCM-MEF naive hESCs, *KLF2* and *PRDM14* (a PR-domain-containing transcriptional regulator), both known to work synergistically to facilitate the reprogramming of EpiSCs towards a naive mESC-like state of pluripotency<sup>17–20</sup> have a similar role in inducing naive pluripotency in hESCs. Interestingly, NHSM-MEF naive hESCs continued to show significantly increased expression of *ESRRB* ( $n = 3$ ,  $P < 0.05$ ) compared to NCM- and RT-MEF naive hESCs (Fig. 2e, Supplementary Figs 1 and 2). This correlates with their elevated *KLF4* expression level (Fig. 2e Supplementary Figs 1 and 2), as *ESRRB* is a direct target of *KLF4*, as seen in mESCs both also being downstream targets of *NANOG* (refs 21,22). The activation of the LIFR $\beta$ /gp130 receptor by LIF can activate the PI3K pathway, which acts on the serine/threonine protein kinase B (PKB or AKT) downstream effectors<sup>23</sup>. *TCL1B* is known to enhance the AKT kinase activity and induce naive pluripotency<sup>24</sup>. This is further facilitated by *PRDM14*, which inhibits FGF signalling and activates AKT-mTORC signalling<sup>17–19</sup>. This was



**Figure 2 | Assessment of naive pluripotency in the newly formed naive hESCs from primed hESC line UGent11-70.** (a) Doubling time and single-cell clonogenicity assays for primed hESCs, NCM-, NHSM- and RT-naive hESCs. a\* compared to a and b\* compared to b,  $P < 0.05$ . (b) Depiction of chromosomal analysis for primed hESCs, NCM-, NHSM- and RT-naive hESCs via array comparative genome hybridization (aCGH). (c) Representative immunofluorescence image for pluripotency makers OCT4 and NANOG in NCM-, NHSM- and RT-naive hESCs; Scale bars, 200  $\mu\text{m}$ . (d) qRT-PCR analysis for ectoderm, mesoderm and endoderm markers on spontaneously differentiated primed hESCs, NCM-, NHSM- and RT-naive hESCs as EBs for 14 days. (e) Gene expression analysis for naive-pluripotency specific genes on undifferentiated primed hESCs, NCM-, NHSM- and RT-naive hESCs. a\* compared to a and b\* compared to b,  $P < 0.05$ . The data are represented as mean  $\pm$  s.d.

further confirmed upon exposing primed, NCM-MEF, NHSM-MEF, RT-MEF naive hESCs to rapamycin (mTORC1 inhibitor)<sup>25</sup> and PP242 (mTORC1/2 inhibitor)<sup>26</sup> (Supplementary Fig. 3). Compared to primed hESCs, all naive hESCs exhibited increased apoptosis in response to rapamycin treatment, which increased in PP242 condition (Supplementary Fig. 3). Similarly, NANOG expression was almost absent in PP242-treated naive hESCs whereas rapamycin-treated naive hESCs showed marked downregulation in NANOG levels (Supplementary Fig. 3). Hence, as naive hESCs express increased levels of *PRDM14* and *TCL1B* in contrast to primed hESCs, albeit at varying levels, and undergo apoptosis upon exposure to mTORC inhibitors (especially mTORC2 inhibitor), our results suggest an intriguing role of the PI3K/AKT/mTORC-signalling pathway in facilitating the naive state in hESCs. This is supported by the knowledge that the self-renewal of naive mESCs is dependent on PI3K/AKT, LIF/STAT3, BMP/SMAD signalling which work together to inhibit primed pluripotency-associated ERK and GSK3 $\beta$  pathways, thereby allowing *Nanog* and *c-Myc* to maintain pluripotency<sup>27,28</sup>. Importance of mTORC has been highlighted in naive mESCs and mouse embryos whereby, disruption of mTORC resulted in their reduced cell size as well as impaired

proliferation<sup>29</sup>. High levels of bFGF also activate the PI3K/AKT/mTOR pathway, leading to the inhibition of MAPK/ERK (MEK) signalling essential for the naive pluripotent state while low FGF continues to activate the MEK pathway<sup>30</sup>. Hence, increased concentration of bFGF (8–12 ng ml<sup>-1</sup>) in all naive media directly relates to the upregulation of PI3K/AKT/mTOR and suppression of the MEK pathway in naive hESCs. This is contrary to the primed state, in which MEK signalling is maintained via PI3K/AKT signalling, due to the lower FGF concentration (4 ng ml<sup>-1</sup>)<sup>31</sup>. To date, only one group has established naive hESCs in the absence of bFGF<sup>13</sup>, hence the role of bFGF in the establishment and maintenance of naive hESC remains to be elucidated.

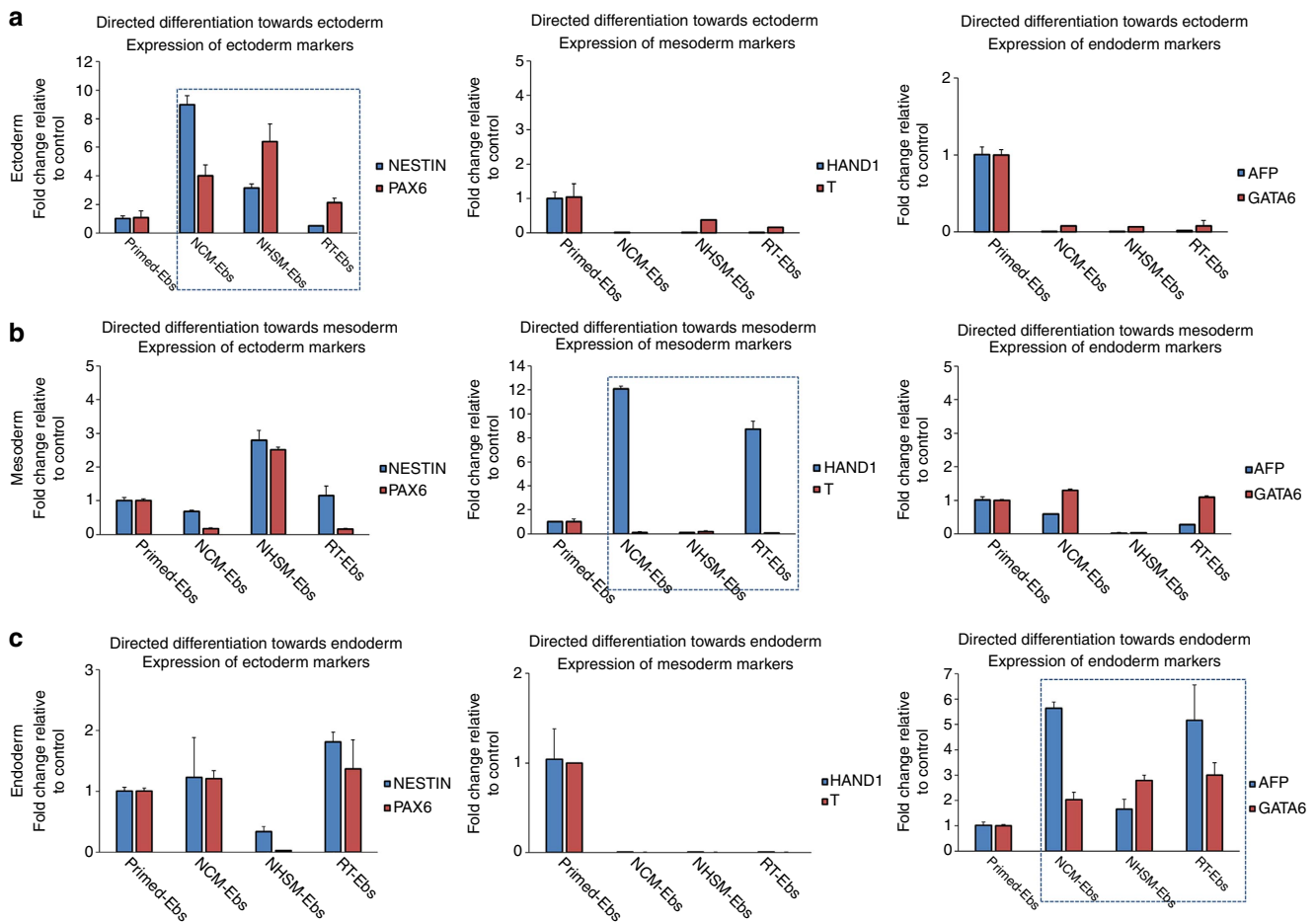
#### Naive hESCs differentiate robustly towards immature cell types.

Further, we aimed to study the lineage-specific differentiation potential of naive hESCs. Upon exposing the EBs derived from NCM-, NHSM- and RT-MEF naive hESCs to ectodermal differentiation cues, we observed high upregulation of *NESTIN* and *PAX6*, by 8- and 6-fold ( $n = 3$ ,  $P < 0.05$ ) in NCM and NHSM-EBs, respectively, compared to RT-EBs and their primed

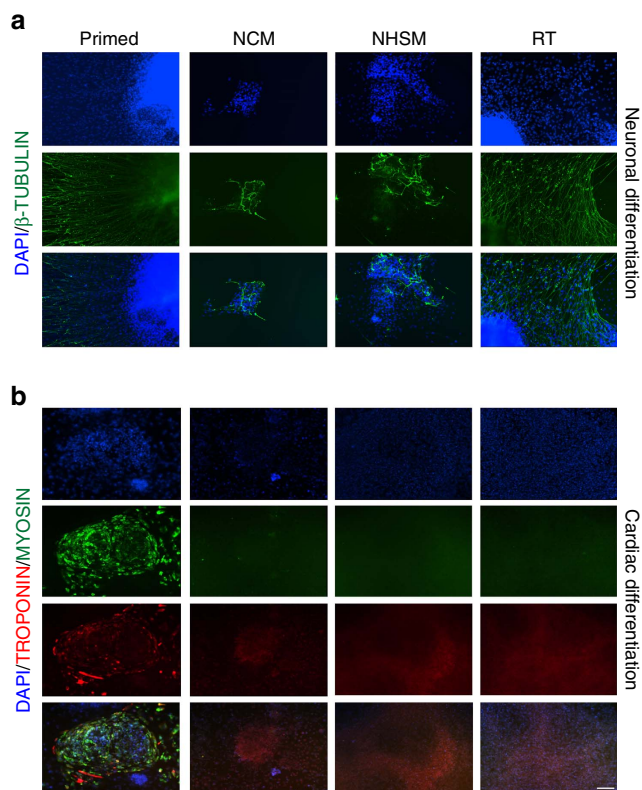
counterparts (Fig. 3a, Supplementary Figs 4 and 5). We observed no expression of mesodermal or endodermal markers in all naive-EBs, whereas EBs from primed hESCs continued to express markers of all lineages, albeit at low levels. This indicates that naive hESCs are transcriptionally more homogenous, similar to mESCs<sup>7,9</sup>. When subjected to mesoderm-directed differentiation, NCM-EBs exhibited highly efficient differentiation, with *HAND1* significantly upregulated by 10-fold ( $n = 3$ ,  $P < 0.05$ ) compared to NHSM-EBs, which showed lower potential to differentiate towards mesoderm and continued to express ectoderm genes. RT-EBs also exhibited efficient mesoderm differentiation, as depicted by a 2 to 8-fold increase in *HAND1* expression compared to NHSM-EBs and primed-EBS (Fig. 3b, Supplementary Figs 4 and 5). Finally, NCM-EBs, NHSM-EBs and RT-EBs underwent robust endoderm-directed differentiation indicated by increase in *AFP* and *GATA6* expression levels ( $n = 3$ ,  $P < 0.05$ ), while primed hESC-EBs continued to express ecto and mesodermal markers (Fig. 3c, Supplementary Figs 4 and 5). To study the functional capacity of these converted naive hESCs compared to their primed origins, we induced neuronal<sup>32</sup> and cardiac differentiation<sup>33</sup> using well-established protocols for hESC. Differentiation towards the neuronal lineage facilitated the induction of neuronal rosettes for all conditions, although only primed hESCs formed well-characterized and recognizable neuronal processes (Supplementary Fig. 6a). For cardiac

differentiation, spontaneously beating population of cells were only observed in cells differentiated from primed hESCs (Supplementary Fig. 6b). Although both naive and primed hESCs expressed early neuronal marker *PAX6* and cardiac progenitor markers *NKX2.5/GATA4* (Supplementary Figs 7,8), only differentiated primed hESCs (and to some extent RT-naive hESCs) expressed the mature neuronal markers  $\beta$ -TUBULIN and TROPONIN/MYOSIN for mature cardiac lineage (Fig. 4). This supports the absence of spontaneously beating clusters in cardiomyocytes differentiated from naive hESCs, in contrast to primed hESC where beating clusters were observed.

Overall, although we see upregulation of lineage-specific markers for naive hESCs, their ability to undergo terminal differentiation towards functional cell types is limited compared to primed hESCs in the current setting. Similar to germ cell differentiation from hESCs<sup>34</sup>, priming of the naive cells to an epiblast-like pluripotency state maybe required. However, further studies are warranted to test whether this epiblast-like state is the equivalent or better than the primed counterpart. It also has to be taken into account that currently used differentiation protocols have been optimized using primed hESC, which mostly differ from the protocols used for directed differentiation in mESC. Hence, these results suggest that the current well-established differentiation protocols for primed hESCs need optimization when applied to naive hESCs to obtain mature cell types.



**Figure 3 | Germ-layer specific differentiation of naive hESC line UGent11-70.** Gene expression analysis for lineage-specific genes upon directed differentiation of naive hESC line UGent11-70 towards (a) ectoderm, (b) mesoderm and (c) endoderm. All germ layer markers were analysed irrespective of the targeted lineage to determine the heterogeneity in differentiation between the different naive medium conditions as well as between primed and naive conditions. Dashed box represents the gene expression for markers specific for that lineage. The data are represented as mean  $\pm$  s.d.



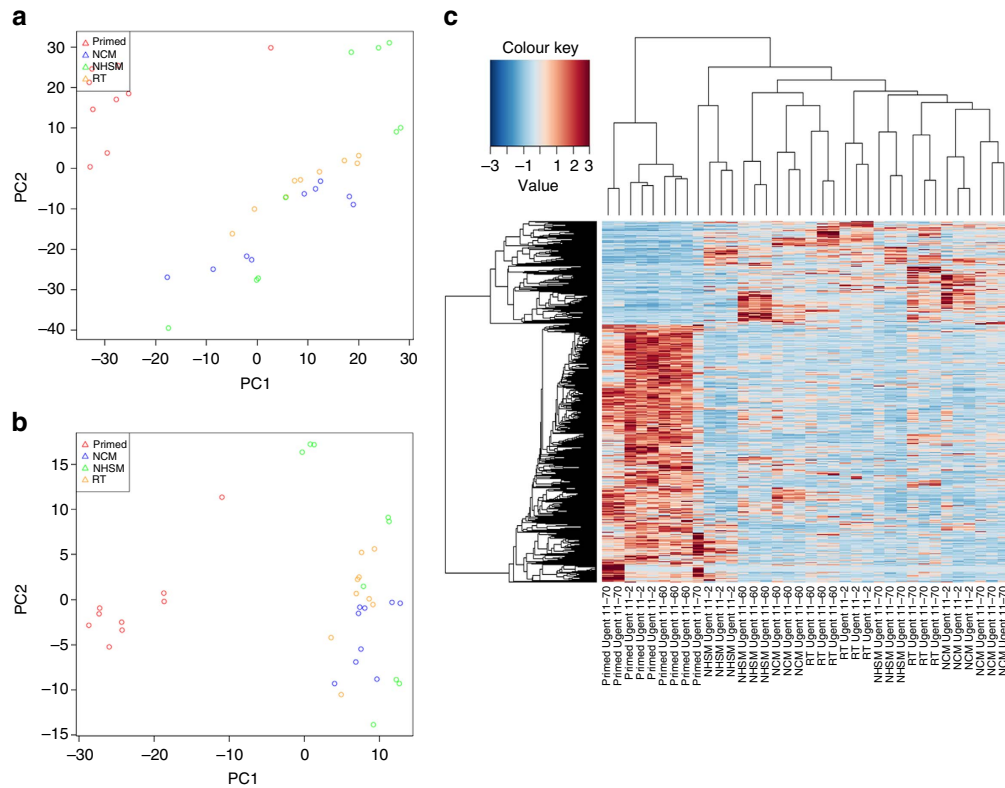
**Figure 4 | Directed differentiation of primed and naive hESCs towards neuronal and cardiac lineages.** (a) Representative images of primed, NCM-, NHSM- and RT-MEF naive hESCs expressing  $\beta$ -TUBULIN when differentiated towards neuronal lineage.  $\beta$ -TUBULIN is expressed higher in neurons derived from primed and RT-MEF naive hESCs compared to NCM- and NHSM-MEF naive hESCs. (b) Representative images of primed, NCM-, NHSM- and RT-MEF naive hESCs expressing mature cardiac markers TROPONIN and MYOSIN when differentiated towards cardiomyocytes. Primed hESCs when differentiated towards cardiac lineage expressing markers TROPONIN and MYOSIN whereas all naive hESCs fail to express the mature markers. Scale bar, 200  $\mu$ m.

**Naive hESCs exhibit unique transcriptional profile.** Naive pluripotency-associated transcriptomic reconfiguration was determined using RNA-seq by comparing multiple and independent cultures of three primed hESC lines and their naive counterparts. A first principal component analysis (PCA) on normalized data showed that the primed cells separated from the converted naive hESC cells (Fig. 5a, Supplementary Data 1). Statistical differential expression analysis comparing primed hESC (selected as the control group) to all the naive converted cells revealed that  $\sim 1,638$  genes (false discovery rate (FDR)  $< 0.05$ ) were differentially expressed between the naive and primed hESCs. PCA and clustering analysis based on the differentially expressed genes separated all naive hESCs cultured in NCM-, NHSM- and RT-MEF naive media and exhibited a unique transcriptome compared to primed hESCs (Fig. 5b,c, Supplementary Data 2). GO analysis showed enrichment for terms related to RNA processing, DNA replication and repair, regulation of cell cycle process and downregulation of terms related to extra cellular adhesion and various developmental processes (Supplementary Data 3,4). Differential KEGG pathway analysis showed upregulation of pathways related to oxidative phosphorylation, cell cycle regulation whereas pathways related to differentiation were downregulated in the naive hESC (Supplementary Data 5). During human embryo development,

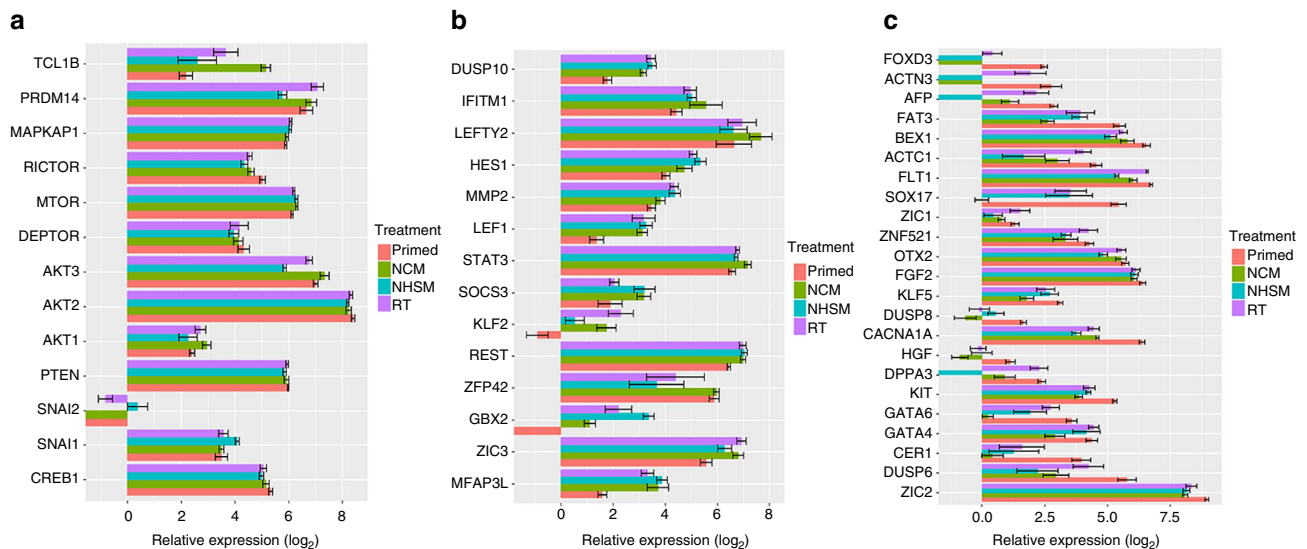
there is an increase in dependence on oxidative phosphorylation as glucose can negatively affect early embryogenesis, and the pre-implantation blastocyst relies on this metabolic state facilitated by mitochondrial biogenesis and maturation<sup>35</sup>. Hence, the dependence of naive hESCs on oxidative phosphorylation instead of glycolytic metabolism, similar to human pre-implantation epiblast, is an important hallmark indicative of their naive state<sup>35–37</sup>. Next, to understand the differences between primed parental hESCs and the individual naive conversion conditions, differential expression was analysed between the primed hESC and the individual treatment groups. This analysis revealed 2,019, 1,996 and 2,007 genes to be differentially expressed when NCM, NHSM and RT conditions were individually compared to the primed condition, respectively (Supplementary Data 6–8). Gene set enrichment analysis showed upregulation of categories related to embryonic stem cell core and stemness in all naive hESCs. Interestingly, categories related to mitochondria, MYC targets and TGF $\beta$  were upregulated in NCM and NHSM conditions only (Supplementary Data 9–11).

In-depth dissection of the data for a subset of key genes required for acquisition of naive pluripotency led to interesting observations. *AKT1* (also known as protein kinase B), is required by the PI3K/AKT pathway to facilitate cell-cell signalling via insulin-like growth factor, which leads to activation of the mTOR and inhibition of the GSK3 $\beta$  pathway<sup>24</sup>. We observed upregulation of *AKT1* in the naive hESCs (specifically in NCM-MEF,  $n = 3$ ,  $P < 0.05$ ), while it continued to be lowly expressed in the primed and NHSM-MEF naive hESCs (Fig. 6a). Increased *AKT1* phosphorylation is facilitated by the mTORC2 subunit, hence promoting cell proliferation and growth. *PTEN*, required in maintaining the negative feedback loop mechanism in PI3K/AKT pathway was expressed in all hESCs<sup>38</sup>. *RICTOR* and *MAPKAP1*, known genes exclusive to mTORC2 subunit<sup>39</sup> showed increased expression compared to *DEPTOR*, a common subunit between mTORC1 and mTORC2 subunits, in the naive lines. Although primed hESCs exhibited a similar expression profile for *PTEN*, *mTOR*, *RICTOR*, *MAPKAP1* and *DEPTOR*, elevated expression of *PRDM14* (NCM- and RT-MEF), *TCL1B* (all naive hESC) and *AKT1* (NCM- and RT-MEF) in naive hESCs shows their increased dependence on PI3K/AKT/mTORC pathway<sup>19,24</sup>, specifically the mTORC2 subunit, as further validated by our inhibitor assay. *SNAI1*, although an epithelial-to-mesenchymal transition factor, is known to enhance reprogramming in human and mouse somatic cells when overexpressed<sup>40</sup>. Interestingly, we observed *SNAI1* expression in naive hESCs increased, specifically in NHSM-MEF, compared to primed hESCs (Fig. 6a). Naive pluripotency markers including *LEF1*, *KLF2*, *ZIC3*, *LEFTY2*, *STAT3*, *HES1*, *SOCS3*, *MMP2*, *GBX2* were significantly increased ( $P < 0.05$ ) in all naive hESCs (Fig. 6b) whereas more primed pluripotency-associated genes such as *DUSP8*, *CACNA1A*, *HGF*, *CER1*, *FGF2* continued to be significantly downregulated<sup>8,10,21,41</sup> (Fig. 6c). Naive hESCs, known to exist in a hypomethylated state<sup>42</sup>, showed a similar expression of *TET* enzymes compared to primed hESCs, however, all naive hESCs showed decreased levels of *de novo* methyltransferase *DNMT3A*, *DNMT3B* compared to primed hESCs whereas *DNMT3L* was significantly increased, similar to as observed in reset cells<sup>13</sup> (Supplementary Data 6–8). Comparison of the different naive conversion conditions with each other (*F*-test, ANOVA,  $FDR < 0.05$ ) 4,195 genes were differentially expressed (Supplementary Fig. 9, Supplementary Data 12), such as naive markers including *GBX2*, *KLF2*, *DPPA3*, *SOAT1*, *PRDM14*, *TCL1B*, *SALL1* and differentiation-associated genes such as *GATA4*, *GATA6*, *AFP*, *SOX17*, *OTX2* between the three naive culture conditions.

Overall, we report an unbiased molecular and transcriptional analysis of naive hESCs generated in different culture conditions.



**Figure 5 | Genome-wide expression analysis of parental primed hESC and converted naive hESC in different conversion conditions. (a)** PCA based on normalized transcriptome data. **(b)** PCA based on rlog values of the differentially expressed genes (FC > 2, FDR < 0.05) between primed hESC and all naive hESC. **(c)** Hierarchical clustering and heatmap of differentially expressed transcriptome data between parental primed and all naive hESCs.



**Figure 6 | Panel of genes of interest. (a)** Genes involved in PI3K/AKT/mTORC pathway. **(b)** Genes related to naive pluripotency upregulated in converted naive hESCs. **(c)** Genes related to lineage specification downregulated in converted naive hESCs. The data are represented as mean ± s.e.m.

Specifically, in-house-derived primed hESCs were converted in the same laboratory, unlike previous studies that conducted comparative analysis exclusively on published data sets. We also demonstrate that although the naive hESCs can differentiate efficiently towards immature cell types, the current differentiation protocols shown to be working efficiently for primed hESCs, probably require further optimization in order to be used for naive hESCs, as these two pluripotency states reflect the different

stages of embryo development. Alternatively, it could be possible that naive pluripotent stem cells require a priming towards a more differentiated stage, before they acquire the full potential of efficient differentiation. Importantly, we confirm the role of PI3K/AKT/mTORC signalling in facilitating the induction of naive pluripotency in not only primed hESCs converted in NCM-MEF conditions, but also in NHSM- and RT-MEF naive hESCs.

## Methods

**Culturing of primed hESCs.** In-house-derived primed hESC lines UGent11-2 (XX), UGent11-60 (XX) and UGent11-70 (XY)<sup>8</sup> were maintained on mitomycin C treated-mitotically inactivated MEFs in low oxygen conditions (5%). The primed cell cultures were fed every day with fresh hESC medium composed of knockout-DMEM, 20% knockout serum replacement, 1% non-essential amino acids (NEAA), 1% l-glutamine, 1% penicillin/streptomycin, 0.1 mM  $\beta$ -mercaptoethanol (Invitrogen) and recombinant human basic fibroblast growth factor (bFGF, 4 ng ml<sup>-1</sup>, Peprotech). Confluent cell cultures (usually 5 days) were passaged using collagenase (1 mg ml<sup>-1</sup>, Invitrogen) and were split at a ratio of 1:2 to 1:4.

**Conversion of primed hESCs to a naive state in different naive media.** Three primed hESC lines were subjected to naive conversion on MEFs in low oxygen conditions in either: (i) naive human stem cell medium (NHSM) composed of basal hESC media supplemented with Albumax1, N2 supplement (Invitrogen), PD0325901 (1  $\mu$ M, Cayman) and CHIR99021 (3  $\mu$ M, Axon Medchem; known together as 2i), leukaemia inhibitory factor (LIF, 1000U, Sigma), bFGF, (8 ng ml<sup>-1</sup>), SP600125 (10  $\mu$ M, Tocris), SB203580 (10  $\mu$ M, Tocris) and TGF $\beta$  (1 ng ml<sup>-1</sup>, Peprotech)<sup>10</sup>; (ii) RT consisting of high glucose DMEM-F12 (Invitrogen) further supplemented with histone deacetylase inhibitors, sodium butyrate (0.1 mM) and SAHA (50 nM) for first three passages followed by 2i and bFGF (10 ng ml<sup>-1</sup>)<sup>11</sup> or (iii) our novel medium naive conversion medium (NCM) containing hESC media with added 2i, LIF, forskolin (10  $\mu$ M, Sigma), ascorbic acid (50 ng ml<sup>-1</sup>, Sigma) and bFGF (12 ng ml<sup>-1</sup>)<sup>8</sup>. Upon the appearance of domed-shaped colonies, the culture was passaged using 0.05% trypsin and once the culture was well established, cells were split at a ratio of 1:4 to 1:8 every 3 days. Single-cell clonogenicity was determined from the number of single-cells plated (upon dissociating primed and naive hESCs using trypsin without ROCKi) and the subsequent number of colonies formed. Doubling time was calculated using the formula:

$$\text{Doubling time} = \frac{\text{duration} \times \log(2) / \log(\text{final concentration})}{-\log(\text{initial concentration})}$$

**Immunocytochemistry.** Undifferentiated naive hESCs in the different media were fixed using 4% paraformaldehyde at room temperature (rt) for 20 min or at 4 °C overnight (O/N). The cells were then permeabilized with 0.1% Triton X-100 in 1  $\times$  phosphate buffered saline (PBS) for 8 min, washed for 5 min in 1  $\times$  PBS and blocked in 1% bovine serum albumin in 1  $\times$  PBS for 1 h at rt. Primary antibodies, anti-goat OCT4 (Santa Cruz Biotechnology), anti-rabbit NANOG (R&D Systems), anti-rabbit PAX6 (Abcam), anti-mouse NKX2.5 (Abcam), anti-rabbit GATA4 (Santa Cruz Biotechnology), anti-mouse  $\beta$ -TUBULIN (Sigma-Aldrich), anti-rabbit TROPONIN (Santa Cruz Biotechnology) and anti-mouse MYOSIN (Chemicon) were diluted in the blocking solution (1:200) and the cells were then incubated with primary antibodies at 4 °C O/N. Next day, the cells were washed with 1  $\times$  PBS three times for 5 min and were incubated with secondary antibodies (diluted in blocking solution) for 1 h at rt. The secondary antibodies applied were donkey anti-goat FITC (1:200, Bioconnect), donkey anti-mouse Alexa 488 (IgG H&L, Life Technologies) and donkey anti-rabbit CY3 (1:200, Bioconnect). Finally, the cells were rinsed again in 1  $\times$  PBS three times for 5 min and mounted onto glass slides in Vectashield with 4',6-diamidino-2-phenylindole (DAPI).

**Differentiation assays.** Spontaneous, directed, neuronal and cardiac differentiation assays were carried out as previously described<sup>8,32,33</sup> in low oxygen conditions. In brief, for neuronal differentiation, cells were maintained on MEFs until they reached ~80% confluency and were passaged to form EBs in 390 ml DMEM/F12, 5 ml NEAA, 5 ml 100  $\times$  GlutaMAX, 36  $\mu$ l 2-mercaptoethanol and 100 ml KOSR (Invitrogen). On day 4, the media was changed to neural induction medium (NIM) composed of 490 ml DMEM/F12, 5 ml NEAA, 5 ml N2 supplement (Invitrogen) and 20 mg ml<sup>-1</sup> Heparin (Sigma). On day 7, EBs were plated in NIM containing 5% fetal bovine serum to promote attachment for 8 h and were refreshed gently with fresh NIM once the EBs were attached. Neural rosettes were observed between days 10–12. For cardiomyocyte differentiation, cells were plated as single cells in the presence of ROCKi (Enzo Life) on matrigel (Corning). Differentiation was initiated once the cells reached 80–90% confluency and were cultured in RPMI/B27 (without insulin, Invitrogen) in the presence of CHIR9902 (12  $\mu$ M, Axon MedChem) for 24 h. Next, the media was changed to only RPMI/B27 (without insulin) and 72 h later, half the media was refreshed with additional supplementation WNT inhibitor IWP2 (5  $\mu$ M, Sigma) in the media. On day 5, the cells were refreshed with fresh RPMI/B27 (without insulin) medium and day 7 onwards, the cells were refreshed with RPMI/B27 (with insulin). Spontaneous contractions were observed between days 10–12. For spontaneous differentiation, primed, as well as different, naive hESCs were trypsinized and re-suspended in differentiation medium (DFM) similar to hESC medium without bFGF and the knockout serum replacement replaced with fetal bovine serum and cultured as EBs on 24-well Costar ultra-low attachment plates (Corning) for 14 days with medium changed every other day. For lineage-specific differentiation, cells were either exposed to DFM with 10  $\mu$ M retinoic acid for ectodermal differentiation for 10 days; to DFM supplemented with 5 ng ml<sup>-1</sup> bFGF and 10 mM nicotinamide

(Sigma) for mesoderm formation for 8 days; or to DFM consisting of Activin A (20 ng ml<sup>-1</sup>, R&D Systems) and BMP4 (50 ng ml<sup>-1</sup>, R&D Systems) for an additional 7 days.

**Quantitative real-time PCR.** RNA isolation, cDNA synthesis, quantitative real-time PCR (qRT-PCR) gene expression and data analysis was performed as previously described<sup>8</sup> for primed hESCs, EBs and naive hESCs and EBs. Briefly, RNA isolation and cDNA synthesis was performed using RNeasy Mini Kit (Qiagen) and iSCRIPT Advanced cDNA synthesis Kit (Biorad), respectively, and the cDNA concentration was determined using Qubit fluorometric quantitation method (ThermoFisher Scientific). For subsequent qRT-PCR gene expression analysis, 10 ng cDNA was used for each reaction per gene, performed in technical and biological replicates on an ABI 7000 sequence detection system (Applied Biosystems). Housekeeping genes included *GAPDH*, *B2M* and *RPL13A*, naive pluripotency markers used were *KLF2*, *KLF4*, *ESRRB*, *PRDM14*, *REX1*, *TCL1B*, *OCT4* and *NANOG*. Germ layer-specific markers used were *NESTIN* and *PAX6* (ectoderm), *T-BRACHYURY*, *HAND1* and *PECAM1* (mesoderm), *AFP*, *GATA4* and *GATA6* (endoderm). CT values were normalized to housekeeping genes and fold change was analysed using the  $\Delta\Delta$ CT method where all samples were compared to their respective controls in each experiment. Statistical significance was calculated using the ANOVA and *t*-test on fold change gene expression values ( $2^{-\Delta\Delta$ CT). No RT (only RNA) or no template (water) were used as negative controls.

**Array-based comparative genome hybridization.** Whole-genome amplification of the samples was performed using the SurePlex Amplification system following manufacturer's instructions (Illumina Bluegenome, United Kingdom). A total of 2.5  $\mu$ l of control DNA (G1521; Promega; 187 ng  $\mu$ l<sup>-1</sup>) was used at a concentration of 25 pg  $\mu$ l<sup>-1</sup>. Whole-genome amplification products were hybridized on 24sure + arrays (Illumina Bluegenome, UK) according to manufacturer's instructions. A male reference sample was used for all tested samples. Briefly, amplified samples (from one colony/naive condition at passage 48) and reference DNA were labelled with Cy3 and Cy5 fluorophores, respectively, for 4 h at 37 °C. After combining sample and reference with Cot1 DNA, the mixture was precipitated for 20 min at -80 °C. The pellets were re-suspended in pre-warmed hybridization buffer and hybridized on a 24sure + array for 16 h (overnight) at 42 °C. After washing, microarrays were scanned on an Agilent G2565CA microarray scanner (Agilent, USA). Data were analysed using the BlueFuse Multi software (Illumina Bluegenome, United Kingdom) and visualized in the Vivar software.

**RNA-seq method and data analysis.** Total RNA was extracted using Qiagen RNeasy mini kit, according to the manufacturers' protocol, including a DNase digestion. Concentration and quality of the total extracted RNA was checked by using the 'Quant-it ribogreen RNA assay' (Life Technologies, Grand Island, NY, USA) and the RNA 6000 nano chip (Agilent Technologies, Santa Clara, CA, USA), respectively. Subsequently, 500 ng of RNA was used to perform an Illumina sequencing library preparation using the QuantSeq 3' mRNA-Seq Library Prep Kits (Lexogen, Vienna, Austria) according to manufacturer's protocol. Libraries were quantified by qPCR, according to Illumina's protocol 'Sequencing Library qPCR Quantification protocol guide', version February 2011. A High sensitivity DNA chip (Agilent Technologies, Santa Clara, CA, USA) was used to control the library's size distribution and quality. Sequencing was performed on a high throughput Illumina NextSeq 500 flow cell generating 75 bp single reads.

Per sample, on average  $5.8 \times 10^6 \pm 1.5 \times 10^6$  reads were generated. First, these reads were trimmed using cutadapt<sup>43</sup> version 1.11 to remove the 'QuantSeq FWD' adaptor sequence. To remove contamination with mouse DNA, the reads were mapped against the *Mus musculus* GRCh38.p5 reference genome using CLC genomics workbench version 9.0.1 (CLC). Only the reads that didn't map against the GRCh38.p5 genome were subsequently mapped against the *Homo sapiens* hg19 GRCh37.74 genome using CLC. Count tables were made by counting the reads that mapped against the genes defined by the Ensembl GRCh37.74 Gene transfer format file and exported from CLC.

To explore if the samples from different treatment groups clustered together and to detect outlier samples, a PCA was performed in R on the  $\log^{44}$  normalized and transformed counts of the genes expressed at c.p.m. above 1 in at least two samples.

Differential gene expression analysis was performed using edgeR<sup>45</sup>. Genes were only retained if they were expressed at a c.p.m. above 1 in at least two samples. In one differential gene expression analysis, a general linear model was made with two treatments 'primed' versus 'non-primed' (including the 'NCM', 'NHSM' and 'RT' samples), defining the effects from the three cell lines 'UGent11-60', 'UGent11-2' and 'UGent11-70' as batch effect. Statistical testing was done using the empirical Bayes quasi-likelihood *F*-test. Genes having a FDR < 0.05 and a fold change > 2 were considered significantly differential. PCA was performed in R on  $\log^{44}$  transformed c.p.m. of the differentially expressed genes. To produce the heat map, the normalized count data of the differentially expressed genes was rescaled between -3 and 3 and clustered based on Pearson correlation<sup>46</sup>. In an alternative differential gene expression analysis, a general linear model was made with four treatments 'primed', 'NCM', 'NHSM' and 'RT', defining the effects from the three



cell lines 'UGent11-60', 'UGent11-2' and 'UGent11-70' as batch effect. The latter model was used to statistically compare 'primed' versus 'NCM', 'primed' versus 'NHSM' and 'primed' versus 'RT'. Statistical testing was done using the Bayes quasi-likelihood *F*-test. Genes with an FDR < 0.05 and a fold change > 2 were considered significantly differential. To compare the three 'non-primed' treatments, a general linear model was made with three treatments 'NCM', 'NHSM' and 'RT', defining the effects from the three cell lines 'UGent11-60', 'UGent11-2' and 'UGent11-70' as batch effect. The latter model was used to statistically test any difference between the three 'non-primed' treatments using empirical Bayes quasi-likelihood *F*-test. Genes with a FDR < 0.05 were considered significantly differential. PCA was performed in R on  $\log^{44}$  transformed c.p.m. of the differentially expressed genes. To produce the heat map, the normalized count data of the differentially expressed genes was rescaled between -3 and 3 and clustered based on Pearson correlation<sup>46</sup>.

Differential expression of pathways in the gene set data collection 'kegggs' derived from the KEGG database, was based on log fold changes of the pairwise comparison of 'primed' versus 'NCM', 'primed' versus 'NHSM' and 'primed' versus 'RT' using the GAGE R package<sup>47</sup>. Pathways were considered differentially expressed when the average log fold change of the pathway was significantly different from the overall average log fold change at an FDR < 0.05. Differentially expressed pathways were visualized using Pathview<sup>48</sup>, showing the log fold change for each comparison and each enzyme in the pathway. Analogously, a differentially expression analysis was performed on the biological and the molecular function GO-terms and the chemical and genetic perturbations gene set from the Molecular Signatures Database<sup>49</sup>. Gene sets were considered differentially expressed when the average log fold change of the gene set was significantly different from the overall average log fold change at an FDR < 0.05.

**Data availability.** The authors declare that all supporting data are submitted under BioSample accession numbers SAMN06111297, SAMN06111298, SAMN06111299, SAMN06111300, SAMN06111301, SAMN06111302, SAMN06111303, SAMN06111304, SAMN06111305, SAMN06111306, SAMN06111307, SAMN06111308 and also supplied as Supplementary Datasets.

## References

- Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145–1147 (1998).
- Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl Acad. Sci. USA* **78**, 7634–7638 (1981).
- Brons, I. G. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191–195 (2007).
- Tesar, P. J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196–199 (2007).
- Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487–492 (2009).
- Nichols, J. & Smith, A. The origin and identity of embryonic stem cells. *Development* **138**, 3–8 (2011).
- Ying, Q.-L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
- Duggal, G. *et al.* Alternative routes to induce naive pluripotency in human embryonic stem cells. *Stem Cells* **33**, 2686–2698 (2015).
- Honda, A. *et al.* Naive-like conversion overcomes the limited differentiation capacity of induced pluripotent stem cells. *J. Biol. Chem.* **288**, 26157–26166 (2013).
- Gafni, O. *et al.* Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282–286 (2013).
- Ware, C. B. *et al.* Derivation of naive human embryonic stem cells. *Proc. Natl Acad. Sci. USA* **111**, 4484–4489 (2014).
- 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).
- Takashima, Y. *et al.* Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* **158**, 1254–1269 (2014).
- Ernst, M. *et al.* Comparative computational analysis of pluripotency in human and mouse stem cells. *Sci. Rep.* **5**, 7927 (2015).
- Huang, K., Maruyama, T. & Fan, G. The naive state of human pluripotent stem cells: a synthesis of stem cell and preimplantation embryo transcriptome analyses. *Cell Stem Cell* **15**, 410–415 (2014).
- 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).
- Grabole, N. *et al.* *Prdm14* promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. *EMBO Rep.* **14**, 629–637 (2013).
- Nakaki, F. & Saitou, M. *Prdm14*: a unique regulator for pluripotency and epigenetic reprogramming. *Trends Biochem. Sci.* **39**, 289–298 (2014).

- Yamaji, M. *et al.* *PRDM14* ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell* **12**, 368–382 (2013).
- Gillich, A. *et al.* Epiblast stem cell-based system reveals reprogramming synergy of germline factors. *Cell Stem Cell* **10**, 425–439 (2012).
- Martello, G. *et al.* *Esrrb* is a pivotal target of the *Gsk3/Tcf3* axis regulating embryonic stem cell self-renewal. *Cell Stem Cell* **11**, 491–504 (2012).
- Guo, G. *et al.* *Klf4* reverts developmentally programmed restriction of ground state pluripotency. *Development* **136**, 1063–1069 (2009).
- Paling, N. R., Wheadon, H., Bone, H. K. & Welham, M. J. Regulation of embryonic stem cell self-renewal by phosphoinositide 3-kinase-dependent signaling. *J. Biol. Chem.* **279**, 48063–48070 (2004).
- Pekarsky, Y. *et al.* *Tcl1* enhances Akt kinase activity and mediates its nuclear translocation. *Proc. Natl Acad. Sci. USA* **97**, 3028–3033 (2000).
- Maiese, K. Stem cell guidance through the mechanistic target of rapamycin. *World J. Stem Cell.* **7**, 999–1009 (2015).
- Feldman, M. E. *et al.* Active-site inhibitors of mTOR target rapamycin-resistant outputs of mTORC1 and mTORC2. *PLoS Biol.* **7**, e38 (2009).
- Singh, A. M., Bechard, M., Smith, K. & Dalton, S. Reconciling the different roles of *Gsk3β* in 'naive' and 'primed' pluripotent stem cells. *Cell Cycle* **11**, 2991–2996 (2012).
- Wang, X. Q. *et al.* *CDK1-PDK1-PI3K/Akt* signaling pathway regulates embryonic and induced pluripotency. *Cell Death Differ.* **24**, 38–48 (2017).
- Murakami, M. *et al.* mTOR is essential for growth and proliferation in early mouse embryos and embryonic stem cells. *Mol. Cell. Biol.* **24**, 6710–6718 (2004).
- Wray, J. *et al.* Inhibition of glycogen synthase kinase-3 alleviates *Tcf3* repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat. Cell. Biol.* **13**, 838–845 (2011).
- Dalton, S. Signaling networks in human pluripotent stem cells. *Curr. Opin. Cell Biol.* **25**, 241–246 (2013).
- Liu, Y. *et al.* Directed differentiation of forebrain GABA interneurons from human pluripotent stem cells. *Nat. Protoc.* **8**, 1670–1679 (2013).
- Lian, X. *et al.* Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating *Wnt/β-catenin* signaling under fully defined conditions. *Nat. Protoc.* **8**, 162–175 (2013).
- Irie, N. *et al.* *SOX17* is a critical specifier of human primordial germ cell fate. *Cell* **160**, 253–268 (2015).
- Folmes, C. D. & Terzic, A. Metabolic determinants of embryonic development and stem cell fate. *Reprod. Fertil. Dev.* **27**, 82–88 (2014).
- Sperber, H. *et al.* The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition. *Nat. Cell Biol.* **17**, 1523–1535 (2015).
- Wu, J. & Izpisua Belmonte, J. C. Metabolic exit from naive pluripotency. *Nat. Cell Biol.* **17**, 1519–1521 (2015).
- Tang, Y. *et al.* Differential effects of Akt isoforms on somatic cell reprogramming. *J. Cell Sci.* **127**, 3998–4008 (2014).
- Pearce, L. R. *et al.* Identification of *protor* as a novel *ricor*-binding component of mTOR complex-2. *Biochem. J.* **405**, 513–522 (2007).
- Unterhaehrer, J. J. *et al.* The epithelial-mesenchymal transition factor *SNAIL* paradoxically enhances reprogramming. *Stem Cell Rep.* **3**, 691–698 (2014).
- Chen, H. *et al.* Reinforcement of *STAT3* activity reprogrammes human embryonic stem cells to naive-like pluripotency. *Nat. Commun.* **6**, 7095 (2015).
- Leitch, H. G. *et al.* Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol. Biol.* **20**, 311–316 (2013).
- Martin, M. *et al.* *Cutadapt* removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
- Love, M. I. *et al.* Moderated estimation of fold change and dispersion for RNA-seq data with *DESeq2*. *Genome Biol.* **15**, 550 (2014).
- Robinson, M. D. *et al.* *edgeR*: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
- Culhane, A. *et al.* *MADE4*: an R package for multivariate analysis of gene expression data. *Bioinformatics* **21**, 2789–2790 (2005).
- Luo, W. *et al.* *GAGE*: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics* **10**, 161 (2009).
- Luo, W. & Brouwer, C. *Pathview*: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* **29**, 1830–1831 (2013).
- Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

## Acknowledgements

S.W., D.D., L.P., A.V.S., P.D.S. and this research are supported by the Concerted Research Actions funding from Bijzonder Onderzoeksfonds University Ghent (BOF GOA 01G01112). M.P. is holder of a doctoral grant provided by the Special Research Funds (Bijzonder Onderzoeksfonds, BOF, University Ghent). P.D.S. is the holder of a fundamental clinical research mandate by the Research Foundation - Flanders (FWO).

We thank Ferring Pharmaceuticals (Aalst, Belgium) for an unrestricted educational grant. D.D., and this research are supported by the Flemish Foundation for Scientific Research (FWO Vlaanderen, Grant Number FWO-3G062910). M.V.d.J. held a PhD grant from the Agency for Innovation by Science and Technology (IWT, Grant Number SB093128), Belgium. P.D.S. is holder of a Fundamental Clinical Research Mandate from the Flemish Foundation of Scientific Research (FWO Vlaanderen), Belgium. S.C.d.S.L. is supported by The Netherlands Organization for Scientific Research (Grant Number NWO ASPASIA 016.121.365) and by the Interuniversity Attraction Poles Program (IUAP-07/07). B.H. and this research are supported by a Ghent University grant, Faculty Health and Medical Sciences; a research grant from Research Foundation - Flanders (FWO; FWO/KAN/ 1507816N) and the Agency for Innovation by Science and Technology (IWT, Project Number 150042). P.M. is a postdoctoral researcher supported by the Foundation for Scientific Research (FWO Vlaanderen). This research has been conducted through collaboration with the Bimetra biobank, a high quality bio-repository for the University Hospital Ghent and the University of Ghent.

### Author contributions

S.W. contributed to conception and design, collection and assembly of data, data analysis and interpretation, manuscript writing; M.V.d.J. contributed to manuscript writing; G.D. contributed in collection of data, data analysis and interpretation, manuscript writing; E.S. helped in collection of data; J.T. contributed to maintenance of stem cell cultures, collection of data; M.P. contributed to manuscript writing; S.L. contributed to maintenance of stem cell cultures; S.C.d.S.L. contributed to data analysis and manuscript writing; A.V.S. contributed to manuscript writing; L.P. contributed to manuscript writing; L.T. contributed to RNA-seq based gene expression data analysis and interpretation; F.V.N. contributed to RNA-seq based gene expression data analysis and interpretation; D.I.M.D.C. contributed to RNA seq based gene expression data analysis and interpretation; B.M. conducted microarray-based gene expression analysis; P.M. contributed to manuscript writing; J.V.d.S. contributed to manuscript writing; D.D.

contributed to data analysis and manuscript writing; P.D.S. contributed to data analysis and manuscript writing; B.H. contributed to conception and design, and manuscript writing.

### Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing interests:** The authors declare no competing financial interest.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Warrior, S. *et al.* Direct comparison of distinct naive pluripotent states in human embryonic stem cells. *Nat. Commun.* **8**, 15055 doi: 10.1038/ncomms15055 (2017).

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



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017