OCT4 cooperates with distinct ATP-dependent chromatin remodelers in naïve and primed pluripotent states in human

Xin Huang
Columbia University

Kyoung-mi Park
Washington University School of Medicine in St. Louis

Paul Gontarz
Washington University School of Medicine in St. Louis

Bo Zhang
Washington University School of Medicine in St. Louis

Laura A. Fischer
Washington University School of Medicine in St. Louis

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Please let us know how this document benefits you.

Recommended Citation
Huang, Xin; Park, Kyoung-mi; Gontarz, Paul; Zhang, Bo; Fischer, Laura A.; Dong, Chen; Dietmann, Sabine; Xing, Xiaoyun; Khan, Shafqat A.; Jensen, Nick; Wang, Ting; Theunissen, Thorold W.; and al, et, "OCT4 cooperates with distinct ATP-dependent chromatin remodelers in naïve and primed pluripotent states in human." Nature Communications. 12, 1. 5123 (2021).
https://digitalcommons.wustl.edu/open_access_pubs/10654
Authors
Xin Huang, Kyoung-mi Park, Paul Gontarz, Bo Zhang, Laura A. Fischer, Chen Dong, Sabine Dietmann, Xiaoyun Xing, Shafqat A. Khan, Nick Jensen, Ting Wang, Thorold W. Theunissen, and et al

This open access publication is available at Digital Commons@Becker: https://digitalcommons.wustl.edu/open_access_pubs/10654
Understanding the molecular underpinnings of pluripotency is a prerequisite for optimal maintenance and application of embryonic stem cells (ESCs). While the protein-protein interactions of core pluripotency factors have been identified in mouse ESCs, their interactome in human ESCs (hESCs) has not to date been explored. Here we mapped the OCT4 interactomes in naïve and primed hESCs, revealing extensive connections to mammalian ATP-dependent nucleosome remodeling complexes. In naïve hESCs, OCT4 is associated with both BRG1 and BRM, the two paralog ATPases of the BAF complex. Genome-wide location analyses and genetic studies reveal that these two enzymes cooperate in a functionally redundant manner in the transcriptional regulation of blastocyst-specific genes. In contrast, in primed hESCs, OCT4 cooperates with BRG1 and SOX2 to promote chromatin accessibility at ectodermal genes. This work reveals how a common transcription factor utilizes differential BAF complexes to control distinct transcriptional programs in naïve and primed hESCs.
Pluripotency, the ability of a single cell to give rise to all cell types found in an organism, is a fundamental characteristic of embryonic stem cells (ESCs). Mouse and human ESCs are both derived from the inner cell mass (ICM) of the pre-implantation blastocyst but differ significantly in their epigenomic, morphological, and transcriptomic features. Mouse ESCs (mESCs) are marked by early developmental properties associated with naïve pluripotency. In contrast, human ESCs (hESCs) derived under conventional culture conditions are developmentally more advanced and resemble mouse epiblast stem cells (EpiSCs); thus, they are considered to represent a primed state of pluripotency. Over the past few years, specific combinations of inhibitors and cytokines have been developed to enable the generation of naïve hESCs by resetting primed hESCs with transposon-associated sequence (3xFLAG tag) recognition sequence (3xFLBio) to the C-terminus of endogenous OCT4–3xFLBio hESCs retained full expression of pluripotency genes and differentiation potential (Supplementary Fig. 1g, h). OCT4-associated proteins were identified by SA and 3xFLAG AP-MS under primed conditions and upon transfer to naïve pluripotency by overexpression of KLF2 and NANOG transgenes (Fig. 1a). In addition, we performed native OCT4 antibody pulldown followed by mass spectrometry in WIBR2 and WIBR3 hESCs before and after resetting using our chemically defined naïve conditions. Candidate OCT4 partners were selected based on detection in at least three of four independent MS experiments, removal of background contaminants using the CRAbome7 database, and spectral count analysis using a combined cumulative probability (CCP) score with a false-discovery rate (FDR) cutoff of 0.1 (Supplementary Fig. 1i, j and Supplementary Data 1). This analysis identified 30 OCT4-associated proteins in naïve hESCs and 14 in primed hESCs, 9 of which were detected in both pluripotent states (Fig. 1b). We compared the OCT4 interactome in naïve hESCs with that in mESCs. Only 7 (23%), 7 (23%), and 13 (43%) proteins were shared between the naïve human OCT4 interactome and three independent mouse studies, respectively (Fig. 1c). While this limited overlap may hint at species-specific differences in OCT4 protein associations, it is important to bear in mind that the mouse data used in these comparisons were generated under serum/LIF conditions, which do not truly represent ground state conditions, but rather a heterogeneous and metastable state that is intermediate between naïve and primed pluripotency. Components of the cohesin complex (SMC1A, SMCHD1), RNA binding protein L1TD1, DNA mismatch repair proteins (MSH2, MSH6), and multiple components of SNF2-family ATP-dependent chromatin remodeling complexes (Fig. 1c, underlined) were conserved between the mouse and human OCT4 networks. Interactions of OCT4 with selected partners in naïve and primed hESCs were validated by co-immunoprecipitation (co-IP) and western blot (Fig. 1d, e). These results indicate that a subset of OCT4 interaction partners is conserved between naïve cells in mouse and human.

**Results**

**Establishing OCT4-centered protein interactomes in naïve and primed hESCs.** To identify OCT4 partners in human pluripotent cells, we captured the dynamic OCT4-centered protein–protein interaction networks (interactomes) under the naïve and primed conditions using affinity purification followed by mass spectrometry (AP-MS) and uncovered extensive associations with ATP-dependent chromatin remodelers. Integrative analysis generated a reference map of stage-specific transcription factors (TFs) and chromatin remodelers, linking enhancer reorganization with concordant transcriptional changes. Our work indicates that a switch in OCT4 partner association contributes to the activation of distinct target genes in naïve and primed hESCs and consequently their distinct developmental potential.

OCT4-associated SNF2-family ATPases are divergently expressed in naïve and primed hESCs. The human OCT4 interactome was strongly enriched in components of mammalian SNF2-family ATP-dependent chromatin remodeling complexes, including members of the Brahma-associated factor (BAF) or mammalian SWI/SNF and imitation SWI (ISWI) complexes (Fig. 2a). The BAF complex contains one of two mutually exclusive ATPase subunits, BRG1 (SMARCA4) or BRM (SMARCA2), while other BAF subunits provide structural or chromatin targeting functions. BRG1 was detected in both the naïve and primed OCT4 interactomes, while its paralog ATPase, BRM, was only detected in the naïve OCT4 interactome (Fig. 1b, Fig. 2a, and Supplementary Fig. 2a). In a similar fashion, ISWI complexes contain one of two ATPase subunits, SNF2H (SMARCA5) or SNF2L (SMARCA1), and 2–4 additional subunits. SNF2H was detected in both the naïve and primed OCT4 interactomes, while SNF2L was only detected in the primed OCT4 interactome (Fig. 1b and Fig. 2a). Hence, OCT4 is associated with distinct SNF2-family ATPases in naïve and primed hESCs.

We examined whether these distinct protein interactions may be explained by differential expression of OCT4 partners. Indeed, BAF and SNF2L were specifically expressed in naïve and primed hESCs, respectively (Fig. 2b, c). Other BAF and ISWI components were expressed at similar levels in naïve and primed hESCs (Fig. 2b, c and Supplementary Fig. 2b), although BRG1 was elevated at the protein level in primed cells (Fig. 2b). Similar
Fig. 1 An OCT4 interactome in human embryonic stem cells (hESCs). a Affinity purification mass spectrometry (AP-MS) strategies for identification of OCT4-associated proteins (interactome) in hESCs. Three strategies are used: (i) streptavidin immunoprecipitation (SA-IP) of WIBR2-OCT4\textsuperscript{3xFLBio} ESCs or BirA control; (ii) FLAG-IP of WIBR2-OCT4\textsuperscript{3xFLBio} ESCs or BirA control; (iii) native OCT4 and IgG antibody (Ab-IP) pulldown in WIBR2 and WIBR3 hESC lines. b OCT4-associated proteins identified in naïve and primed hESCs. Proteins are listed as official gene symbols. c Proteins in the naïve human OCT4 interactome are compared to OCT4-associated proteins identified in three independent studies in mouse ESCs\textsuperscript{29–31}. Components of SNF2-family chromatin remodeling complexes are underlined. d OCT4 co-immunoprecipitation (co-IP) followed by western blot of selected OCT4-associated proteins in naïve and primed hESCs. e BRM (SMARCA2) and BRG1 (SMARCA4) co-IP followed by western blot analysis of BAF155 (SMARCC1) and OCT4. d, e Experiment is repeated independently twice with similar results.
expression patterns were observed in naïve hESCs generated using t2i/L/Gö, an alternative culture formulation for naïve hESCs (Fig. 2d). Some BAF subunits that were identified in the naïve OCT4 interactome nonetheless showed an interaction with OCT4 by co-IP assay in both naïve and primed cells (Supplementary Fig. 2c), suggesting that the stringent cutoffs used in our AP-MS analysis may have caused some bona fide interactions to be omitted. Similarly, the topoisomerase TOP2A, which synergizes with the BAF complex to recruit pluripotency factors in mouse ESCs, was identified in the naïve OCT4

**Fig. 2 Differential expression of OCT4-associated SNF2-family ATPases in naïve and primed hESCs.**

*a* Illustration of naïve-specific (red/pink), primed-specific (blue), and common (yellow) components of the mammalian SWI/SNF family of subunits (represented as one complex rather than three subcomplexes cBAF, PBAF, and nCBaF) and ISWI complexes identified in the naïve and primed OCT4 interactomes. Pink color indicates proteins that are identified with marginal significance (0.1 < FDR < 0.2) in the naïve OCT4 interactome.

*b* Western blot analysis of ATPase expression in naïve and primed hESCs. Experiment is repeated independently twice with similar results.

*c* Log2 ratios of mRNA expression of BAF and ISWI components in naïve and primed hESCs based on a microarray analysis.

*d* Expression of BRG1 (SMARCA4), BRM (SMARCA2), and SNF2L (SMARCA1) mRNA transcripts in the alternative t2iLGoY culture condition for naïve human pluripotency (red) and parental primed hESCs (blue) according to RNA-seq data. *SNF2L* was quantified by MS only in WIBR2 cells.
interactome but its CCP score was only slightly below the cutoff for inclusion in the primed OCT4 interactome (Supplementary Fig. 1j). Conversely, its isoform TOP2B was upregulated in primed cells and exclusively detected in the primed OCT4 interactome (Supplementary Fig. 1j, k). Co-IP assays indicated that SNF2L and TOP2B form a direct interaction with BRG1 in primed hESCs (Supplementary Fig. 2d). These data suggest that some of the protein associations with OCT4 may be bridged by third partners, such as BRG1.

To determine whether other OCT4-associated proteins were differentially expressed between naïve and primed hESCs, we performed stable isotope labeling with amino acids in cell culture (SILAC) followed by MS to determine global protein expression levels (Supplementary Fig. 2e). This analysis confirmed that BRG1 was upregulated at the protein level in primed hESCs, whereas several naïve-specific OCT4 partners (BRM, EPPK1, and NLRP2/7) were highly induced in naïve hESCs (Fig. 2e and Supplementary Fig. 2f). However, most OCT4-associated proteins did not show a significant shift in expression between naïve and primed hESCs. We conclude that cell-type-specific association with OCT4 can be attributed to differential protein expression in only a handful of cases, including BRM (naïve specific) and SNF2L (primed specific).

OCT4 and BAF complexes shape the naïve- and primed-specific chromatin landscapes. To explore the role of the BAF complexes in regulating naïve versus primed human pluripotency, we performed chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) analysis using antibodies specific for BRG1 and BAF155 (SMARCC1), a common scaffold protein of the BAF complex42 under naïve and primed conditions (ChIP-seq QC metrics in Supplementary Data 2 and source data provided with this paper). A total of 69,007 BAF (BRG1/BAF155 merged) peaks were identified in both states, among which 26,428 peaks (53.0%, 26,428/49,846 of naïve peaks, or 58.0%, 26,428/45,589 of primed peaks) were shared (Fig. 3a). Consistent with the known role of BAF complexes in promoting an accessible chromatin architecture, we observed a high overlap of BAF peaks with open chromatin regions in naïve and primed cells, as defined by the assay for transposable-accessible chromatin followed by sequencing (ATAC-seq)48 (Fig. 3b). BAF peaks specific to naïve and primed cells showed higher ATAC intensities in their respective states, as exemplified by genes active in naïve (TFAP2C) and primed (SALL1) states (Fig. 3c, d).

In line with the physical interaction between OCT4 and the BAF components, OCT4 peak regions from ChIP-seq35 overlapped substantially (57.9%, 14,659/25,329 in naïve state or 74.1%, 15,839/21,359 in primed state) with BAF peaks (Fig. 3e). Using a clustering method based on the ChIP-seq intensities of OCT435 and histone H3 modification marks K4me3, K27ac, and K27me34 at all BAF peak regions (N = 69,007), we defined 10 discrete BAF-bound clusters with distinct occupancy patterns of OCT4 and active and repressive histone marks in naïve and primed hESCs (Fig. 3f, Supplementary Fig. 3a, b, and Supplementary Data 3). Of particular interest for understanding differential gene regulation, clusters 1, 6, and 8 (C1, C6, C8) represent naïve-specific, primed-specific, and shared enhancers, respectively, that were enriched in OCT4 and the active histone mark H3K27ac but lacked the promoter mark H3K4me3 and the repressive mark H3K27me3. Additionally, clusters 4 and 10 (C4, C10) represent shared and naïve-specific active promoters, respectively, that were enriched in OCT4, H3K27ac, and H3K4me3, but lacked H3K27me3, while cluster 7 (C7) represents shared bivalent promoters that were enriched in both H3K4me3 and H3K27me3 (Fig. 3f and Supplementary Fig. 3a, b). As expected, annotations of these peak regions agreed with their predicted activity in naïve vs. primed hESCs based on chromatin accessibility and expression of the nearest genes (Supplementary Fig. 3c, d).

By examining the presence of TF binding motifs, we identified enrichment of TFAP2A/C and KLF4/5 motifs at naïve-specific OCT4/BAF-bound enhancers (C1) and promoters (C10) (Fig. 3g and Supplementary Fig. 3e). KLF4 and TFAP2C were previously shown to activate enhancers specific to naïve hESCs4,6. In contrast, primed-specific OCT4/BAF-bound enhancers (C6) were enriched in motifs associated with ZIC2/3 and SOX2/3 (Fig. 3g). Zic2/3 have been implicated as transcriptional determinants of primed mouse EpiSCs49,50, while SOX2 is a master regulator of pluripotent cells that also contributes to neural differentiation51,52. These findings suggest that the BAF complex is recruited to discrete sets of target genes in naïve and primed hESCs through association with TFs that are differentially expressed in these two pluripotent states.

To determine the in vivo relevance of these OCT4/BAF-bound enhancers, we examined stage-specific enhancers that were identified in early human embryos using ATAC-seq analysis33. Distal ATAC-seq peaks that were identified specifically in human 8-cell embryos, ICM, and primed hESCs were compared to naïve-specific (C1), primed-specific (C6), and naïve/primed-shared (C8) enhancer regions. C1 regions were associated with 8-cell and ICM-specific ATAC peaks, while C8 regions were only associated with ICM-specific ATAC peaks in vivo (Fig. 3h, i). These data confirm the biological relevance of the identified OCT4/BAF-bound enhancers and demonstrate that naïve hESCs partially recapitulate enhancer signatures of human 8-cell-to-ICM embryos10.

Naïve and primed-specific BAF-bound enhancers harbor distinct chromatin remodelers and TFs. Our OCT4 interactome studies identified BRM as a naïve-specific partner (Fig. 1b). ChIP-seq analysis identified 17,413 BRM peaks in naïve hESCs, most of which (88.8%, 15,463/17,413) were shared with BAF (BRG1/BAF155) peaks in naïve cells (Supplementary Fig. 4a). Previous studies suggested that the BRM- and BRG1-containing BAF complexes are mutually exclusive54,55. In naïve hESCs, we detected an interaction between these two ATPases by reciprocal co-IP experiments (Supplementary Fig. 4b). However, the interaction of BRM with BRG1, but not other BAF components (BAF53A, BAF47), was abolished using a more stringent co-IP condition (Supplementary Fig. 4c). Hence, the two ATPases are likely not assembled in the same BAF complex, which is consistent with structural studies42,56, and their association in naïve hESCs may be bridged by a third interactor, such as OCT4.

The enrichment of the KLF4 motif at naïve-specific (C1) enhancers and of the SOX2 motif at primed-specific (C6) enhancers (Fig. 3g) prompted us to further investigate the genome-wide location of KLF4 and SOX2 in naïve and primed hESCs by CUT&Tag and ChIP-seq analysis, respectively (Supplementary Data 2). Interestingly, expression of KLF4 and SOX2 was largely naïve and primed specific, respectively (Supplementary Fig. 4d, e). Accordingly, association of KLF4 to chromatin was only detected in naïve hESCs, while that of SOX2 was detected with a much larger proportion in primed than naïve hESCs (Fig. 4a and Supplementary Fig. 4f). In naïve hESCs, BRM and KLF4 were enriched at naïve specific and common BAF (BRG1/BAF155) peaks, while SOX2 occupied all BAF peaks with a low intensity. In contrast, SOX2 was more enriched at primed specific and common BAF regions in primed hESCs (Fig. 4b). These data suggest that SOX2 may function as a largely primed-specific TF in the human context, although its genome-wide
Fig. 3 Genome-wide binding profiles of OCT4 and BAF complexes shape the naïve- and primed-specific enhancer landscapes. a Overlap of BAF (BRG1/BAF155) peaks in naïve and primed hESCs identified by ChIP-seq analysis. b Overlap of BAF and chromatin accessibility (ATAC-seq) peaks in naïve (left) and primed (right) hESCs. The two tracks of each factor between naïve and primed hESCs are normalized to the same RPM (reads per million) values.

c K-means clustering analysis of all BAF (N = 69,007) peaks (shown in panel a) in naïve and primed hESCs identified 10 clusters with distinct patterns of OCT4 and histone marks enrichment.

d Highly enriched TF binding motifs in regions C1 (left), C6 (middle), and C10 (right) from clustering analysis (shown in panel f). OSTN: OCT4/SOX2/TCF/NANOG consensus motif.

e Overlap of 8-cell-specific, ICM-specific, and primed hESC-specific ATAC-seq peaks (h) and plots of mean ATAC intensity (i) at clusters C1, C6, and C8 peak regions. The stage-specific distal ATAC regions are from a recent study of chromatin accessibility in human preimplantation embryos.

---

ARTICLE NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-021-25107-3

NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-021-25107-3 | www.nature.com/naturecommunications

6 (2021) 12:5123 | https://doi.org/10.1038/s41467-021-25107-3 | www.nature.com/naturecommunications
Fig. 4 Naïve and primed-specific enhancers harbor distinct chromatin remodelers and TFs. a Mean intensity plots of BRM, SOX2 (ChIP-seq), and KLF4 (CUT&Tag) signals in naïve and primed hESCs. b Mean intensity plots of BRM, KLF4, and SOX2 signals at naïve-only, primed-only, and naïve/primed-common BAF peaks. c Correlation analysis of OCT4 and SOX2 chromatin association in naïve and primed hESCs grouped by naïve-only, primed-only, and naïve/primed-common OCT4 or SOX2 peaks. d Intensity heatmaps of BAF subunits, KLF4, SOX2, OCT4, and TFAP2C signals at clusters C1, C6, and C8 regions in naïve (left) and primed (right) hESCs. e Summary of activities of chromatin remodeling complexes and TFs in naïve (C1), primed (C6), and shared (C8) enhancers in hESCs. f Expression of OCT4-associated chromatin remodelers and TFs in single-cell RNA-seq analysis of human embryos cultured in vitro between days 6 and 14 of human development. Average gene expression levels and the percentage of cells that express each gene are presented with differential color intensities and circle sizes, respectively. Note that KLF17 and SUSD2 are markers associated with naïve pluripotency, while SFRP2 and SOX11 are associated with primed pluripotency.
binding in naïve hESCs was still weakly correlated with that of OCT4 (Fig. 4c). We then examined the association of stage-specific chromatin remodelers and TFs with BAF-bound enhancers. BRG1, BRM, KLF4, OCT4, and TFAP2C were enriched at naïve-specific (C1) and naïve/primed-shared (C8) enhancers, while BRG1, SOX2, and OCT4 were enriched at primed-specific (C6) and naïve/primed-shared (C8) enhancers (Fig. 4d and Supplementary Fig. 4g). Thus, the naïve and primed-specific enhancers harbor distinct ATP-dependent chromatin remodelers and TFs (Fig. 4e).

To investigate whether such a switch in OCT4 partner association could also be observed in transcriptome studies of human embryos, we analyzed single-cell RNA-seq (scRNA-seq) data from human embryos between days 6 and 14 of development.3,35,36 By profiling cells corresponding to the epiblast (EPI) lineage over this period, we observed elevated expression of the naïve-associated factors BRM, KLF4, and TFAP2C during early timepoints, which diminished after day 9 (Fig. 4f). The primed-specific OCT4 partner, SNF2L, was induced starting at day 9 together with SOX2 and subsequently, ZIC2. In contrast, OCT4, BRG1, and SNF2H maintained a relatively constant expression level over this time course (Fig. 4f). These transcriptional dynamics are consistent with the dynamic reorganization of OCT4 protein partners observed in naïve and primed hESCs in vitro.

Naïve- and primed-specific BAF-bound enhancers drive expression of blastocyst and ectodermal signatures, respectively. We examined the identity of the target genes of the OCT4/BAF-bound enhancers in naïve and primed hESCs. Naïve-specific OCT4/BAF-bound enhancers (C1) were located nearby genes that are involved in blastocyst development, blastocyst formation, and trophectoderm differentiation. Examples of such genes include NLRP7, TEAD4, and TFAP2C (Fig. 5a–c). In contrast, primed-specific OCT4/BAF-bound enhancers (C6) were located nearby genes that are involved in differentiation towards ectoderm (Fig. 5d), which is consistent with recent evidence in the mouse that ectodermal enhancers become primed in the early post-implantation epiblast.3,35 Examples of such genes include SAL1, PTPRZ1, and ZIC3 (Fig. 5e, f). As expected, naïve/primed-shared OCT4/BAF-bound enhancers (C8) were located nearby genes that are involved in stem cell population maintenance (Supplementary Fig. 5a, b). Comparison with scRNA-seq data from human embryos confirmed that naïve-specific enhancer target genes (C1) were more highly expressed during early timepoints, whereas primed-specific enhancer target genes (C6) were more highly expressed after day 9 (Fig. 5g and Supplementary Data 3). Similar expression dynamics were observed in non-human primate embryos (Supplementary Fig. 5c). In accordance with prior studies that examined the developmental identity of naïve hESCs,46 naïve hESCs displayed a human ICM-specific gene expression signature (Fig. 5h), which was confirmed by principle component analysis (PCA) (Fig. 5i). These data indicate that OCT4/BAF-bound enhancers drive the expression of blastocyst-specific genes in naïve hESCs.

Since the primed-specific OCT4/BAF-bound enhancers (C6) were associated with genes involved in ectoderm development, we investigated the activity of these enhancers in human neural progenitor cells (hNPCs). Primed-specific OCT4/BAF-bound enhancers overlapped substantially with SOX2 peaks in hNPCs (Fig. 5j, k). Chromatin accessibility, BRG1 binding, and the enhancer-associated histone marks H3K27ac and H3K4me1 were all highly enriched at primed-specific enhancers (C6) compared to naïve-specific enhancers (C1) in hNPCs (Fig. 5l and Supplementary Fig. 5e). In line with activity of these enhancers, expression of primed-specific enhancer target genes was significantly higher in primed hESCs and hNPCs compared to naïve hESCs (Supplementary Fig. 5d). In addition, we putatively identified a naïve-specific (C1, distance to TSS: 50 kb) and a primed-specific (C6, distance to TSS: 45 kb) enhancer within downstream intergenic regions of SOX2. BRG1 and OCT4 co-occupied the C1 enhancer in naïve hESCs, whereas BRG1 and OCT4/SOX2 co-occupied the C6 enhancer in primed hESCs, and finally BRG1 and SOX2 remained present at the C6 enhancer in hNPCs (Supplementary Fig. 5f). These data reveal a regulatory cascade at distinct enhancers controlled by stage-specific TFs and BAF chromatin remodelers to drive expression of SOX2 in hESCs and hNPCs.

Functional redundancy of BRG1 and BRM in naïve hESCs. We examined the functional significance of the BAF complex in controlling naïve and primed human pluripotency. We performed short hairpin RNA (shRNA)-mediated knockdown (KD) of BRG1 in naïve and primed hESCs, and KD of BRM in naïve hESCs (Fig. 6a). Knockdown resulted in more than 50% reduction in mRNA transcript and protein levels of each target gene (Supplementary Fig. 6a, 6b). Approximately two hundred genes (101 Down, 94 Up, fold change > 2, p value < 0.05) were differentially expressed upon BRG1 KD in primed hESCs, but little effect was seen in naïve hESCs upon depletion of either BAF ATPase (Fig. 6b and Supplementary Data 4). In agreement with prior findings,31,32 BRG1 knockdown in primed hESCs caused upregulation of genes with bivalent chromatin domains, as identified in C7 of our clustering analysis (Fig. 3f and Supplementary Fig. 3b). In particular, regulators of endodermal lineage differentiation, such as SOX17, GATA4, FOXA2, and LGR5, were significantly upregulated in response to BRG1 depletion in primed hESCs (Fig. 6b and Supplementary Fig. 6c–e).

Since KD of neither ATPase had a significant effect on transcription in naïve hESCs, we examined whether BRG1 and BRM may have an overlapping functional role in naïve human pluripotency. We first introduced guide RNA (gRNAs) targeting the ATPase domains of BRG1 or BRM into primed hESCs and genotyped individual clones (Fig. 6c). One third of clones nucleasefected with BRG1 gRNAs contained homologous frameshift mutations in BRG1, but none contained homologous frameshift mutations (Fig. 6d). In contrast, nucleasefection with BRM gRNAs generated equivalent proportions of clones that were wild-type (WT), heterozygous (HET), or knockout (KO) for BRM (Fig. 6d). These results confirm that BRG1, but not BRM, is critical for self-renewal of primed hESCs. We then reprogrammed these primed BRM+/+, BRM−/−, or BRM−/− hESCs to the naïve state and found that cells of all three genotypes expressed the naïve cell-surface markers CD75 and SUSD2, but CD75 intensity was marginally reduced in the absence of BRM (Fig. 6e, f and Supplementary Fig. 6f). Western blot analysis indicated that BRG1 protein levels increased in the absence of BRM, suggesting a potential compensatory mechanism (Fig. 6g). We then attempted to deplete BRG1 in naïve hESCs that were also deficient in BRM. gRNAs targeting the ATPase domain of BRG1 were introduced into WT and two independent clones of BRM−/− naïve hESCs. The frequency of insertion–deletions (indels) in each genetic background was determined by sequencing. While we detected a substantial (43.4%) frequency of BRG1 alleles containing indels in WT naïve cells, very few (<9%) BRG1 indels were detected in the BRM−/− background (Fig. 6h). In
addition, KD of BRG1 in BRM−/− naïve cells resulted in severe collapse of colony morphology and loss of the CD75+/SUSD2+ population compared to KD of BRG1 in WT naïve cells (Fig. 6i). These results indicate that naïve cells deficient in both BAF ATPases are selectively eliminated during culture. Based on the above findings we predicted that it should be possible to completely disrupt BRG1 in naïve hESCs that contain functional BRM. We subcloned BRM+/+ naïve cells that were transfected with gRNAs targeting BRG1 and isolated three clones containing homozygous mutations in the BRG1 ATPase domain (Fig. 7a). These cells lacked detectable BRG1 protein expression (Fig. 7b) but maintained expression of CD75 and SUSD2 (Fig. 7c). Furthermore, RNA-seq analysis indicated few differentially expressed genes (DEGs) between WT, BRM−/−, and BRG1−/− naïve hESCs (Supplementary Fig. 7a). We surmise that neither BRM nor BRG1 is required for maintenance of naïve hESCs, which supports our hypothesis that these two ATPases play a functionally redundant role in naïve human pluripotency.
BRG1 regulates chromatin accessibility during the exit from naïve pluripotency. Since we were unable to derive BRG1−/−-primed hESCs (Fig. 6d), we hypothesized that BRG1 is essential for establishment of primed pluripotency. Naïve hESCs containing different BRG1 genotypes were treated with primed media, a process called “re-priming” (Fig. 7d). WT and BRG1−/− naïve hESCs rapidly acquired a flattened epithelioid morphology typical of primed hESCs within 5 days of repriming. In contrast, BRG1−/− naïve hESCs initially retained a dome-shaped colony morphology (Fig. 7d). In agreement with this delayed morphological transition, WT and BRG1−/− naïve hESCs showed more rapid induction of the primed-specific cell-surface marker CD90 (Fig. 7e). ATAC-seq analysis revealed few differentially accessible regions (DARs, FDR < 0.05) between WT, BRG1−/−, or BRG1−/− cells under naïve conditions (Fig. 7f). In contrast, chromatin accessibility was substantially reduced (3940 out of 3945, or 99.9% DARs) in BRG1−/− cells compared to WT/BRG1−/+ cells during repriming (Fig. 7f). These reduced DARs (N = 3940) were enriched in the primed specific (C6) and naïve/primed-shared (C8) OCT4/BF-bound enhancers identified in our ChIP-seq analysis (Fig. 7g and Supplementary Fig. 7b). The C6 enhancers with reduced chromatin accessibility were associated with target genes that are upregulated in both primed hESCs and hNPCs (Supplementary Fig. 7c). Hence, the removal of BRG1 alters the epigenomic landscape during the naïve-to-primed transition (Fig. 7h and Supplementary Fig. 7d).

Finally, we performed RNA-seq analysis on WT, BRG1−/−, and BRG1−/− hESCs during 5 days of repriming. BRM transcript levels were maintained in WT cells (Supplementary Fig. 7e). There were 1897 and 2064 DEGs significantly (fold change > 2, p value < 0.05) upregulated in WT naïve and primed cells, respectively (Fig. 7i and Supplementary Data 5). Clustering analysis based on these DEGs indicated that cells undergoing repriming acquired an intermediate transcriptional identity (Supplementary Fig. 7f). 499 (out of 1897, 26.3%) and 573 (out of 2064, 27.8%) of these DEGs were significantly downregulated and upregulated, respectively, during repriming compared with the naïve state (Fig. 7i). Consistent with the retention of a naïve colony morphology, the reduction of naïve-specific genes and activation of primed-specific genes were delayed in BRG1−/− cells (Fig. 7j) and Supplementary Fig. 7g). Examples of genes showing delayed repriming kinetics in BRG1−/− cells include the naïve-specific TF DPPA3 and the primed-specific genes GAP43, ROR1, and ZIC3, which also exhibited reduced chromatin accessibility at enhancer and/or promoter regions (Fig. 7k–m and Supplementary Fig. 7h). Hence, while BRG1 appears to be dispensable for maintenance of naïve hESCs, it plays a critical role in ensuring adequate chromatin accessibility during the exit from naïve pluripotency.

Discussion

In summary, we have generated an interactome of protein–protein interactions around OCT4 in primed hESCs and then reconstructed this interactome under recently devised conditions for naïve human pluripotency. Our results indicate that OCT4 engages in dynamic interactions with ATG-dependent chromatin remodelers in human pluripotent states. While the interaction between OCT4 and BRG1 has been well documented in mouse ESCs44,64, its association with distinct chromatin remodelers under naïve and primed conditions was not previously reported. The currently prevailing view is that BRG1 is the key catalytic subunit of BAF complexes in ESCs64 and early embryos65,66, whereas BRM is thought to play a greater role in differentiated cells that are less proliferative.67 By assessing the expression, genome-wide binding, and function of BAF subunits in naïve and primed hESCs, we propose that both BRG1 and BRM contribute to the activation of enhancers involved in blastocyst development and trophectoderm specification in naïve cells. Co-binding of KLF4 and TAP2P2C to OCT4/BF-bound enhancers suggests that these two TFs provide specificity to OCT4 binding in the naive state, as was previously reported for KLF4 during the early stages of human fibroblast reprogramming14. In contrast, OCT4 engages with the BRG1-containing BAF complex and SOX2 to activate ectodermal enhancers in primed hESCs. This precocious activation of ectodermal enhancers in primed hESCs is consistent with recent evidence from mouse embryos that ectodermal enhancers become primed in the early post-implantation epiblast58. Thus, our interactome analysis has uncovered a switch in OCT4 protein partners that orchestrates a dynamic reconfiguration of the enhancer landscape during the interconversion between naïve and primed pluripotent states (Fig. 8). This remodeling of the epigenetic landscape may explain the distinct developmental competencies of naïve and primed hESCs towards trophoblast and neural fates, respectively17.

The temporal expression dynamics of OCT4-associated factors between days 6 and 14 of human embryonic development recapitulate the expression changes observed during the naïve-to-primed transition in hESCs. While BRM and KLF4 are activated in the preimplantation epiblast, SOX2 is most significantly induced at later stages. Indeed, our ChIP-seq data suggest that SOX2 functions largely as a primed-specific TF in the human context, in contrast to its central role in naïve mouse ESCs68. Whether BRM also plays a previously underappreciated role in mouse ESCs and early embryogenesis will require future investigation. However, blastocyst outgrowths from double heterozygous intercrosses indicated a combined gene dosage requirement for Brg1 and Brm during mouse peri-implantation development69. Such functional redundancy may provide a failsafe mechanism to ensure adequate chromatin accessibility at the blastocyst stage before BRM expression is extinguished and BAF

ARTICLE
catalytic activity becomes exclusively reliant on BRG1 in the post-implantation epiblast. This transition is modelled during the naïve-to-primed transition in vitro, which we have shown here is critically dependent on BRG1 to promote adequate chromatin accessibility at primed-specific enhancers and promoters. We propose that BRM and BRG1 have overlapping functions in naïve hESCs, but not during later stages of development. This would explain why many human diseases are caused by haploinsufficiency of BAF subunits\textsuperscript{41}. Consistent with this interpretation, Brm was required to prevent implantation defects but not neural tube defects in Brg1 heterozygous mice\textsuperscript{69}.

An important focus for future studies will be to identify functional enhancers that are activated by the BAF complex during the primed-to-naïve transition. Since BRG1 is essential for the self-renewal of primed cells, such experiments will likely require the use of inducible systems to deplete one or both catalytic subunits of the BAF complex at discrete stages of resetting. In addition, the OCT4 interactome data in this study provide a
rich resource to explore other candidate regulators of human pluripotency. For example, it will be instructive to determine how ISWI complexes (containing SNF2H or SNF2L as ATPases) contribute to epigenetic remodeling in human pluripotent cells. Intriguingly, Snf2h−/− mouse embryos die during the peri-implantation stage, which has been attributed to growth arrest in both ICM and trophectoderm lineages.70 Furthermore, evidence from mouse ESCs indicates that SNF2H, but not BRG1, plays a dominant role in orchestrating chromosome folding and insulation of topologically associated domains.43 Other candidates of interest include the chromatin remodeler BEND371 and the RNA markers CD75 and SUSD2. This experiment is representative of the interest include the chromatin remodeler BEND3 and the RNA markers CD75 and SUSD2. This experiment is representative of

Methods

Cell culture

Primed hESC culture. WIBR2 and WIBR3 primed hESCs were maintained on mitomycin C-inactivated mouse embryonic fibroblast (MEF) feeder layers in primed hESC medium (hESM) and passed using Collagenase IV (Gibco, 1.5 mg/mL), as previously described67,68. hESCs were cultured in DMEM/F12 supplemented with 15% fetal bovine serum, 5% KnockOut Serum Replacement, 1 mM GlutaMax (Gibco, 35050), 1% nonessential amino acids (Gibco, 11140), 1% penicillin-streptomycin (Gibco, 15140), 0.1 mM β-mercaptoethanol and 4–8 ng/mL βFGF. H9 primed hESCs were cultured in mTeSR Plus (STEMCELL Technologies, 05825) on Matrigel hESC-QualiMat (Corning, 354277) coated wells and passaged using Gentle Cell Dissociation Reagent (STEMCELL Technologies, 07174) every 4 to 6 days.

Primed-to-naïve resetting. Chemically induced resetting of primed hESCs to naïve pluripotency was performed as previously described15,16. 2 × 10^5 single primed cells were seeded on a six-well plate with MEF feeder layer in 2 mL primed hESC medium (for WIBR2/3) or mTeSR1 (for H9) supplemented with 10 µM Y-27632 (Stemgent, 04-0012). Two days later, medium was switched to 5%/1% naïve media (see below for details). Ten days after seeding, the cells were expanded polyclonally using Accutase (Gibco, A1110501) or TrypLE Express (Gibco, 12604) on a MEF feeder layer. Media were changed every 1–2 days. For AP-MS experiments, we also performed transgene-mediated resetting to naïve pluripotency as previously described.69 WIBR2-BirA and WIBR2-OC47ΔFLPlus human ESCs were infected with Doxycycline (Dox)-inducible KLF2 and NANOG lentiviral transgenes and maintained in 1 µM PD0325901 (Stemgent, 04-0006), 1 µM CHIR99021 (Stemgent, 04-0004), 20 ng/mL hLIF (Peprotech, 300-05), and 0.5 µg/mL ROCK inhibitor Y-27632 (Stemgent, 04-0012). We also used 5α-reductase activity assay to assess the expression of the 5α-reductase activity assay.

Naïve hESC culture. Naïve hESCs were cultured on mitomycin C-inactivated MEF feeder cells and were passaged by a brief PBS wash followed by single-cell dissociation using 5 min treatment with Accutase or TrypLE Express (Gibco, 12604) and centrifugation in fibroblast medium (DMEM [Millipore Sigma, #D-MEM-021-B] supplemented with 10% FBS [Millipore Sigma, ES-009-B], 1x GlutaMAX, and 1% penicillin-streptomycin). Naïve hESCs were cultured in 5%/1% naïve media as previously described.15 Five hundred milliliters of 5%/1% naïve media was generated by combining 240 mL DMEM/F12 (Gibco, 11320), 240 mL Neurobasal (Gibco, 21103), 5 mL N2 100x supplement (Gibco, 17202), 10 mL B27 50x supplement (Gibco, 17504), 1x GlutaMAX, 1x MEM NEAA (Gibco, 11140), 0.1 mM β-mercaptoethanol (Millipore Sigma, 8.05740), 1% penicillin-streptomycin, 50 µg/mL BSA Fraction V (Gibco, 15260), and the following small molecules and cytokines: 1 µM PD0325901 (Stemgent, 04-0006), 1 µM IM-12 (Enzo, BML-WN102), 0.5 µM SB590885 (Tocris, 2650), 1 µM WH4-023 (A Chemtech, H620061), 10 µM Y-27632 (Stemgent, 04-0012), 20 ng/mL hLIF (Peprotech, 300-05), and 10 ng/mL Activin A (Peprotech, 120-14). For MS experiments, the GSK3 inhibitor IM-12 was omitted from the 5%/1% media to achieve enhanced proliferation of 5%/1% media. As previously described,69,74 targeted and BRG1-targeted naïve hESCs were cultured in 5i/L/IA supplemented with 10 µM Y-27632 and differentiation medium better supported clonal expansion following CRISPR/Cas9-mediated gene editing. mTeSR media consisted of N2B27 supplemented with 1 µM PD0325901, 2 µM XAV939 (Selleckchem, 51180), 20 ng/mL hLIF, 10 µM Y-27632 was added to the culture. All naïve hESC experiments were conducted in 5% O2, 5% CO2, 10% N2, and 1% of each OCT4 TALEN encoding plasmid. To target the OCT4 locus, we cloned these plasmids into pCRII-TOPO vector (Invitrogen) using standard cloning methods. Repriming of naïve hESCs. The naive-to-primitive pluripotency transition (repriming) was performed as previously described.69 Naïve hESCs were single cell dissociated using TrypLE Express and 5 × 10^3 cells were seeded per Matrigel-coated well in mTeSR Plus supplemented with 10 µM Y-27632. The cells were cultured in 5% CO2 and 20% O2. After 2 days, Y-27632 was withdrawn and cells were collected on d5 for RNA-seq analysis. To improve survival and sample quality for ATAC-seq analysis, the reprogramming protocol was modified by seeding naïve hESCs on a MEF feeder layer and cells were harvested 4 days after transfer to mTeSR Plus media.

Donor vectors for gene targeting in hESCs. AAVS1 and OCT4 TALEN plasmids were designed and assembled as previously described.69 Intron 1 of the AAVS1 locus was targeted to constitutively express biotin ligase BirA. The OCT4 locus was targeted to express the 3FLAG sequence followed by a 23 amino acid recognition for biotin ligase BirA, which becomes biotinylated on the lysine residue. Donor vectors were constructed by PCR amplifying homology arms from the corresponding loci using genomic DNA isolated from hESCs. The homology arms followed by either biotin ligase of E. coli BirA or 3FLAG-Biotin sequences were cloned into pCR1-TOPO vector (Invitrogen) using standard cloning methods.

TALEN-mediated gene targeting in hESCs. To target the CAGGS-BirA-V5His sequence to the AAVS1 locus, WIBR2 human ESCs were cultured with ROCK inhibitor (10 µM) 24 h prior to electroporation. Cells were harvested using 0.25% trypsin/EDTA solution (Invitrogen) and resuspended in phosphate buffered saline (PBS). Ten million cells were electroporated with 40 µg of donor vector and 5 µg of each G0213 TALEN encoding plasmid (Gene Pulser Xcell System, Bio-rad (Pulsed E. coli Xcell System, Bio-rad) at 250 V, 500 µF, 0.4 cm cuvettes). Cells were subsequently plated on 3D MEFs feeder layer with neomycin for selection in hESC medium supplemented with 10 µM ROCK inhibitor (Y-27632) for the first 24 h. Individual colonies were picked and expanded after neomycin selection (300 µg/mL) 10–14 days after electroporation. Gene targeting efficiency was verified by Southern blotting (EcoRV digested). To introduce a 3FLAG-Biotin sequence at the C-terminus of endogenous OCT4, WIBR2-BirA hESCs were used and the same electroporation protocol was applied with 40 µg of donor vector and 5 µg of each OCT4 TALEN encoding plasmid. Correctly targeted clones were confirmed by Southern blot analysis (BamHI digested). Then the PGK-Puro selection cassette was removed by transient expression of Cre recombinase in WIBR2-OC47ΔFLPlus hESCs. Briefly, hESCs were transfected with 40 µg of pTurbo-Cre (GenBank accession number AF334827)
and 10 μg fluorescent FUW-Tomato vector with the same electroporation protocol. Two days later, Tomato-expressing cells were FACS-purified and replated at low density on MEF feeder layers. Individual colonies were picked and expanded 10–14 days after FACS sorting. The excision of the PGK-puro selection cassette was confirmed by Southern blot analysis.

Southern blot. Genomic DNA was extracted from WIBR2, WIBR2-BirA, and WIBR2-OCT4aΔFLhESCs (before and after Cre excision of the PGK-Puro selection cassette) by tail lysis buffer (10 mM Tris-Cl (pH8.0), 100 mM NaCl, 10 mM EDTA, 0.5% SDS and 0.4 mg/ml proteinase K) and incubated overnight at 37°C. The next morning an equal volume of isopropyl alcohol was added and centrifuged at 6000 x g for 5 min. The supernatant was discarded, and the pellet resuspended in 1 ml of 70% ethanol for washing and centrifuged at 6000 x g for 5 min. At this point the supernatant was carefully decanted and the pellet air dried before being resuspended in 200 μl TE buffer. Five micrograms aliquots were digested overnight with the appropriate enzymes (EcoRV for AAVS1 targeting, BamHI for OCT4 targeting) at 37°C. Digested genomic DNA was separated on a 0.8% agarose gel followed by capillary transfer to a Hybond N+ membrane (Amersham Biosciences). For making an AAVS1-specific internal probe, the donor
Fig. 7 BRG1 regulates chromatin accessibility during the exit from naïve pluripotency. a Schematic overview of CRISPR/Cas9-mediated generation of BRG1 KO naïve hESCs and representative images. n = 3 independent BRG1 KO clones were generated. Scale bar is 200 μm. b Western blot analysis for BRG1, BRM, NANOG, and ACTIN in WT (BRG1+/+), HET (BRG1+/−), and KO (BRG1−/−) naïve hESCs (three independent clones). c Flow cytometry analysis for the naïve cell-surface markers CD75 and SUSD2 in primed hESCs and naïve BRG1 KO hESCs. n = 3 independent BRG1 KO clones were analyzed. d Schematic overview of the naïve-to-primer pluripotency transition and representative images with WT, BRG1+/+, and BRG1−/− cells on day 5 of repriming. Scale bar is 200 μm. e Histogram representing the mean fluorescence intensity (MFI) of CD90 cell-surface antibody staining in BRG1 WT, HET, and KO cells on day 10 of repriming compared to naïve and primed hESCs. n = 3 independent BRG1 KO clones were analyzed. f Scatterplots showing the distribution of ATAC-seq peaks in BRG1 KO vs. WT/HET cells under naïve conditions (top) or on day 4 of repriming (bottom). Sites with significantly different ATAC intensities were determined by FDR < 0.05. g Mean ATAC-seq intensity at primed-specific (C6) and naïve-primer-shared (C8) enhancers of the BAF peaks during repriming in BRG1 KO vs. WT/HET cells. h Principal component analysis (PCA) of ATAC-seq profiles from BRG1 WT, HET, and KO cells under naïve conditions and during repriming compared to published ATAC-seq data of naïve and primed hESCs.18,53 i Volcano plot (left) showing differentially expressed genes (fold change > 2, p value < 0.05, from unpaired two-sided t-test) between naïve and primed (WT) hESCs. Pie charts (right) for expression of naïve genes with decreased expression and primed genes with increased expression on day 5 repriming, compared with day 0 of naïve cells. j Expression of the subsets of naïve genes decreased in repriming (left) and primed genes increased in repriming (right) from RNA-seq analysis of BRG1 WT/HET and KO cells under naïve, primed conditions, and during repriming. k Data are obtained from n = 3 biologically independent experiments. l Expression of BRG1-activated (N = 87) and -repressed (N = 99) genes on day 5 of repriming in naïve and primed hESCs. m RNA-seq and ATAC-seq tracks at the ZIC3 locus showing reduced expression and chromatin accessibility at enhancer and promoter regions in BRG1−/− cells during repriming. Boxplots (j, l) present the 25th, median, and 75th quartiles, and the whiskers extend to the 1.5 of interquartile ranges. P value is from two-sided Mann-Whitney test.

Fig. 8 Model of enhancer regulation by common (OCT4, BRG1) and state-specific TFs and chromatin remodelers in naïve and primed hESCs. In naïve hESCs OCT4 regulates enhancers of blastocyst-specific and stem cell maintenance genes together with the BAF ATPases BRG1 and BRM. Specificity of OCT4/BAF targeting to these naïve-specific genomic sites is conferred by the naïve TFs TAFAP2C and KLF4, which are downregulated during the naïve-to-primed transition in hESCs and in human embryos cultured through implantation stages in vitro.57 In contrast, in primed hESCs OCT4 and SOX2 control expression of pan-ectodermal lineage and stem cell maintenance genes together with the BAF ATPase BRG1. Hence, the BAF complex has overlapping and divergent roles in naïve and primed pluripotent states in human.

**Immunostaining.** Cells in 12-well plates were fixed in PBS supplemented with 4% paraformaldehyde for 20 min at room temperature. The cells were then permeabilized and blocked using 0.2% Triton X-100, 0.1% Tween-20 and 3% donkey serum in PBS for 30 min. Primary antibody against human OCT4 (1:500, Santa Cruz, sc-5279), human SOX2 (1:500, R&D Systems, AF2018), and human NANOG (1:250, goat polyclonal, R&D Systems, AF1997) was diluted in 0.2% Triton X-100 in PBS and incubated with the samples overnight at 4 °C. The cells were then washed with continuous mixing. For SA-IP, nuclear extracts were incubated with pre-optimized ANTI-FLAG M2 antibody (200 μl slurry per IP, Sigma, F2426) for 3 h at 4 °C. For OCT4 antibody IP, precleared extracts were incubated with 20 μg
OCT4 primary antibody (Santa Cruz, sc-5279) or mouse IgG (Millipore, sc-101162), BAF60A (Santa Cruz, sc-514400), BAF170 (Santa Cruz, sc-137863), OCT4 (Santa Cruz, sc-5279). SNF2L (Cell Signaling Tech., #12483), BRM and BRG1 CRISPR targeting. 

Outputs of protein identification from Proteome Discoverer were imported into a local Microsoft Access database. Duplicated records were removed by unique protein symbol. Common contamination proteins (trypsin, keratin, Actin, Tubulin, and Ribosomal Protein S1) and protein lists were identified by identified scores >10, and number of identified peptides >2. A protein needed to be identified in three out of four experiments to be considered an interactor candidate. Proteins present in more than 20% of CRAPome27 (available at http://www.crapome.org/) experiments were removed prior to calculation of empirical p-values. OCT4 interactors were chosen from the list of proteins with a combined cumulative probability (CCP) score as described previously38, with the following minor modifications to the algorithm. We multiplied the ratio of spectral counts between pull-down and control experiments by the number of spectral counts in pull-down experiments, prior to calculation of the empirical p-value. This rewards proteins with high spectral counts in pull-down experiments, which is necessary when working with a more sensitive MS platform. A false-detection-rate (FDR) of 0.1 was applied as the significance cutoff of OCT4 interactors.

Co-immunoprecipitation (co-IP) and western blot. Co-IP in regular condition was performed with the same protocol of OCT4 antibody IP–MS experiments but with a lower number of cells. Usually, confluent 15 cm diameter dishes of WIBR2 and WIBR3 cells in the primed and naive (4i/L/A or 5i/L/A) conditions were harvested for co-IP. In high-salt co-IP, the cell nuclei were harvest with the sample strainer. Cell debris was excluded by FSC vs. SSC gates and single cells were gated into 8 with buffer D-NP. Bound material was eluted by boiling for 5 min in Laemmli buffer and fractionated on a 10% SDS-PAGE. The gel lanes were horizontally cut into 8–10 pieces and each piece was subjected to digestion with porcine trypsin (Promega) as previously described39. The resulting peptides from each piece were dried down and analyzed by LC-MS/MS.

Mass spectrometry and proteomics data analysis. The samples were reconstituted in 5–10 μl of HPLC solvent A (2.5% acetonitrile, 0.1% formic acid). A nano-scale reverse-phase HPLC capillary column was created by packing 5 μm C18 spherical silica beads into a fused silica capillary (100 μm inner diameter x 12 cm length) with a flame-draw tip. After equilibrating the column each sample was loaded onto the column. A gradient of acetonitrile from 2.5 to 97.5% was used to elute peptides and then protein G agarose beads eluted, subjected to electron transfer dissociation, and then they entered into an LTQ-Orbitrap–Velo mass spectrometer (Thermo Finnigan) with collision-induced dissociation (CID). Eluting peptide were detected, isolated, and fragmented to produce a tandem mass spectrum of specific fragment ions for each peptide. MS data were processed by Thermo Proteome Discoverer software (available at https://www.uniprot.org/).

 Knockdown by lentiviral shRNAs. VSV–glycoprotein pseudotyped lentiviral vector particles were produced by co-transfection of HEK293T cells with packaging and envelope plasmids as previously described47. The plK01-puro (SIGMA) plasmids targeting BRM (TRCN0000358828, TRCN0000020329, TRCN0000367681), BRG1 (TRCN0000155499, TRCN000015550) and two Non–Mammalian shRNA Control Plasmid (SIGMA, #SHC002 and #SHC016) were used. Viral supernatants were harvested after 48 h and filtered through a 0.45 μm membrane. WIBR3 hESCs maintained in primed hESC medium and WIBR3 naïve hESCs maintained in 5i/L/A were infected with the lentivirus in the presence of 6 μg/ml polybrene (Millipore). Infected cells were plated on mitomycin C–inactivated DR4 feeder cells and selected with 0.5 μg/ml of puromycin starting 2 days after transduction. Total RNA was extracted after 5 days of puromycin selection.

Quantitative real-time PCR. To assess the expression of Bira transgene following AAVS1 targeting in primed hESCs, total RNA was isolated using the Rneasy Kit (QIAGEN) and reversed transcribed using the Superscript III First Strand Synthesis Kit (Invitrogen). Quantitative RT–PCR analysis was performed in triplicate using the ABI 7900 HT system with FAST SYBR Green Master Mix (Applied Biosystems). Gene expression was normalized to GAPDH. Error bars represent the standard deviation (SD) of the mean of triplicate reactions. Primer sequences are included in Supplementary Data 6.

RNA-seq analysis. RNA-seq analysis was performed in naive and primed hESCs in which the BAF ATPases were generally perturbed by either shRNA-mediated knockdown of WIBR3 cells and CRISPR/Cas9-mediated knockdown of H9 cells. Total RNA was extracted using E.Z.N.A. Total RNA kit (Omega Bio-Tek, R6834) with DNase I following the manufacturer’s instructions. Samples were prepared according to the library kit manufacturer’s protocol, indexed, pooled, and sequenced on an Illumina HiSeq at the Genome Technology Access Center at the University of Illinois at Urbana–Champaign.

Flow cytometry analysis. Cells were single cell dissociated using TrypLE Express and washed once in FACS buffer [PBS supplemented with 5% FBS]. The cells were then resuspended in 100 μl fresh FACS buffer, and incubated with antibodies for 30 min on ice. The following naïve-specific cell-surface antibodies were used: anti–SUSD2-PE, 1:100 (BioLegend, 327406) and anti–CD75-eFluor 660, 1:100 (Thermo Fisher, 20-0759-42). A positive staining control for the human generation antibody line was used. Flow cytometry analysis was performed using a BD LSFortessa X-20 and the data were analyzed using the FlowJo software.
genome, as well as the transcript annotation, was downloaded from the iGenomes site. The aligned bam files were sorted. Transcript assembly and differential expression analyses were performed using the Cuffdiff program. In the gene expression matrix value of RPKM +0.1 was applied to the samples of single-end data and a value of RPKM +1 was applied to the samples of paired-end data, to minimize the effect of low-expression genes. P values were calculated using a t-test. Differentially expressed genes (DEGs) were identified by two-sided FDR (false discovery rate) < 0.05 and fold change > 2. Boxplots for expression were generated using R. P value was calculated from two-sided Mann–Whitney test.

Transcriptome analyses of human and non-human primate embryos. Figures 4 and 5c: Processed single-cell RNA-seq data from 3D-cultured human pre-gastrulation embryos (GSE136447) were downloaded. Genes with very low expression (average(log(FPKM + 1)) < 1) potentially due to low capture rates over all replicates and genes with no reads were removed. Plotting the expression of selected markers by color and the percentage of epiblast cells that express the marker by point size was generated by using the ‘DotPlot’ function of the R Seurat package. Expression of C1 and C6 target genes in human embryos were obtained from CUT&Tag annotation of the BAF peaks (refer to ChiP-seq analysis, Supplementary Data 3). Boxplots were generated using R.

Supplementary Figure 5c: To examine the expression of C1 and C6 target genes (Supplementary Data 3) in non-human primate embryos, single pre- and post-implantation cynomolgus monkey gene expression data (GSE74767) were acquired and filtered for mapped unique gene IDs and valid numeric expression values. Using a one-to-one gene ID lookup table (Nakamura et al. 2016, Supplementary Data 2), orthologs of expression target genes (from C1 and C6) filtered for mapped unique gene IDs and valid numeric expression values. Using a one-to-one gene ID lookup table (Nakamura et al. 2016, Supplementary Data 2), orthologs of expression target genes (from C1 and C6) filtered for mapped unique gene IDs and valid numeric expression values. Using a one-to-one gene ID lookup table (Nakamura et al. 2016, Supplementary Data 2), orthologs of expression target genes (from C1 and C6) filtered for mapped unique gene IDs and valid numeric expression values. Using a one-to-one gene ID lookup table (Nakamura et al. 2016, Supplementary Data 2), orthologs of expression target genes (from C1 and C6) filtered for mapped unique gene IDs and valid numeric expression values. Using a one-to-one gene ID lookup table (Nakamura et al. 2016, Supplementary Data 2), orthologs of expression target genes (from C1 and C6) filtered for mapped unique gene IDs and valid numeric expression values. Using a one-to-one gene ID lookup table (Nakamura et al. 2016, Supplementary Data 2), orthologs of expression target genes (from C1 and C6) filtered for mapped unique gene IDs and valid numeric expression values.
Data availability

The data that support this study are available from the corresponding authors upon reasonable request. The OCT4 affinity purification mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD026556. The ATAC-seq, ChIP-seq, CUT&Tag, and RNA-seq data generated in this study have been deposited in the Gene Expression Omnibus (GEO) database (GEO accession codes GSE47751 and GSE16802). The ATAC-seq and TFAP2C ChIP-seq data (GSE101074), OCT4 and H3K27ac ChIP-seq data (GSE96947), and H3K4me3 and H3K27me3 ChIP-seq data (GSE59435) in naive and primed hESCs were analyzed in this study. The ATAC-seq and RNA-seq data of human embryos (GSE101571), and ATAC-seq and ChIP-seq data (GSE122631) of human NPCs, and Sox2 ChIP-seq data (GSE80479) in human NPCs were analyzed in this study. Source data are provided with this paper.

Received: 19 April 2021; Accepted: 12 July 2021;
Published online: 26 August 2021

References

74. Bredenkamp, N. et al. Wnt inhibition facilitates RNA-mediated transcription by BRG1 and BRM, two mutually exclusive SWI/SNF ATPase subunits. 


Acknowledgements

This work was supported by the National Institutes of Health (NIH) Director’s New Innovator Award (DP2 GM137418) and a grant from the Children’s Discovery Institute (CDI-LI-2019-819) to T.W.T. and grants from the NIH (R01GM127957, R01HD095538, and R01HD197268) and New York State (NYSTEM) (C32538G and C32569G) to J.W. S.J. was funded by the Suanaree University of Technology under the National Research University (NRU) project of Thailand, Office of the Higher Education Commission. We thank Frank soldier and Haoyi Wang for advice on gene targeting. Qing Gao for assistance with teratoma analysis, Angela Bowman and Liannna Solniha-Kreutz for critical reading of the manuscript, the Genome Engineering and iPSC Center (GEiC) at the Washington University School of Medicine (WUSM) for gRNA validation services, and the Genome Technology Access Center (GTAC) in the Department of Genetics at WUSM for help with genome sequencing. GTAC is partially supported by NCI Cancer Center Support Grant #P30 CA91842 to the Siteman Cancer Center and by ICTS/CTSA Grant# UL1 TR000448 from the National Center for Research Resources (NCRR), a component of the NIH, and NIH Roadmap for Medical Research. This publication is solely the responsibility of the authors and does not necessarily represent the official view of NCRR or NIH. All experiments involving hESC were approved by the Institutional Biological and Chemical Safety Committee and Embryonic Stem Cell Research Oversight Committee at Washington University School of Medicine and Columbia University Irving Medical Center.

Author contributions

T.W.T. and X.H. conceived, designed and conducted the studies, and wrote the manuscript; K.P. performed functional studies of BAF subunits and interpreted the results; J.P. and Z.M. performed ChIP-seq experiments under direction of C.K.; P.V.S. analyzed MS data together with X.H.; I.A.F., C.D., J.Y., D.L., J.D., T.L., M.M., S.A.K. and S.I. contributed to experiments; X.H., P.G., B.Z., S.D. and N.J. performed bioinformatics analysis. R.J. designed experiments, interpreted the results, and approved the manuscript; J.W. conceived the project, designed the experiments, interpreted the results, prepared and approved the manuscript. X.H. and K.P. contributed equally as joint first authors. J.W. and T.W.T. are co-corresponding authors.

Competing interests

R.J. is a co-founder of Fate Therapeutics, Fulcrum Therapeutics, and Omega Therapeutics, and an advisor to Development Therapeutics. C.K. is the Scientific Founder, fiduciary Board of Directors member, Scientific Advisory Board member, shareholder, and consultant for Foghorn Therapeutics. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-25107-3.

Correspondence and requests for materials should be addressed to J.W. or T.W.T.

Peer review information Nature Communications thanks Alvaro Rada-Iglesias, Gerald Crabtree and the other, anonymous, reviewer for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s), 2021

Nature Communications | (2021) 12:5123 | https://doi.org/10.1038/s41467-021-25107-3 | www.nature.com/naturecommunications