VGLL1 contributes to both the transcriptome and epigenome of the developing trophoblast compartment

Ruben I. Calderon[1,2,], Nirvay Sah[1], Molly Huang[3], Sampada Kallol[1], Ryan H. Kittle[1,7], Walee B. Shaik[1], Ahmed Abdelbaki[4,13], Jennifer N. Chousal[1], Robert Morey[1], Tony Bui[1], Alejandra Mitre[1], Norah M.E. Fogarty[5,8], Claudia Gerri[5,9,10], Zoe Manalo[1], Claire Zheng[1,2], Peter De Hoff[3], Pratik Home[6,11], Kathy K. Niakan[5,12,13,14], Heidi Cook-Andersen[2,3], Kathleen M. Fisch[3], Soumen Paul[6], Francesca Soncin[1,*]

Affiliations:

1. Department of Pathology, University of California San Diego, La Jolla, CA 92093, USA
2. School of Biological Sciences, University of California San Diego, La Jolla, CA 92093, USA
3. Department of Obstetrics, Gynecology, and Reproductive Sciences, School of Medicine, University of California San Diego, La Jolla, CA, 92093, USA.
4. Department of Zoology, Faculty of Science, Zagazig University, Zagazig 44519, Egypt.
5. Human Embryo and Stem Cell Laboratory, The Francis Crick Institute, London, NW1 1AT, UK.
6. Department of Pathology & Laboratory Medicine and Institute for Reproductive and Developmental Sciences, University of Kansa Medical Center, Kansas City, KS 66160.

Current addresses:

7. Department of Genetics and Development, Columbia University Irving Medical Center, New York, NY 10032, USA.
8. Centre for Gene Therapy and Regenerative Medicine, King's College London, London, SE1 9RT, UK.
9. Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstrasse 108, 01307, Dresden, Germany.
10. Center for Systems Biology Dresden, Pfotenhauerstrasse 108, 01307, Dresden, Germany.
11. BioIVT, 1101 W Cambridge Circle Dr. Kansas City, KS 66103.
12. The Loke Centre for Trophoblast Research, Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3EG, UK.
13. Wellcome Trust – Medical Research Council Stem Cell Institute, University of Cambridge, Jeffrey Cheah Biomedical Centre, Puddicombe Way, Cambridge CB2 0AW, UK.
14. Epigenetics Programme, Babraham Institute, Cambridge CB22 3AT, UK.

* Lead contributor and Corresponding author: fsoncin@health.ucsd.edu

Conflict of Interest statement

All authors have no conflict of interest to declare.

Classification

Biological Sciences/Developmental Biology

Keywords

trophectoderm, trophoblast stem cells, pluripotent stem cells, VGLL1

Abstract

The trophectoderm (TE), the first lineage specified during mammalian development, initiates implantation and gives rise to placental trophoblasts. While animal models have elucidated key conserved signaling pathways involved in early TE specification, including BMP, WNT, and HIPPO, species-specific differences during early development emphasize the need for human-specific models. We previously identified VGLL1, a coactivator of TEAD transcription factors, as a human-specific placental marker. In this study, we employed a pluripotent stem cell (PSC)-based model of TE induction by BMP4 to investigate chromatin remodeling and transcriptional dynamics during TE formation. BMP4-induced chromatin accessibility changes promoted a trophoblast gene expression program, while mesoderm lineage markers were only transiently expressed upon canonical WNT activation. We found that VGLL1 was expressed downstream of key TE transcription factors (GATA2/3, TFAP2A/C) but was essential for establishment of full trophoblast identity by up-regulating EGFR and reinforcing GATA3 expression through positive feedback. Notably, VGLL1 enhanced canonical WNT signaling via direct regulation of WNT receptors and effectors. We also identified KDM6B, a histone demethylase that removes H3K27me3 repressive marks, as a direct VGLL1 target. KDM6B facilitated activation of bivalent promoters associated with TE markers, linking epigenetic regulation to lineage identity. Our findings establish a mechanistic framework positioning VGLL1 as a central regulator that integrates HIPPO, BMP, and WNT signaling pathways to drive establishment of human TE.

Statement of Significance

Early development of the human placenta is essential for pregnancy success, yet the mechanisms that guide placental lineage specification remain poorly defined. Using human stem cells, we show how signaling pathways and chromatin remodeling programs work together to direct formation of the trophectoderm, the earliest placental cell type. We identify VGLL1 as a key regulator linking multiple signaling networks to gene expression and epigenetic control. Our findings reveal a species-specific mechanism of placental initiation with broad implications for understanding reproductive disorders, pregnancy loss, and advancing stem cell–based models to study and potentially treat human placental disease.

Introduction

The trophectoderm (TE) is the first lineage specified during mammalian development and is important for embryo cavitation, implantation, and pregnancy initiation (1). After implantation, the TE gives rise to the trophoblast compartment of the placenta, contributing to the formation of the maternal/fetal interface (2). Early pregnancy loss and several placental-associated disorders are linked to abnormalities in the establishment and differentiation of the TE compartment during early development (2). However, ethical and technical considerations have limited our understanding of the mechanisms of TE specification and trophoblast development in humans. Studies in animal models have determined that tight spatio-temporal regulation of highly conserved signaling pathways, including BMP, WNT, and HIPPO, are required to establish the TE in the pre-implantation embryos and for early placental development (3–8). However, increasing evidence points to species-specific differences in the composition of regulatory networks that contribute to the development of the extra-embryonic ectoderm compartment (3, 9–11). Therefore, species-specific models are better suited to investigate TE specification and early placental development (3, 9, 12).

The HIPPO signaling pathway is a conserved cascade that initiates the specification of the TE lineage following activation of GATA3 and CDX2 expression by the TEAD4/YAP1 complex (10, 13, 14). In early mouse embryos, differential HIPPO pathway activation between the inner and outer cells leads to the expression of CDX2, which initiates a TE expression program, including induction of GATA3 and ELF5 (13, 15, 16). EOMES is expressed down-stream of CDX2 and, while not required for TE specification, it is fundamental for the maintenance of the TE transcriptional program and maintenance of mouse trophoblast stem cells (TSCs) (17, 18). In humans, four transcription factors, GATA2, GATA3, TFAP2A, and TFAP2C, form a transcriptional network required for TE initiation (TEtra factors) downstream of the HIPPO signaling pathway (19). This network drives the establishment of an extra-embryonic ectoderm signature, which includes mature markers such as VGLL1, TP63, and EGFR, but not EOMES, which is dispensable for human TE and trophoblast development (9, 20). In particular, we have previously identified VGLL1 as a human-specific marker

expressed in the trophoblast progenitor compartments of the placenta (9). VGLL1 is a conserved binding partner of the TEAD family of transcription factors and competes with YAP1 to form a transcriptional complex with TEAD4, therefore potentially modulating HIPPO signaling in human placenta (21).

Canonical WNT signaling is a well-characterized developmental pathway mediated by nuclear localization of β-catenin, where it binds to TCF/LEF complexes and drives expression of WNT-specific target genes (22). This pathway is required for maintenance and expansion of various human epithelial stem cells, including intestinal and skin stem cells (23, 24). In human pre-implantation embryos, the TE compartment is enriched in WNT components, including ligands and receptors (25–29). Moreover, compared to mouse, both human and non-human primate embryos show higher canonical WNT-related gene expression in the TE (30). WNT genes are also highly expressed in the trophoblast compartment of the placenta. *In vitro*, both human TSCs and trophoblast organoids require activation of canonical WNT signaling for proliferation (31–33) while its transient inhibition is required to initiate differentiation of extravillous trophoblasts (EVT) (32, 34).

Bone Morphogenetic Protein (BMP) signaling has well described roles in the embryonic compartment during early development, including specification of the primitive endoderm from NANOG+ cells, the embryonic mesoderm lineage during gastrulation, and the primordial germ cells (5, 35). However, recent evidence demonstrated BMP activation in the TE of human, but not murine, blastocysts (5, 12), though autocrine BMP signaling has been reported in both mouse and human TSCs (36). *In vitro*, exogenous BMP4 has been shown to drive differentiation of human pluripotent stem cells (PSCs) into several lineages, namely embryonic mesoderm, primordial germ cells, or extra-embryonic trophoblast lineage progenitors, based on the cellular context and the activation of other signaling pathways (5). In particular, embryonic mesoderm differentiation requires BMP-induced WNT activation, while inhibition of β-catenin-dependent WNT signaling downstream of BMP effectively prevents mesoderm differentiation and promotes the trophoblast lineage (37). However, the mechanisms underlying WNT activation by BMP4 and the switch from mesoderm to trophoblast lineage in PSC-derived TE-like cells have not been elucidated yet.

127 Regulatory networks drive establishment and maintenance of cellular identity by
128 acting on a lineage-specific epigenetic landscape downstream of signaling pathways (38).
129 DNA methylation and histone modifications surrounding lineage-specific genes determine
130 chromatin accessibility for specific transcription factor (TF) binding (38). Recent studies in
131 pre-implantation embryos have described epigenome changes across early development
132 from zygote to blastocyst, including segregation of TE and ICM (39–41). However, there is
133 little mechanistic understanding on how epigenetic marks are laid out in a lineage-specific
134 manner during trophoblast development. Moreover, it is largely unknown how multiple
135 signaling pathways like HIPPO, WNT, and BMP converge towards a lineage-specific outcome
136 and drive trophoblast development.

137 PSCs represent a unique resource to study species-specific genetic and epigenetic
138 mechanisms of TE and trophoblast specification as they are tractable models for genetic
139 and chemical manipulation. Multiple protocols have been published to derive TE from
140 human PSCs (42–46). We have established a 2-step protocol based on BMP4 treatment of
141 PSCs to first convert them into TE-like cells, followed by either further differentiation into
142 mature trophoblasts or derivation of *bona fide* hTSCs (47, 48). Based on prior literature (37),
143 we further optimized the first step by adding a WNT inhibitor, IWP2, to block concurrent
144 activation of a mesoderm program and obtain a more homogeneous population of TE-like
145 cells (37, 45). We previously reported that VGLL1 is one of the most highly upregulated genes
146 in the BMP4-directed conversion of human embryonic stem cells (hESCs) into TE-like cells
147 (9). Here, we use this model to first investigate the BMP4-dependent chromatin
148 modifications that occur during the transition from PSC to TE-like cells in relation to gene
149 expression changes and WNT activation status. Then, we present our findings on how VGLL1
150 directly and indirectly drives a trophoblast-specific gene expression and epigenomic
151 program in this PSC-derived trophoblast model.

152 Results

153 BMP4 mediates chromatin accessibility to stabilize a trophoblast-specific
154 expression program

155 We have previously optimized a protocol for lineage conversion of PSCs into TE-like
156 cells by treatment with 10ng/mL BMP4 and the WNT inhibitor IWP2 over 4 days, which
157 prevents mesoderm differentiation and promotes a uniform EGFR+ population (45). To start
158 dissecting the transcriptional changes towards TE-like cells during this conversion protocol,
159 we first performed bulk RNA-seq analysis of H9 hESCs at day 0, day 2, day 4 following
160 treatment with BMP4 or BMP4/IWP2 (**Suppl. Fig. S1A**) and identified differentially expressed
161 genes for each timepoint and treatment compared to d0 (**Suppl. Table S1**). Genes
162 downregulated in both treatments included pluripotency and embryonic lineage markers,
163 i.e. *NANOG*, *POU5F1*, and *SOX2* (**Suppl. Table S1**). Upregulated genes common across
164 timepoints and treatments included the TEtra factors *GATA2*, *GATA3*, *TFAP2A*, and *TFAP2C*.
165 By day 4, common upregulated genes included *ENPEP*, and trophoblast markers *VGLL1* and
166 *TP63 (***Suppl. Fig. S1B and Suppl. Table S1**). Genes uniquely upregulated by BMP4
167 treatment on day 2 were associated with various signaling pathways, including WNT
168 signaling (**Fig. 1A, Suppl. Table S1**) and, as expected, we detected upregulation of WNT-
169 dependent mesoderm markers *EOMES*, *TBXT*, and *TBX6* (**Fig. 1B**). However, by day 4, the
170 expression of these markers decreased to levels comparable to the treatment with the
171 addition of IWP2 (**Fig. 1B**). Accordingly, WNT signaling was no longer one of the top
172 significantly enriched pathways at day 4 of BMP4 alone treatment (**Suppl. Fig. S1C**). To
173 confirm the activation timeframe of the WNT pathway during BMP4 treatment, we used a
174 TCF/LEF-eGFP reporter hESC line (HUES9) where detection of GFP fluorescence is
175 dependent on canonical WNT activation (49). We detected WNT signaling activation starting
176 at day 2, a peak at day 3, and decreased activity at day 4 of BMP4 treatment (**Suppl. Fig.**
177 **S1D**). The WNT inhibitor IWP2 completely blocked eGFP expression, while another inhibitor,
178 IWR1, showed only partial suppression of WNT activity at the dose used (**Suppl. Fig. S1D**).
179 Accordingly, we detected a dose-dependent inhibition of both WNT targets, *LEF1* and
180 *WNT5B*, as well as mesoderm markers *TBXT* and *EOMES* between the two inhibitors (**Suppl.**
181 **Fig. S1E**). Finally, we compared our data to the human embryo reference created by Zhao *et*
182 *al.* (50) (**Suppl. Fig. S1F-G**). As expected, undifferentiated hESCs at d0 were annotated as
183 Epiblast (**Suppl. Fig. S1F-G** and **Suppl. Table S2**). Although treatment with BMP4 and

184 BMP4/IWP2 for 2 days did not sufficiently modify the transcriptome, the annotation predictor
185 scores were lower than for hESCs at d0 (**Suppl. Fig. S1F-G** and **Suppl. Table S2**).
186 BMP4/IWP2-treated cells at d4 were annotated as TE but only one of the BMP4 d4 replicates
187 was also identified as TE, albeit with a lower predictor score, while the other replicates
188 remained ambiguous and mapped close to the hypoblast (**Suppl. Fig. S1F-G** and **Suppl.**
189 **Table S2**). Overall, these data suggest that low dose of BMP4 transiently activates WNT
190 signaling during treatment of hESCs, but that both WNT activity and mesoderm gene
191 expression decrease over time as the TE identity is established in this model.

192 To investigate chromatin accessibility changes associated with the differential
193 establishment of the mesoderm and TE lineage during BMP4 treatment of PSCs, we
194 performed Assay for Transposase-Accessible Chromatin (ATAC) followed by sequencing in
195 hESCs treated with BMP4 or BMP4/IWP2 at days 0, 2, and 4. We found that peaks resided at
196 promoter (31%), intergenic (24%), and intronic (23%) regions (**Suppl. Fig. S1H**). Following
197 differential peak analysis, both at day 2 and day 4 of treatment, we found open chromatin
198 regions specific to BMP4 or BMP4/IWP2, as well as regions common to both treatments (**Fig.**
199 **1C, Suppl. Table S3**). We performed Hypergeometric Optimization of Motif EnRichment
200 (HOMER) analysis on common and day/treatment-specific regions to identify accessible TF
201 binding motifs. Common regions at both days 2 and 4 contained binding motifs for well-
202 known TE and trophoblast TFs, including TEAD and AP1-complex members (i.e. JUN and
203 FOS) (**Suppl. Fig S1I**). Binding motifs of both TEtra factors GATA3 and TFAP2C appeared in
204 both common and treatment-specific regions. At day 2 of BMP4-only treatment, we found
205 open chromatin at LEF-1 and TCF3, TFs downstream of the WNT pathway. We then looked at
206 the chromatin regions around the transcription start sites of well-known trophoblast and
207 mesoderm genes (**Fig.1D, Suppl. Fig. S1J**). The open chromatin region across
208 treatments/days showed a sharp peak and remained consistent around trophoblast markers
209 (**Fig. 1D**), including GATA3 and EGFR (**Suppl. Fig. S1J**). Conversely, open chromatin around
210 mesoderm genes showed a broader distribution that decreased between day 2 and day 4 in
211 response to both treatments (**Fig. 1D**), including at EOMES and TBXT (**Suppl. Fig. S1J**) These

212 chromatin accessibility data are largely consistent with the gene expression patterns
213 observed above.

214 Finally, we looked at the differential expression of known epigenetic modifiers that
215 could account for these changes in chromatin accessibility. We found multiple lineage-
216 specific chromatin modifiers differentially expressed during TE-specification. These
217 included down-regulation of *de novo* DNA methyltransferase *DNMT3A* at day 2, a switch
218 from *TET1* to *TET2* (51) expression, and up-regulation of histone demethylases like *KDM6B*,
219 *KDM4B,* and *KDM1B* (**Fig. 1E**). These data suggest that low levels of BMP4 preferentially
220 enable a trophoblast transcriptional signature via a combination of lineage-specific gene
221 expression activation and finely tuned regulation of the epigenetic landscape.

222 VGLL1 regulates a TE-specific gene expression program

223 In the context of BMP4 treatment of human PSCs, the TEtra factors GATA2, GATA3,
224 TFAP2A, and TFAP2C have been identified as initiators of the TE/trophoblast transcriptional
225 program, upstream of more mature trophoblast markers like TP63, EGFR, ELF5, and
226 VGLL1(9, 19, 20). To further tease apart the timeframe of activation of these TE/trophoblast
227 genes in this BMP4-based model, we investigated the expression of TEtra factors, and
228 mature trophoblast markers by qPCR at short intervals within the first 72 hours of
229 BMP4/IWP2 treatment. As previously reported, TEtra factors *GATA3* and *TFAP2C* were the
230 first to be activated, with significant upregulation as early as 6h after the start of the
231 treatment and continuous increase over time (**Fig. 2A**). *VGLL1* expression started to
232 significantly rise at 24 hours of BMP4/IWP2 treatment (**Fig. 2A**) while other mature
233 TE/trophoblast markers were observed at 48 hours (*TP63*) and 72 hours (*ELF5* and *EGFR*),
234 respectively (**Fig. 2A**). By immunofluorescence, VGLL1 protein was detected in the nucleus
235 after 24 hours of BMP4/IWP2 treatment, where it co-localized with TEAD4 and YAP1
236 throughout TE induction (**Suppl. Fig. S2A**). While previous literature considered VGLL1 a
237 mature trophoblast marker (like TP63), these temporal activation data suggest that VGLL1
238 might play an early role in TE induction downstream of TEtra factors.

239 To understand the role of VGLL1 in this context, we established hESC lines that
240 constitutively express VGLL1-specific shRNA for targeted gene knock-down (or shScramble
241 as control, hereafter referred to as Scramble). Compared to Scramble hESCs, treatment of
242 shVGLL1 hESCs with BMP4/IWP2 led to blunted up-regulation of late trophoblast markers,
243 *TP63*, *ASCL2*, and *GCM1* (**Fig. 2B**). Interestingly, TEtra factor *GATA3* and the early TE marker
244 *CDX2* also showed lower up-regulation. In addition, surface levels of EGFR levels were also
245 lower in the BMP4/IWP2-treated shVGLL1-hESCs, compared to similarly-treated Scramble-
246 hESCs (**Suppl. Fig. S2B**), suggesting an incomplete TE conversion. To further explore the
247 effect of VGLL1 knock-down in TE induction, we performed bulk RNA-seq analysis on two
248 independent shVGLL1 clones and one Scramble control at days 0 and 4 of BMP4/IWP2
249 treatment (**Suppl. Fig. S2C**). Overall, shVGLL1 hESC clones showed a marked delay in
250 lineage conversion compared to Scramble hESCs (**Suppl. Fig. S2C**). We found 7,284 genes
251 that were differentially dysregulated in BMP4/IWP2-treated shVGLL1 hESCs compared to
252 Scramble hESCs (adj. p value<0.05), organized in 4 distinct clusters based on pattern of
253 expression (**Suppl. Table S4**). Specifically, 3,056 genes showed blunted down-regulation
254 (cluster #2) compared to Scramble and were associated with transcriptional regulation of
255 pluripotent stem cells and embryo development (**Suppl. Fig. S2D-E**), including pluripotency
256 markers *POU5F1*, *SOX2*, and *SALL2*. The 4,192 up-regulated genes that showed
257 misregulation in shVGLL1 hESCs compared to Scramble hESCs (**Fig. 2C**) organized in 3
258 clusters of expression (**Suppl. Fig. S2D-E**). Cluster #3 contained genes that failed to be up-
259 regulated, including TE/trophoblast markers *GATA2*, *NR2F2* (52), and *CCR7* (53), while
260 cluster #4 included genes with blunted expression compared to Scramble, including TEtra
261 factors *GATA3* and *TFAP2C*, TE markers *ENPEP* (54), and trophoblast markers *EGFR*, *ARID3A*,
262 and *ARID5B* (55). Gene ontology analysis indicated that genes that failed to be fully up-
263 regulated were associated with response to hormones and cell migration and were highly
264 enriched for placental genes (**Suppl. Fig. S2F**). Interestingly, the remaining cluster #1
265 contained genes that were more highly up-regulated in shVGLL1 hESCs compared to
266 Scramble hESCs and included TFAP2A (**Suppl. Fig. S2D-E**), a potential compensatory
267 response to lack of VGLL1 and lower expression of other TEtra factors. Finally, we compared

268 these cells to the human embryo reference in Zhao *et al*. (50) (**Suppl. Fig. S2G**). As expected,
269 both Scramble and shVGLL1 undifferentiated hESCs (d0) were annotated as Epiblast
270 (**Suppl. Fig. S2G** and **Suppl. Table S2**). BMP4/IWP2 treated cells were annotated as TE;
271 however, Scramble and shVGLL1 cells mapped to different areas of the reference dataset,
272 with Scramble representing more mature TE compared to shVGLL1 cells (**Suppl. Fig. S2G**
273 and **Suppl. Table S2**). These data suggest VGLL1 acts through a positive feedback
274 mechanism downstream of the TEtra factors to complete a TE/trophoblast-specific
275 program.

276 We then investigated if VGLL1 was required for TSC derivation. Recently, we
277 demonstrated that hESC-derived TE-like cells (hESCs following 4 days of BMP4/IWP2
278 treatment) can adapt to hTSC culture conditions, resulting in cells that show characteristics
279 of *bona fide* primary hTSCs (48). hTSC culture medium includes a canonical WNT agonist
280 and EGF signaling activation combined with TGFβ and ROCK signaling inhibition. Therefore,
281 we attempted to derive hTSC-like cells from both Scramble and two shVGLL1 hESC clones.
282 Similar to our initial report with wild type PSCs, both Scramble and shVGLL1 hESCs rapidly
283 lost both epithelial morphology and cell surface expression of EGFR during the early stages
284 of adaptation (**Suppl. Fig. S3A-B**). However, while Scramble cells acquired a hTSC-like
285 colony morphology and regained expression of surface EGFR by passage 6, shVGLL1 clones
286 maintained a single cell morphology and failed to regain EGFR expression (**Fig. 2D-E**).
287 Moreover, shVGLL1 clones expressed lower levels of TEtra factors *GATA2*, *GATA3*, and
288 *TFAP2C*, compared to Scramble hTSCs (**Fig. 2F, Suppl. Fig. S3D**). Interestingly, while the TE
289 marker *ENPEP* also decreased in shVGLL1 cells, *CDX2* increased compared to Scramble
290 (**Suppl. Fig. S3C-D**). We also tested the ability of TSC-adapted shVGLL1 cells to differentiate
291 into STB and EVT following established protocols (31, 48). Unlike Scramble TSCs, TSC-
292 adapted shVGLL1 cells failed to fuse following STB differentiation and did not secrete hCG
293 into the media (**Suppl. Fig. S3E-F**). In EVT differentiation media, TSC-adapted shVGLL1 cells
294 showed no morphological changes and failed to up-regulate HLA-G (**Suppl. Fig. S3G-H**).
295 These data suggest that VGLL1 knock-down PSCs failed to form *bona fide* TSCs and that

296 VGLL1 is required for the maintenance of a trophoblast signature, in particular related to the
297 expression of some TEtra factors and EGFR.

298 We further evaluated VGLL1 activity, uncoupled from BMP4 treatment, to investigate
299 the function of VGLL1 in promoting the TE-specific transcriptional program. We leveraged a
300 system composed of a doxycycline (DOX)-inducible dCas9-VPR and VGLL1-specific guide
301 RNAs for CRISPR-mediated gene activation in undifferentiated hESCs (iVGLL1a) (**Fig. 3A**). In
302 the absence of exogenous BMP4, ectopic expression of VGLL1 in hESCs initiated changes in
303 cell morphology towards a cobblestone-like shape (**Fig. 3B**), and increased surface
304 expression of EGFR (**Fig. 3C**). Moreover, we observed up-regulation of the TEtra factor *GATA3*
305 and trophoblast marker *EGFR* accompanied by down-regulation of pluripotency marker
306 *NANOG* by qPCR (**Fig. 3D**). We then performed bulk RNA-seq analysis on three independent
307 iVGLL1a clones treated with DOX (or DMSO as vehicle control) (**Suppl. Fig. S4A**). We found
308 6,530 differentially expressed genes (DEGs) in DOX-treated cells compared to DMSO (adj. p-
309 value 0.05), of which 3,053 were up-regulated and 3,477 down-regulated (**Fig. 3E, Suppl.**
310 **Table S5**). Significantly down-regulated genes were enriched in pluripotency genes,
311 including *MYC*, *NANOG*, *POU5F1*, and *SOX2* (**Fig. 3E and Suppl. Fig. S4B**). Up-regulated
312 DEGs included TE/trophoblast markers *ARID5B*, *EGFR*, and *GATA3* (**Fig. 3E**). Moreover, gene
313 ontology analysis of pathways and biological processes revealed the activation of signaling
314 pathways important for TE and trophoblast specification and function, including TGFβ,
315 MAPK, and PI3K-Akt pathways (**Suppl. Fig. S4C**). Interestingly, DOX-treated iVGLL1a cells
316 did not transform into TSCs following culture in TSC media: their morphology quickly
317 reverted to compact cells and were not distinguishable from DMSO-treated cells in TSC
318 media (**Suppl. Fig. S4D**). DOX-treated cells lost expression of VGLL1 when passaged into
319 TSC media without DOX induction (**Suppl. Fig. S4E**). They also lost expression of the
320 trophoblast marker TP63 as determined by qPCR (**Suppl. Fig. S4E**) and cell surface
321 expression of EGFR as determined by flow cytometry. Interestingly, both DMSO and DOX-
322 treated cells in TSC media showed low expression levels of the pluripotency marker *NANOG*
323 (**Suppl. Fig. S4E**) without showing signs of differentiation into any other embryonic lineage.
324 These data point to a specific role of VGLL1 in controlling a TE/trophoblast transcriptional

program but demonstrated that ectopic expression of VGLL1 alone is insufficient for TSC establishment from hPSCs in this model.

Unlike primed hPSCs, which requires BMP4, naïve hPSCs can be readily converted into TSCs following culture into TSC media (56). Therefore, we also tested the effect of ectopic VGLL1 expression in naïve hESCs on activation of TE-specific genes. We converted primed iVGLL1a hESCs to the naïve state in PXGL media (**Suppl. Fig. S4F-G**) (43). Interestingly, we observed higher basal expression of some TE/trophoblast markers in naïve hESCs compared to the primed state, including *VGLL1* and *GATA3* (**Suppl. Fig. S4G**). Similar to what we observed in primed hESCs, treatment with DOX and consequent VGLL1 up-regulation caused up-regulation of TE markers *GATA3* and *EGFR* as well cell surface expression of EGFR (**Suppl. Fig. S4H-I**). These data demonstrate that the program initiated by VGLL1 is conserved independently of pluripotency state.

VGLL1 regulates canonical WNT pathway effectors

Gene ontology analysis of up-regulated genes in response to DOX treatment of iVGLL1a hESCs suggested activation of WNT signaling pathway (**Suppl. Fig. S4C**). Indeed, VGLL1 expression induced up-regulation of WNT pathway components and mediators, including *EPHA2*, *FZD2*, *LEF1*, *LRP4*, *WLS*, *WNT4/5A/5B/6/10A/11*, as well as WNT signaling targets like *FST* (**Fig. 4A and Suppl. Fig. S4J**). We also confirmed β-catenin dependent WNT activation using a reporter expressing mCherry under constitutive promoter and GFP under TCF/LEF promoter (57). We observed GFP signal in a subset of mCherry expressing cells following DOX treatment compared to DMSO control (**Fig. 4B**). Interestingly, we observed up-regulation of mesoderm markers on DOX-treated iVGLL1a hESCs by qPCR, including *EOMES* and *TBXT* (**Fig. 4C**). During BMP4-induced hESC differentiation, co-treatment with WNT inhibitor IWP2 effectively suppressed mesoderm marker expression. Therefore, we tested the dependence of mesoderm gene expression down-stream of VGLL1 activation in iVGLL1a hESCs on activation of the WNT signaling pathway. Addition of IWP2 during doxycycline induction of VGLL1 effectively reduced the expression of WNT genes *FST*, *LEF1*, and *LGR5* (**Suppl. Fig. S4JH**). Moreover, this resulted in reduced expression of mesoderm markers

353 *EOMES* and *TBXT* (**Fig. 4C**) and improved expression levels of *GATA3* and *EGFR* by qPCR and
354 flow cytometry (**Fig. 4D-E**). We compared gene expression profiles of DOX and DOX/IWP2
355 treated cells (**Fig. 4F, Suppl. Table S6**) and found that DOX/IWP2 treated cells showed
356 significant loss of mesoderm-associated genes, including *EOMES* and *TBXT* (**Suppl. Table**
357 **S6**). Interestingly, ectopic expression of VGLL1 in naïve hESCs showed similar up-regulation
358 of WNT-related genes but not mesoderm markers, *EOMES* and *TBXT* (**Suppl. Fig. S4L**). These
359 data support a role for VGLL1 in regulating the expression of β-catenin-dependent WNT
360 signaling genes, an important pathway in trophoblast development and placental function.

361 VGLL1 directly regulates trophoblast-associated genes at placenta-specific 362 distal regulatory regions

363 We compared the differentially expressed genes in both the constitutive knock-down
364 and gene activation systems to further narrow down VGLL1 contributions to the TE
365 transcriptional program (**Fig. 5A**). We found 528 genes that were concordantly regulated by
366 VGLL1 expression, specifically down-regulated in VGLL1sh hESCs (relative to Scramble)
367 after TE induction, and up-regulated in the iVGLL1a hESCs treated with DOX/IWP2 (**Suppl.**
368 **Table S5**). These genes were involved in cancer pathways, focal adhesion, TGFβ and HIPPO
369 signaling pathways (**Suppl. Fig. S4E**) and were highly associated with placental gene
370 expression (**Suppl. Fig. S5A**). We then set out to investigate whether VGLL1 directly regulates
371 the transcription of a subset of these genes. VGLL1 lacks a DNA binding domain but is a
372 transcriptional co-factor for TEAD family members. TEAD4 is expressed in both hESCs and
373 hESC-derived TE-like cells and plays a role in trophoblast differentiation of hESCs as well as
374 in maintenance of stemness in primary hTSCs (58, 59). Therefore, we predicted that VGLL1
375 binding to regulatory regions might be mediated through its association with TEAD4. To test
376 this, we performed immunoprecipitation with either an anti-TEAD4 or anti-VGLL1 antibody
377 on the choriocarcinoma cell line JEG3 and confirmed co-precipitation of VGLL1 and TEAD4,
378 respectively (**Suppl. Fig. S5B**). We further confirmed VGLL1 binding to TEAD4 in BMP4-IWP2-
379 treated TE-like cells by co-immunoprecipitation (**Suppl. Fig. S5B**). We also confirmed

380 binding of VGLL1 and TEAD4 to a validated 5'UTR region of *GATA3* by ChIP-qPCR (**Suppl. Fig.**
381 **S5C**) (60).

382 Having validated the interaction between VGLL1 and TEAD4, we next set out to
383 identify binding regions of VGLL1 in complex with TEAD4 in hESC-derived TE-like cells.
384 Therefore, we performed ChIP-seq using both anti-VGLL1 and anti-TEAD4 antibodies on
385 BMP4/IWP2-treated hESCs at day 4. VGLL1 ChIP peaks were distributed mainly between
386 intronic (44.4%), intergenic (30.6%), and promoter (20%) regions (**Suppl. Fig S5D**),
387 suggesting that VGLL1 might regulate target genes through transcription activation at
388 promoters as well as through interactions at distal enhancer regions. In line with this, we
389 found evidence of VGLL1 occupancy at *EGFR* and *ENPEP* promoter regions (**Fig. 5B)**, as well
390 as in distal regulatory regions of *GATA3* and *EGFR* identified as putative active enhancers in
391 the human placenta (**Suppl. Fig. 5E, Suppl. Table S6**) (61). VGLL1 peaks were enriched in
392 TEAD family member motifs, including TEAD4 and TEAD1 (**Fig. 5C**). In our ChIP-seq data,
393 VGLL1 and TEAD4 co-occupied regions at trophoblast-associated genes *EGFR*, *ARID3A*,
394 *CDH1*, *ENPEP*, *KRT7*, *KRT8*, *PTPN14*, and *TEAD1* (**Fig. 5D**, **Suppl. Fig. S5F, Suppl. Table S6**).
395 In addition, we identified VGLL1-specific binding at TE/trophoblast regulators *ASCL2*,
396 *ARID5B*, *CCR7*, CITED2, *NODAL*, *TP63,* and WNT signaling effectors, such as the long non-
397 coding *LEF-AS1*, and components *EPHA2*, *FZD2*, *ITGB1, LGR4*, *LRP5/6*, *WLS*, *WNT5A*, and
398 *WNT5B* (**Fig. 5D and Suppl. Table S6**). The genes bound by VGLL1 were highly associated
399 with the human TE transcriptional signature and not with mesoderm-derived cell types
400 (**Suppl. Fig. S5G**). To confirm that these regulatory regions are specific for VGLL1 in the TE
401 context independently of the model used, we mined published VGLL1 CUT&TAG data from
402 Yang *et al.* in naïve TE cells (62) and found conserved binding regions in TE-specific genes,
403 including at *GATA3*, *EGFR*, and *ENPEP loci* (**Suppl. Fig. S5H**). Together, these data
404 demonstrate that VGLL1 directly regulates a TE/trophoblast-specific program at promoter
405 regions in complex with TEAD4 and at distal regulatory regions potentially with a novel
406 partner.

407 ## VGLL1 contributes to BMP4-dependent epigenetic changes by regulating the
408 histone demethylation enzyme KDM6B

409 To investigate if VGLL1 might contribute to the epigenetic changes for TE induction of
410 hESCs downstream of BMP4-treatment, we first looked at chromatin modifiers whose
411 expression patterns followed VGLL1 manipulation (**Fig. 5A, Suppl. Table S5**). Among the 528
412 genes concordantly regulated, we identified *TET2*, involved in DNA demethylation, and
413 *KDM6B*, an enzyme responsible for demethylation of histone 3 lysine residue 27
414 (H3K27me3/2/1). Furthermore, we detected VGLL1 occupancy at *KDM4B* and *KDM6B* loci in
415 BMP4/IWP2-treated hESCs (**Fig. 5D**), indicating these lysine demethylases are direct targets
416 of VGLL1 in our model. Interestingly, we found overlapping binding regions at *KDM6B* loci in
417 the VGLL1 CUT&TAG dataset in naïve TE cells from Yang *et al.* (62), suggesting a conserved
418 mechanism (**Suppl. Fig. S5H**). KDM6B has been reported to regulate the demethylation of
419 H3K27me3 at promoter regions of *Gata3* and *Cdx2*, thus promoting their expression in the
420 TE of pre-implantation murine embryos (63). Therefore, we treated hESCs with a KDM6B
421 demethylase activity inhibitor, GSKJ4, during BMP4/IWP2-mediated differentiation, and,
422 similar to VGLL1 knock-down, we found decreased levels of TE markers *GATA3* and *CDX2*,
423 compared to vehicle control (**Fig. 5F**). To confirm maintenance of the H3K27me3 repressive
424 mark at VGLL1 binding peaks of TE-specific genes, we performed CUT&RUN with an anti-
425 H3K27me3 antibody followed by qPCR (64). We found decreased H3K27me3 binding at both
426 *GATA3* and *EGFR* regions occupied by VGLL1 following 4 days of BMP4/IWP2 treatment
427 compared to d0 in control samples (**Fig. 5G**). In contrast, H3K27me3 levels at these *loci* were
428 maintained or even increased at d4 compared to d0 upon addition of GSKJ4 (**Fig. 5G**).
429 Collectively, these data suggest that VGLL1 further contributes to the establishment of a
430 TE/trophoblast transcriptional signature by regulating the expression of chromatin
431 remodeler enzymes.

432 ## VGLL1 and KDM6B are co-expressed in the TE compartment of pre-
433 implantation embryos

434 To extend our findings beyond PSC-based models, we examined VGLL1 expression in
435 human embryos. Using published datasets, we investigated its temporal and spatial
436 expression during early embryonic development. Yan *et al*. previously performed RNA
437 sequencing (RNA-seq) of individual early human embryos from the oocyte to pre-
438 implantation blastocyst (27). Analysis of this data revealed that *VGLL1* was expressed at very
439 low levels until the blastocyst stage, at which point its expression began to increase (**Suppl.**
440 **Fig. S6A**). Blakeley *et al*. investigated the transcriptional signature of the three main cell
441 types found in the pre-implantation blastocyst: TE, epiblast (EPI), and primitive endoderm
442 (PE) with single-cell RNA-seq (3). Examination of datasets from these cell types revealed that
443 *VGLL1* was highly expressed in TE, with some expression in the PE but not in EPI (**Suppl. Fig.**
444 **S6B**). To further distinguish between polar TE (adjacent to the inner cell mass) and mural TE
445 (adjacent to the blastocoele) clusters, we re-analyzed single-cell RNA-seq data of
446 dissociated human embryo cells from another study that performed extended embryo
447 culture (65). We found that *VGLL1* expression was increased in TE relative to the EPI and PE,
448 with slightly higher expression in the polar TE compared to the mural TE (**Fig 6A**). By contrast,
449 *CDX2*, another TE marker central to the TE program in humans and mice, was similarly
450 increased in TE compared to EPI and PE cells but was higher in the mural compared to the
451 polar compartment (**Fig 6A**). To assess if protein expression correlated with our
452 transcriptomic results, we co-stained pre-implantation embryos with VGLL1 and both extra-
453 embryonic (GATA3, TP63) and embryonic (NANOG) markers. We observed VGLL1 positive
454 cells at the morula stage in both inside and outside cells and co-staining with GATA3 in a
455 subset of outside cells (**Suppl. Fig. S6C**). At the blastocyst stage (day 5), VGLL1 expression
456 was higher in the nuclei of TE cells compared to ICM (**Fig 6B-C, Suppl. Fig. S6D**). These data
457 confirm VGLL1 activity in the TE compartment of human embryos, further suggesting a role
458 for this transcriptional co-factor in establishment/maintenance of this lineage.

459 Finally, to determine if VGLL1 and KDM6B were found in similar compartment *in vivo*,
460 we examined KDM6B expression in both human embryos and first trimester placentas. In
461 RNA-seq datasets from human pre-implantation embryos, *KDM6B* was highly expressed in
462 the TE, with increased expression in the polar compared to mural TE (**Fig. 6D**). However, it

463 was also expressed at high levels in PE and at lower level in the ICM. In day 5 blastocysts,
464 protein levels recapitulated observations in single cell RNA-seq datasets, with significantly
465 higher expression of KDM6B in cells of the TE compared to those of the ICM (**Fig. 6E-F**).
466 Finally, in the first trimester placenta, immunostaining showed ubiquitous expression of
467 KDM6B in the trophoblast compartment with lower expression in cytotrophoblasts (CTB)
468 compared to expression in both the proximal cell column (CC) and the syncytiotrophoblast
469 compartment (STB) (**Fig. 6G**). Taken together with our ChIPseq, these data point to a direct
470 role for VGLL1 in regulating KDM6B expression in the TE and the trophoblast compartment
471 of first trimester placenta. Overall, this study demonstrates that VGLL1 regulates both
472 genetic and epigenetic features to support a trophoblast lineage-specific transcriptional
473 program.

474 Discussion

475 Correct establishment of the TE lineage is fundamental for placental development
476 and a successful pregnancy outcome. Because of the ethical and technical limitations of
477 working with human embryos, various protocols have been reported to differentiate human
478 pluripotent stem cells into trophoblast in order to study species-specific mechanisms of
479 development. In this work, we used a BMP4-model for TE induction from primed PSCs
480 previously established in our lab, which includes addition of the WNT inhibitor IWP2 to
481 suppress mesoderm differentiation, followed by TSC culture adaptation (48). First, we
482 further explore the transcriptional programs and lineages elicited by BMP4 in the presence
483 and absence of IWP2. We demonstrated that activation of the canonical WNT pathway and
484 consequent up-regulation of mesoderm markers was limited to the first 2 days of treatment,
485 with transcription directed towards a TE-like program by day 4. Universal TE markers
486 including *GATA3*, *ENPEP*, and *TACSTD2* were all induced by BMP4 treatment. Our ATAC-seq
487 data comparing BMP4-driven differentiation with and without WNT signaling inhibition
488 indicates that, in primed hPSCs, the chromatin surrounding trophoblast lineage genes was
489 accessible for transcription throughout differentiation. The changes in chromatin reported
490 here accompany the temporal increase of TE/trophoblast markers that we observed in our

491 RNA-seq data. In contrast, we showed that, at these concentrations of BMP4, chromatin
492 near mesoderm lineage genes closed over time even in the absence of WNT inhibition.

493 Temporal activation of TE/trophoblast lineage in our model was similar to previously
494 published studies of BMP4-dependent induction of the trophoblast lineage, with TEtra
495 factors GATA2/3 and TFAP2A/C expressed early during differentiation followed by more
496 mature markers (19). However, in their study Krendl and colleagues defined VGLL1 as a late
497 marker. In this study, we determined that *VGLL1* induction occurs after the TEtra factors but
498 before other mature markers like *EGFR* and *TP63*. Knock-down of VGLL1 prevented a
499 complete exit from pluripotency with maintenance of factors like *POU5F1* and *SOX2*, and
500 blunted expression TE markers including *EGFR*, *ENPEP*, and *NR2F2*. Interestingly, we also
501 observed reduced expression of TEtra factors *GATA3*, *GATA2*, and *TFAP2C*. These data
502 suggest that while VGLL1 is not required for the initiation of the TE/trophoblast program, it
503 belongs to a second wave of factors necessary for the completion of this program in part via
504 a positive feedback loop onto the TEtra factors.

505 Such a role is not unique to VGLL1. In the mouse, Eomesodermin (Eomes),
506 downstream of Cdx2, is not required to initiate the TE program (18, 66) but instead plays a
507 key role as regulator of TE differentiation and maintenance of murine TSCs in the extra-
508 embryonic compartment (17). Mouse embryos lacking Eomes are arrested in the early post-
509 implantation period, initiating formation of TE (e.g. by cavitation), but failing to form a proper
510 chorion (17). While EOMES is not involved in human TE/trophoblast development (3, 9), our
511 data suggest that VGLL1 function similarly to reinforce the TE/TSC identity downstream of
512 early key TE lineage regulators and allow hTSC expansion. In fact, VGLL1 knock-down TE-like
513 cells adapted to TSC culture conditions showed decreased expression of *GATA2* and *GATA3*,
514 lack of cell surface expression of EGFR, and a failure to differentiate into either STB or EVT
515 cells, demonstrating lack of *bona fide* hTSC derivation.

516 Canonical WNT signaling is a crucial pathway to establish and maintain various
517 epithelial stem cells (23). As mentioned, canonical WNT signaling supports mesoderm
518 differentiation of BMP4-treated hPSCs (37). VGLL1 is uniquely expressed in the

519 TE/trophoblast lineage during early human development, thus the up-regulation of
520 mesoderm genes *EOMES*, *MIXL1*, and *TBXT* in hPSCs as a consequence of ectopic
521 expression of VGLL1 was unexpected. As in our BMP4-based differentiation protocol,
522 addition of IWP2 significantly decreased the expression of mesoderm genes in iVGLL1a
523 hESCs, suggesting that canonical WNT activation, but not mesoderm genes, may be directly
524 regulated by VGLL1. In fact, temporal activation of the canonical WNT reporter at 48-72h of
525 BMP4-treatment correlated with the timeline of VGLL1 induction. We also detected active
526 WNT signaling in DOX-treated iVGLL1a hESCs. However, our data also suggest that VGLL1
527 alone cannot sustain long-term autocrine activation of the canonical WNT pathway. We
528 detected VGLL1 occupancy at WNT coreceptors *LRP5/6* as well as the WNT agonist R-
529 Spondin receptor *LGR4*. Canonical WNT pathway activation is necessary for hTSC and
530 trophoblast organoid establishment and maintenance from placental tissue. Our data
531 therefore suggest that VGLL1 is key in supporting and maintaining the trophoblast stem cell
532 state, potentially via up-regulation of canonical WNT signaling pathway receptors and
533 coreceptors.

534 Recently, Yang and colleagues explored the role of VGLL1 and TEAD4 in TE
535 specification, starting from naïve PSCs, which represent cells of the pre-implantation
536 embryonic compartment (62). Similar to our results, they demonstrated that loss of VGLL1
537 led to maintenance of pluripotency genes, blunted expression of TE genes, including *GATA3*,
538 *ENPEP* and *NR2F2*, and inability to derive hTSCs. They demonstrated that the VGLL1/TEAD4
539 complex directly regulates expression of markers such as *GATA3* and *ENPEP* at promoter and
540 intergenic regions. In contrast, our model utilizes primed PSCs, which represent a post-
541 implantation epiblast state (50). Despite the different initial states, our data largely
542 corroborate their findings with VGLL1 directly regulating markers such as *GATA3* and *ENPEP*.
543 Interestingly, we showed VGLL1 occupancy at putative distal regulatory regions of *GATA3*
544 and *EGFR*. These regions were recently mapped to predicted enhancers active in the human
545 placenta across gestation (61), suggesting a conserved role for VGLL1 in trophoblast-
546 specific distal regulatory regions. In fact, when we looked at the presence of these distal

547 enhancer peaks in the Yang *et al* data, we found shared peaks, corroborating that they might
548 represent functionally relevant regulatory regions for TE and trophoblast identity.

549 The TE and ICM lineages segregate early during mammalian development and are
550 distinguished by unique transcriptomes and epigenomes. Both naïve and primed hPSCs
551 contain bivalent promoters at developmental genes, which maintain them in a
552 transcriptionally poised state (19, 67, 68). Such bivalent regions are characterized by
553 H3K4me3, a post-translational histone modification mark associated with active promoters,
554 and H3K27me3, a marker correlated with transcriptional repression. Global distribution of
555 H3K27me3 differs between primed and naive hPSCs. However, many of the genes marked
556 by H3K27me3 that are shared between these two pluripotency states are trophoblast
557 regulators, including GATA3 (67, 68). In their study, Yang and colleagues found a correlation
558 between loss of VGLL1 and lower levels of H3K27ac, a histone mark associated with active
559 promoters and enhancers, at TE-specific genes (62). However, the authors did not propose
560 a direct link between VGLL1 expression and H3K27ac levels at TE genes. Moreover, before
561 H3K27 can be modified by acetylation, the methyl groups need to be removed from the
562 bivalent promoter occupied by H3K27me3. We have also observed decreased levels of
563 global H3K27ac in shVGLL1 hESCs in our primed model of TE conversion and identified
564 KDM6B, one of two H3K27me3-specific demethylases in mammals, as a potential mediator
565 of VGLL1-dependent chromatin remodeling. The transcript levels of *KDM6B* were largely
566 dependent on VGLL1 expression and we observed direct VGLL1 binding on *KDM6B* gene in
567 our ChIP-seq data. In the murine blastocyst, KDM6B removes the repressive H3K27me3 at
568 the promoter regions of *Gata3* and *Cdx2* specifically in the TE (63). Loss of *Kdm6b* results in
569 maintenance of H3K27me3 at TE-specific gene regulatory regions, lower expression of Gata3
570 and Cdx2 in the TE, and an impaired ability for the embryos to implant into the uterus. In
571 humans, *VGLL1* and *KDM6B* are highly expressed in the polar TE of the human blastocyst,
572 which attaches to the uterine wall during implantation. It remains unclear whether KDM6B
573 is involved in human implantation, but our results show that chemical inhibition of KDM6B
574 demethylase activity phenocopies loss of VGLL1 and results in the blunted expression of
575 *GATA3* and *CDX2*, *in vitro*. Thus, we posit a model whereby VGLL1 directly up-regulates the

576 expression of KDM6B, which contributes to demethylation of H3K27me3 at bivalent/poised
577 promoters prior to deposition of the acetyl group by acetyltransferases at trophoblast-
578 specific regulatory regions to mark a transcriptionally permissive chromatin state (**Fig.7**).
579 Evidence suggests that, while mouse TE is fully specified in the pre-implantation blastocyst,
580 human TE lineage commitment might not be completed until post-implantation (52). VGLL1-
581 dependent regulation of chromatin remodelers like KDM6B might provide a mechanism
582 explaining the higher plasticity of human post-implantation pluripotent cells. In the human
583 placenta, KDM6B was expressed across the trophoblast compartments of first trimester
584 tissues, including villous CTB, cell column trophoblasts, and STB. Therefore, our data
585 suggests that KDM6B may play a role not just in TE specification but also in functional mature
586 trophoblasts.

587 Overall, our data extend previous findings and add a substantial mechanistic hypothesis on
588 the role of VGLL1 in TE and trophoblast lineage specification. We also demonstrated that
589 primed PSCs and the BMP4-dependent conversion protocol remain a functional and
590 practical *in vitro* model to study fundamental aspects of genetic and epigenetic regulation
591 of TE and trophoblast cell identity in humans (69).

592 Material and Methods

593 **Human embryo and Placental Tissue Collection.** Human embryos were donated to the
594 research project under the UK Human Fertilisation and Authority Licence number R0162.
595 The full study protocol was approved by the Health Research Authority's Research Ethics
596 Committee IRAS projects 308099. Human placental tissues were collected under a protocol
597 approved by the Human Research Protections Program Committee of the University of
598 California, San Diego Institutional Review Board. All patients gave informed consent for the
599 collection and use of embryos/placental tissues and samples were de-identified prior to
600 use.

601 **Trophoblast differentiation of hESC.** hESCs were first converted into TE-like cells followed
602 by adaptation to hTSC culture conditions according to a previously published protocol (70).
603 Briefly, hESCs were plated at a density of 10,000 cells per cm^2 on Geltrex-coated plates in

604 mTeSR Plus with 5μM Y27632 (Selleck Chemicals, cat. no. S1049) and differentiated in basal
605 media with 10 ng/mL recombinant human BMP4 (R&D Systems, cat. no. 314-BP-010), and 2
606 μM IWP2 (Selleck Chemicals, cat. no. S7085). At day 4, cells were either collected for
607 analysis or re-plated on collagen IV (5μg/mL)-coated plates (MilliporeSigma, cat. no. C0543-
608 1VL) in iCTB media. Other reagents include 2 μM IWR1 (MilliporeSigma, cat. no. I0161) and
609 1 μM GSK-J4 (MedChem Express, cat. no. HY-15648B).

610 Data availability

611 All datasets are available in Geo: GSE295121, GSE295122, GSE295124, GSE295126.

612 Acknowledgments

613 All sequencing was conducted at the IGM Genomics Center, University of California, San
614 Diego, La Jolla, CA. We thank Dr. Karl Willert for the generous donation of the TOPGFP HUES9
615 cell line, and Dr. Marisa Korody for the IWR1. We are grateful to Dr. Mana Parast for the
616 continuous encouragement and support. The contents of this publication are solely the
617 responsibility of the authors and do not necessarily represent the official views of CIRM or
618 any other agency of the State of California.

619 Funding

620 FS and RC were supported by the Eunice Kennedy Shriver National Institute of Child Health
621 and Human Development (NICHD) grant R01HD096260 to FS and Diversity supplement to
622 RC. This research was also supported by the California Institute for Regenerative Medicine
623 (CIRM) awards DISC0-13757 to FS and DISC0-13816 to HCA. NS was funded by CIRM EDUC4
624 – 12804 Interdisciplinary Stem Cell Training Grant. KF was partially supported by NIH Grant
625 UL1TR001442 and MH was partially supported by T32 GM145427. AM was funded by the
626 CIRM Bridges program. PDH was supported by the UC San Diego OB/GYN and Reproductive
627 Sciences. Work in the laboratory of K.K.N. was supported by the Wellcome Trust
628 (221856/Z/20/Z). Work in the laboratory of KKN was also supported by the Francis Crick
629 Institute, which receives its core funding from Cancer Research UK CC2074, the Medical

630 Research Council CC2074, and Wellcome Trust CC2074. This research was supported by
631 NIH grants HD113673, HD103161, HD062546, HD101319 and HD119510 to SP. HC was
632 supported by NIH Grant R35GM157073. This publication includes data generated at the UC
633 San Diego IGM Genomics Center utilizing an Illumina NovaSeq 6000 that was purchased
634 with funding from a National Institutes of Health SIG grant (#S10 OD026929).

635

636 # Bibliography

637 1. V. B. Alarcon, Y. Marikawa, Trophectoderm formation: regulation of morphogenesis
638 and gene expressions by RHO, ROCK, cell polarity, and HIPPO signaling HHS Public
639 Access. *Reproduction* **164**, 75–86 (2022).

640 2. M. Y. Turco, A. Moffett, Development of the human placenta. (2019).
641 https://doi.org/10.1242/dev.163428.

642 3. P. Blakeley, *et al.*, Defining the three cell lineages of the human blastocyst by single-
643 cell RNA-seq (Development, (2015) 142, 3151-3165). *Dev.* **142**, 3613 (2015).

644 4. C. Gerri, *et al.*, Initiation of a conserved trophectoderm program in human, cow and
645 mouse embryos. *Nature* **587**, 443–447 (2020).

646 5. R. M. Roberts, *et al.*, The role of BMP4 signaling in trophoblast emergence from
647 pluripotency. *Cell. Mol. Life Sci.* **79** (2022).

648 6. N. Nishioka, *et al.*, The Hippo Signaling Pathway Components Lats and Yap Pattern
649 Tead4 Activity to Distinguish Mouse Trophectoderm from Inner Cell Mass. *Dev. Cell*
650 **16**, 398–410 (2009).

651 7. A. C. Denicol, *et al.*, Canonical WNT signaling regulates development of bovine
652 embryos to the blastocyst stage. *Sci. Reports 2013 31* **3**, 1–7 (2013).

653 8. K. B. Lee, J. K. Folger, S. K. Rajput, G. W. Smith, Temporal regulation of mRNAs for
654 select bone morphogenetic proteins (BMP), BMP receptors and their associated
655 SMAD proteins during bovine early embryonic development: Effects of exogenous
656 BMP2 on embryo developmental progression. *Reprod. Biol. Endocrinol.* **12**, 1–10
657 (2014).

658 9. F. Soncin, *et al.*, Comparative analysis of mouse and human placentae across
659 gestation reveals species-specific regulators of placental development. *Dev.* **145**
660 (2018).

661 10. C. Gerri, *et al.*, A conserved role of the Hippo signalling pathway in initiation of the
662 first lineage specification event across mammals. *Dev.* **150** (2023).

663 11. M. Regin, *et al.*, Lineage segregation in human pre-implantation embryos is specified
664 by YAP1 and TEAD1. *Hum. Reprod.* **38**, 1484–1498 (2023).

665 12. B. A. T. Weatherbee, *et al.*, Distinct pathways drive anterior hypoblast specification
666 in the implanting human embryo. *Nat. Cell Biol.* **26**, 353–365 (2024).

667 13. H. Sasaki, Roles and regulations of Hippo signaling during preimplantation mouse
668 development. *Dev. Growth Differ.* **59**, 12–20 (2017).

669 14. M. Azagury, Y. Buganim, Unlocking trophectoderm mysteries: In vivo and in vitro
670 perspectives on human and mouse trophoblast fate induction. *Dev. Cell* **59**, 941–
671 960 (2024).

672 15. A. Ralston, *et al.*, Gata3 regulates trophoblast development downstream of Tead4
673 and in parallel to Cdx2. *Development* **137**, 395–403 (2010).

674 16. H. Sasaki, Position- and polarity-dependent Hippo signaling regulates cell fates in
675 preimplantation mouse embryos. *Semin. Cell Dev. Biol.* **47–48**, 80–87 (2015).

676 17. A. P. Russ, *et al.*, Eomesodermin is required for mouse trophoblast development and
677 mesoderm formation. *Nature* **404**, 95–99 (2000).

678 18. D. Strumpf, *et al.*, Cdx2 is required for correct cell fate specification and
679 differentiation of trophectoderm in the mouse blastocyst. *Development* **132**, 2093–
680 2102 (2005).

681 19. C. Krendl, *et al.*, GATA2/3-TFAP2A/C transcription factor network couples human
682 pluripotent stem cell differentiation to trophectoderm with repression of
683 pluripotency. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E9579–E9588 (2017).

684 20. Y. Li, *et al.*, BMP4-directed trophoblast differentiation of human embryonic stem
685 cells is mediated through a ΔNp63+ cytotrophoblast stem cell state. *Dev.* **140**, 3965–
686 3976 (2013).

687 21. H. M. Sonnemann, B. Pazdrak, D. A. Antunes, J. Roszik, G. Lizée, Vestigial-like 1
688 (VGLL1): An ancient co-transcriptional activator linking wing, placenta, and tumor
689 development. *Biochim. Biophys. Acta - Rev. Cancer* **1878**, 188892 (2023).

690 22. K. Qin, *et al.*, Canonical and noncanonical Wnt signaling: Multilayered mediators,
691 signaling mechanisms and major signaling crosstalk. *Genes Dis.* **11**, 103–134 (2024).

692 23. A. Haegebarth, H. Clevers, Wnt signaling, lgr5, and stem cells in the intestine and
693 skin. *Am. J. Pathol.* **174**, 715–721 (2009).

694 24. D. Pinto, H. Clevers, Wnt control of stem cells and differentiation in the intestinal
695 epithelium. *Exp. Cell Res.* **306**, 357–363 (2005).

696 25. J. Adjaye, *et al.*, Primary Differentiation in the Human Blastocyst: Comparative
697 Molecular Portraits of Inner Cell Mass and Trophectoderm Cells. *Stem Cells* **23**,
698 1514–1525 (2005).

699 26. L. Aghajanova, *et al.*, Comparative transcriptome analysis of human trophectoderm
700 and embryonic stem cell-derived trophoblasts reveal key participants in early
701 implantation. *Biol. Reprod.* **86**, 11–12 (2012).

702 27. L. Yan, *et al.*, Single-cell RNA-Seq profiling of human preimplantation embryos and
703 embryonic stem cells. *Nat. Publ. Gr.* (2013). https://doi.org/10.1038/nsmb.2660.

704 28. M. Krivega, W. Essahib, H. Van de Velde, WNT3 and membrane-associated β-catenin
705 regulate trophectoderm lineage differentiation in human blastocysts. *Mol. Hum.*
706 *Reprod.* **21**, 711–722 (2015).

707 29. S. Sonderegger, J. Pollheimer, M. Knöfler, Wnt signalling in implantation,

708 decidualisation and placental differentiation--review. *Placenta* **31**, 839–847 (2010).

709 30. Y. Hu, *et al.*, Single-cell analysis of nonhuman primate preimplantation development
710 in comparison to humans and mice. *Dev. Dyn.* **250**, 974–985 (2021).

711 31. H. Okae, *et al.*, Derivation of Human Trophoblast Stem Cells. *Cell Stem Cell* **22**, 50-
712 63.e6 (2018).

713 32. S. Haider, *et al.*, Self-Renewing Trophoblast Organoids Recapitulate the
714 Developmental Program of the Early Human Placenta. *Stem Cell Reports* **11**, 537–
715 551 (2018).

716 33. M. Y. Turco, *et al.*, Trophoblast organoids as a model for maternal–fetal interactions
717 during human placentation. *Nature* **564**, 263–281 (2018).

718 34. B. Dietrich, S. Haider, G. Meinhardt, J. Pollheimer, · Martin Knöfler, WNT and NOTCH
719 signaling in human trophoblast development and differentiation. **79**, 292 (2022).

720 35. A. S. Brumm, *et al.*, Initiation and maintenance of the pluripotent epiblast in pre-
721 implantation human development is independent of NODAL signaling. *Dev. Cell* **60**
722 (2025).

723 36. J. Au, *et al.*, Role of autocrine bone morphogenetic protein signaling in trophoblast
724 stem cells. *Biol. Reprod.* (2021). https://doi.org/10.1093/biolre/ioab213.

725 37. D. Kurek, *et al.*, Endogenous WNT signals mediate BMP-induced and spontaneous
726 differentiation of epiblast stem cells and human embryonic stem cells. *Stem Cell*
727 *Reports* **4**, 114–128 (2015).

728 38. N. Sah, F. Soncin, Conserved and divergent features of trophoblast stem cells. *J.*
729 *Mol. Endocrinol.* **72** (2024).

730 39. L. Li, *et al.*, Single-cell multi-omics sequencing of human early embryos. *Nat. Cell*
731 *Biol. 2018 207* **20**, 847–858 (2018).

732 40. L. Liu, *et al.*, An integrated chromatin accessibility and transcriptome landscape of
733 human pre-implantation embryos. *Nat. Commun. 2019 101* **10**, 1–11 (2019).

734 41. A. L. Wilkinson, I. Zorzan, P. J. Rugg-Gunn, Epigenetic regulation of early human
735 embryo development. *Cell Stem Cell* **30**, 1569–1584 (2023).

736 42. S. Io, *et al.*, Capturing human trophoblast development with naive pluripotent stem
737 cells in vitro. *Cell Stem Cell* **28**, 1023-1039.e13 (2021).

738 43. G. Guo, *et al.*, Human naive epiblast cells possess unrestricted lineage potential.
739 *Cell Stem Cell* **28**, 1040-1056.e6 (2021).

740 44. J. K. Cinkornpumin, *et al.*, Naive Human Embryonic Stem Cells Can Give Rise to
741 Cells with a Trophoblast-like Transcriptome and Methylome. *Stem Cell Reports* **15**,
742 198–213 (2020).

743 45. M. Horii, T. Bui, O. Touma, H. Y. Cho, M. M. Parast, An Improved Two-Step Protocol
744 for Trophoblast Differentiation of Human Pluripotent Stem Cells. *Curr. Protoc. Stem*
745 *Cell Biol.* **50**, 1–21 (2019).

746 46. M. Amita, *et al.*, Complete and unidirectional conversion of human embryonic stem
747 cells to trophoblast by BMP4. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 1–10 (2013).

748 47. M. Horii, *et al.*, Human pluripotent stem cells as a model of trophoblast
749 differentiation in both normal development and disease. *Proc. Natl. Acad. Sci. U. S.*
750 *A.* **113**, E3882–E3891 (2016).

751 48. F. Soncin, *et al.*, Derivation of functional trophoblast stem cells from primed human
752 pluripotent stem cells. *Stem Cell Reports* **17**, 1303–1317 (2022).

753 49. N. Moya, J. Cutts, T. Gaasterland, K. Willert, D. A. Brafman, Endogenous WNT
754 signaling regulates hPSC-derived neural progenitor cell heterogeneity and specifies
755 their regional identity. *Stem Cell Reports* **3**, 1015–1028 (2014).

756 50. C. Zhao, *et al.*, A comprehensive human embryo reference tool using single-cell
757 RNA-sequencing data. *Nat. Methods* | **22**, 193–206 (2025).

758 51. W. Huang, *et al.*, Uncovering the role of TET2-mediated ENPEP activation in
759 trophoblast cell fate determination. *Cell. Mol. Life Sci.* **81** (2024).

760 52. D. Meistermann, *et al.*, Integrated pseudotime analysis of human pre-implantation
761 embryo single-cell transcriptomes reveals the dynamics of lineage specification.
762 *Cell Stem Cell* **28**, 1625-1640.e6 (2021).

763 53. X. Luan, *et al.*, Down-regulation of CCR7 via AKT pathway and GATA2 inactivation
764 suppressed trophoblast migration and invasion in recurrent spontaneous abortion†.
765 *Biol. Reprod.* **102**, 424–433 (2020).

766 54. M. Drukker, *et al.*, Isolation of primitive endoderm, mesoderm, vascular endothelial
767 and trophoblast progenitors from human pluripotent stem cells. *Nat. Biotechnol.* **30**,
768 531–542 (2012).

769 55. C. Dong, *et al.*, A genome-wide CRISPR-Cas9 knockout screen identifies essential
770 and growth-restricting genes in human trophoblast stem cells. *Nat. Commun. 2022*
771 *131* **13**, 1–16 (2022).

772 56. C. Dong, *et al.*, Derivation of trophoblast stem cells from naïve human pluripotent
773 stem cells. *Elife* **9**, 1–26 (2020).

774 57. D. Horst, *et al.*, Differential WNT activity in colorectal cancer confers limited
775 tumorigenic potential and is regulated by MAPK signaling. *Cancer Res.* **72**, 1547–
776 1556 (2012).

777 58. B. Saha, *et al.*, TEAD4 ensures postimplantation development by promoting
778 trophoblast self-renewal: An implication in early human pregnancy loss. *Proc. Natl.*
779 *Acad. Sci. U. S. A.* **117**, 17864–17875 (2020).

780 59. G. Meinhardt, *et al.*, Pivotal role of the transcriptional co-activator YAP in trophoblast
781 stemness of the developing human placenta. *Proc. Natl. Acad. Sci. U. S. A.* **117**,
782 13562–13570 (2020).

783 60. P. Home, *et al.*, Altered subcellular localization of transcription factor TEAD4
784 regulates first mammalian cell lineage commitment. *Proc. Natl. Acad. Sci. U. S. A.*
785 **109**, 7362–7367 (2012).

786 61. D. M. Owen, *et al.*, Genome-wide identification of transcriptional enhancers during
787 human placental development and association with function, differentiation, and
788 disease†. *Biol. Reprod.* **109**, 965–981 (2023).

789 62. Y. Yang, *et al.*, VGLL1 cooperates with TEAD4 to control human trophectoderm
790 lineage specification. *Nat. Commun.* **15** (2024).

791 63. B. Saha, *et al.*, EED and KDM6B coordinate the first mammalian cell lineage
792 commitment to ensure embryo implantation. *Mol. Cell. Biol.* **33**, 2691–2705 (2013).

793 64. A. Panday, R. Elango, N. A. Willis, R. Scully, A modified CUT&RUN-seq technique for
794 qPCR analysis of chromatin-protein interactions. *STAR Protoc.* **3**, 101529 (2022).

795 65. F. Zhou, *et al.*, Reconstituting the transcriptome and DNA methylome landscapes of
796 human implantation. *Nature* **572**, 660–664 (2019).

797 66. H. Niwa, *et al.*, Interaction between Oct3/4 and Cdx2 determines trophectoderm
798 differentiation. *Cell* **123**, 917–929 (2005).

799 67. B. Kumar, *et al.*, Polycomb repressive complex 2 shields naïve human pluripotent
800 cells from trophectoderm differentiation. *Nat. Cell Biol. 2022 246* **24**, 845–857
801 (2022).

802 68. D. W. Zijlmans, *et al.*, Integrated multi-omics reveal polycomb repressive complex 2
803 restricts human trophoblast induction. **24**, 858–871 (2022).

804 69. A. S. Seetharam, *et al.*, The product of BMP-directed differentiation protocols for
805 human primed pluripotent stem cells is placental trophoblast and not amnion. *Stem*
806 *Cell Reports* **17. 10.**, /j.stemcr.2022.04.014 (2022).

807 70. V. C. Cheung, *et al.*, Current Strategies of Modeling Human Trophoblast Using
808 Human Pluripotent Stem Cells in vitro. *Curr. Protoc.* **3**, e875 (2023).

809

Figure Legends

Figure 1. BMP4 treatment of hESCs induces chromatin remodeling. A. Venn diagram of up-regulated genes following treatment of hESCs with BMP4 with and without IWP2 for 2 days. Gene list enrichment analysis on the KEGG pathway database (2021) for genes uniquely up-regulated with BMP4 alone treatment is presented, a=WNT-mediated b=FGF-mediated c=TGFβ-mediated. **B.** Gene expression levels of markers for the mesoderm lineage across treatment/days by bulk RNA-seq. **C.** Heatmaps illustrating open chromatin loci by ATAC-seq. **D.** Signal intensity of open chromatin at mesoderm- and TE-specific genes. **E.** Heatmap of differentially expressed chromatin modifiers genes across treatment/days.

Figure 2. Loss of VGLL1 prevents completion of the TE/trophoblast program. A. Gene expression levels of early and late trophoblast markers in the first 72h of BMP4/IWP2 treatment of hESCs by qPCR. Statistical analysis relative to 0h **B.** Gene expression levels of early and late trophoblast markers in shVGLL1 compared to Scramble hESCs during BMP4/IWP2 treatment by qPCR. Asterisks represent statistical differences within cell lines compared to d0. Letters represent statistical significance of shVGLL1 compared to same timepoint on Scramble. **C.** Heatmaps of differentially expressed genes between shVGLL1 and Scramble hESCs at d0 and d4 of BMP4/IWP2 treatment by RNA-seq organized by gene cluster analysis. **D.** Bright field images of Scramble and shVGLL1 hESC-derived TSCs. **E.** Cell surface expression of EGFR in hESC-derived TSCs by flow cytometry. **F.** Gene expression levels of TEtra factors and EGFR in hESC-derived TSCs by qPCR. Statistical analysis compared to Scramble. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

Figure 3. VGLL1 regulates a TE/trophoblast program independently of BMP4 treatment. A. Schematic representation of the inducible CRISPR-mediated gene activation system, which includes a doxycycline (DOX)-inducible dCas9-VPR and VGLL1-specific guide RNAs. **B.** Bright field images of DMSO (control) and DOX-treated hESCs at day 4. **C.** Cell surface expression of EGFR in DOX-treated iVGLL1a hESCs at day 4 by flow cytometry. **D.** Gene expression levels of lineage markers in DOX-treated iVGLL1a hESCs at day 4 by qPCR. Statistical analysis compared to DMSO. **E.** Heatmap of differentially expressed genes in iVGLL1a hESCs at day 4 of DMSO or DOX treatment. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$

Figure 4. VGLL1 regulates expression of canonical WNT effectors. A. Heatmap of WNT target genes differentially expressed in iVGLL1a hESCs at day 4 of DMSO or DOX treatment. **B.** β-catenin-dependent WNT activation in DOX-treated hESC using a TCF/LEF-GFP reporter plasmid. **C.** Gene expression levels of mesoderm markers in DOX-treated iVGLL1a hESCs with and without IWP2 by qPCR. **D.** Gene expression levels of early and late trophoblast markers in DOX-treated iVGLL1a hESCs with and without IWP2 by qPCR. **E.** Cell surface

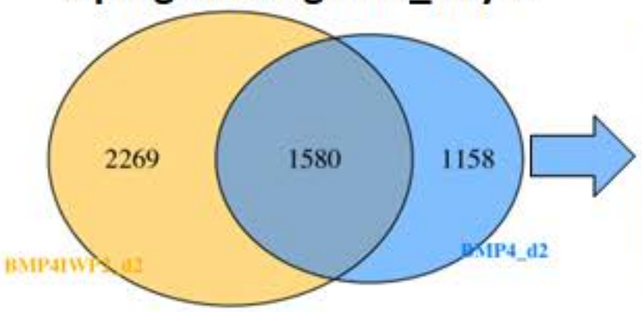
expression of EGFR in DOX-treated iVGLL1a hESCs with and without IWP2 at day 4 by flow cytometry. **F.** Heatmap of differentially expressed genes in iVGLL1a hESCs at day 4 of DMSO or DOX treatment, with and without IWP2. * p≤0.05, ** p≤0.01, *** p≤0.001, **** p≤0.0001.

Figure 5. VGLL1 directly regulates chromatin modifier gene KDM6B. A. Venn diagram of concordantly regulated genes downregulated in shVGLL1 clones and up-regulated in DOX-treated iVGLL1a hESCs. **B.** Genome browser tracks showing VGLL1 and TEAD4 genomic enrichment peaks (ChIP-seq) and chromatin accessibility peaks (ATAC-seq) for representative TE gene loci of BMP4/IWP2-treated hESCs at day 4. **C.** Motif analysis using HOMER showing the significantly enriched DNA-binding motifs at VGLL1-binding sites in BMP4/IWP2-treated hESCs at day 4. **D.** Co-occupancy analysis by signal density pileups of VGLL1 and TEAD4 peaks in ChIP-seq of BMP4/IWP2-treated hESCs at day 4. **E.** Genome browser tracks showing VGLL1 genomic enrichment peaks (ChIP-seq) and chromatin accessibility peaks (ATAC-seq) for chromatin modifiers, KDM6B and KDM4B loci of BMP4/IWP2-treated hESCs at day 4. **F.** Gene expression levels of early trophoblast markers after 4 days of treatment of hESCs with BMP4/IWP2 in the presence of KDM6B inhibitor GSKJ4 (or DMSO as control). **G.** CUT&RUN for repressive histone mark H3K27me3 followed by qPCR for regulatory regions occupied by VGLL1 of TE markers GATA3 and EGFR. * p≤0.05, ** p≤0.01, **** p≤0.0001, ns = non-significant

Figure 6. VGLL1 and KDM6B are highly expressed in the polar TE of pre-implantation blastocysts. A. Violin plots showing expression levels of VGLL1 and CDX2 in the mural TE, polar TE, epiblast, and primitive endoderm (PE) cells of d6 blastocysts (data from Zhou *et al.* 2019). **B.** Immunofluorescence staining of a day 5 human blastocyst with VGLL1, TP63, and NANOG. **C.** Quantification of VGLL1 staining in the TE and ICM compartments of day 5 human blastocyst (n=3) ***p=0.0005. **D.** Violin plots showing expression levels of KDM6B in the mural TE, polar TE, epiblast, and primitive endoderm (PE) cells of d6 blastocysts (data from Zhou *et al.* 2019). **E.** Immunofluorescence staining of a day 5 human blastocyst with KDM6B and TP63. **F.** Quantification of KDM6B staining in the TE and ICM compartments of day 5 human blastocyst (n=3) *p=0.0117. **G.** Immunofluorescence staining of KDM6B in a week 6 placenta. CTB= cytotrophoblast, STB=syncytiotrophoblast, CC=cell column

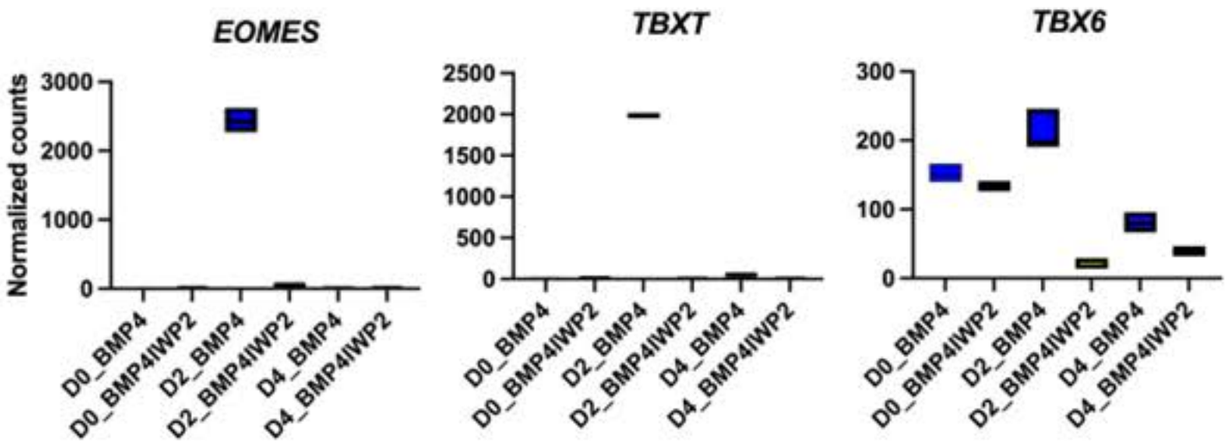
Figure 7. Graphic summary of the results. Created in BioRender. Soncin, F. (2025).

A

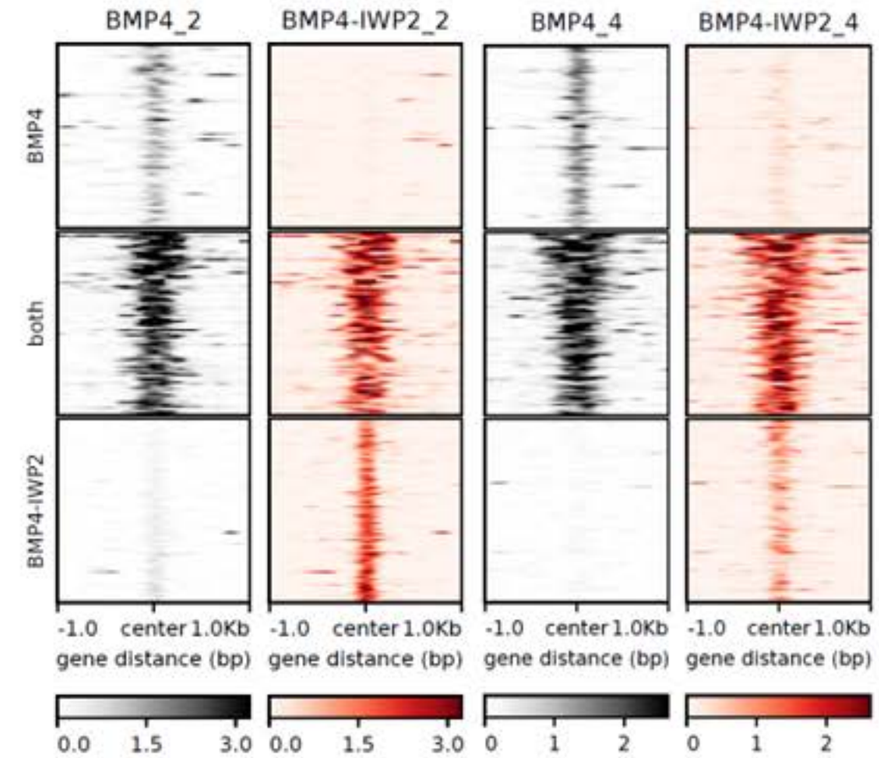


Term	Adj. P-value
Gastric cancer[a,b,c]	2.96E-04
Hippo signaling pathway	6.77E-04
Wnt signaling pathway	1.56E-03
Basal cell carcinoma[a]	1.56E-03
Breast cancer[a,b]	6.95E-03
Regulation of actin cytoskeleton	7.32E-03

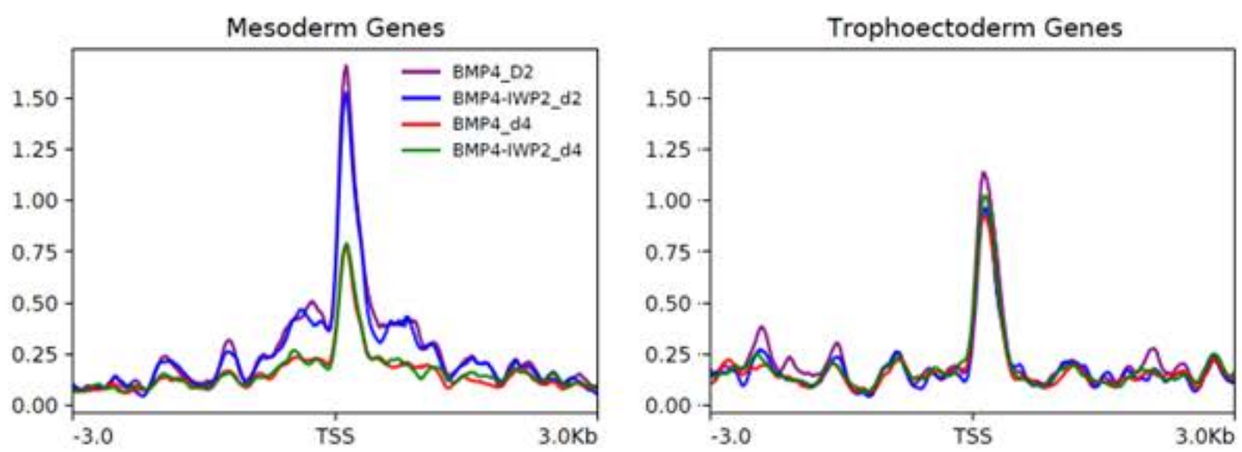
B



C



D



E

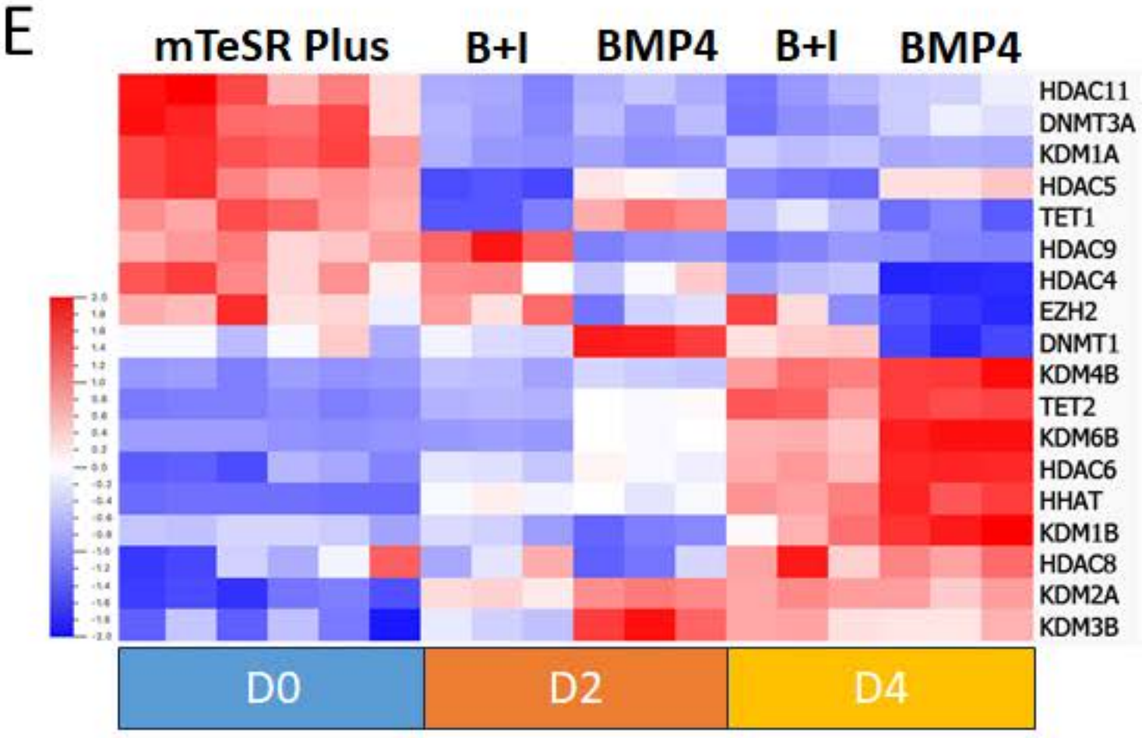


Figure 1.

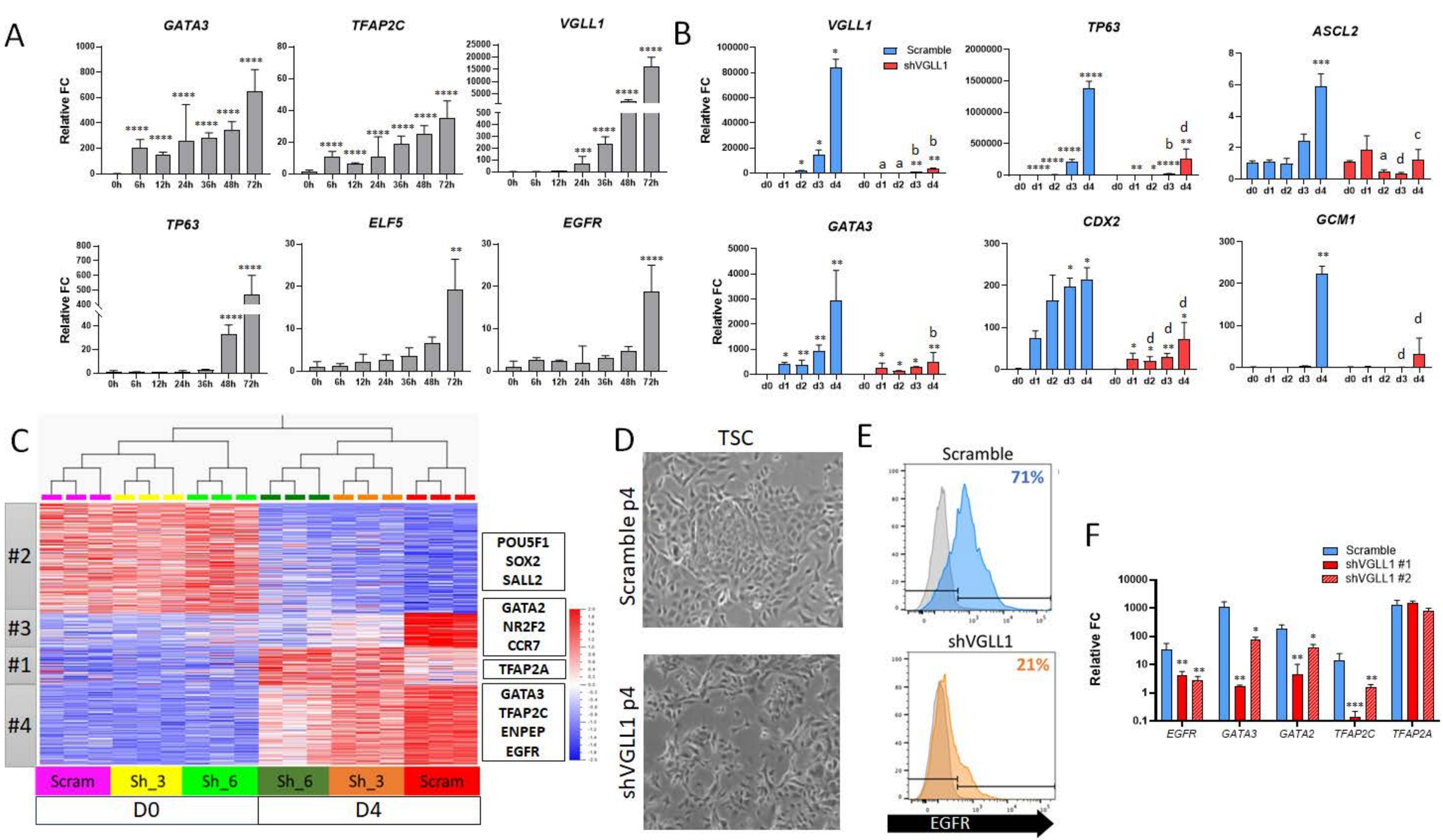

A
GATA3
TFAP2C
VGLL1
TP63
ELF5
EGFR
Relative FC
0h 6h 12h 24h 36h 48h 72h
B
VGLL1
TP63
ASCL2
GATA3
CDX2
GCM1
Scramble
shVGLL1
d0 d1 d2 d3 d4
C
#2
#3
#1
#4
POU5F1
SOX2
SALL2
GATA2
NR2F2
CCR7
TFAP2A
GATA3
TFAP2C
ENPEP
EGFR
Scram
Sh_3
Sh_6
Sh_6
Sh_3
Scram
D0
D4
D
TSC
Scramble p4
shVGLL1 p4
E
Scramble
71%
shVGLL1
21%
EGFR
F
Scramble
shVGLL1 #1
shVGLL1 #2
Relative FC
EGFR
GATA3
GATA2
TFAP2C
TFAP2A

Figure 2.

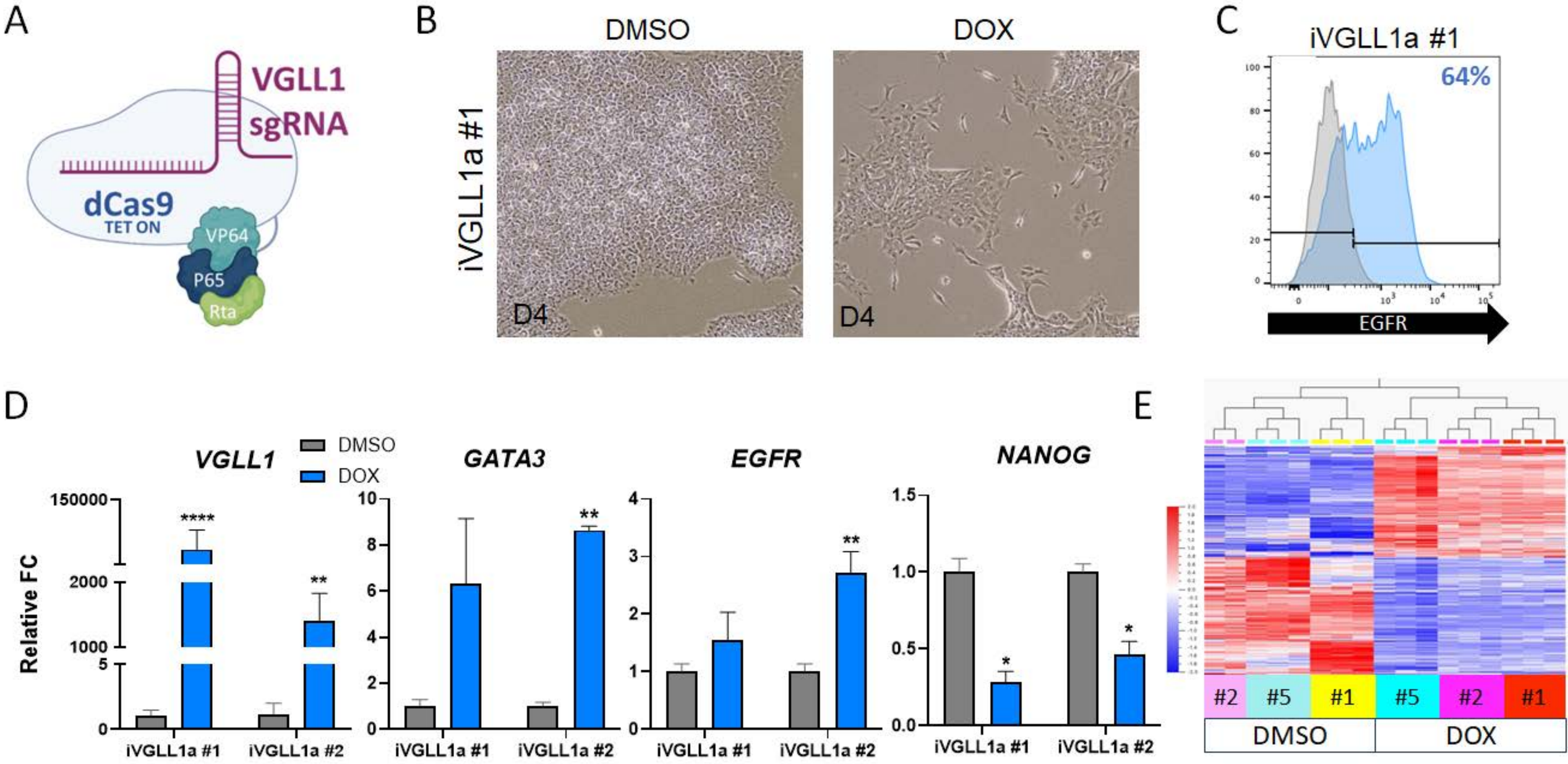

A
VGLL1
sgRNA
dCas9
TET ON
VP64
P65
Rta
B
DMSO
DOX
iVGLL1a #1
D4
D4
C
iVGLL1a #1
64%
EGFR
D
VGLL1
GATA3
EGFR
NANOG
DMSO
DOX
Relative FC
iVGLL1a #1
iVGLL1a #2
E
#2
#5
#1
#5
#2
#1
DMSO
DOX

Figure 3.

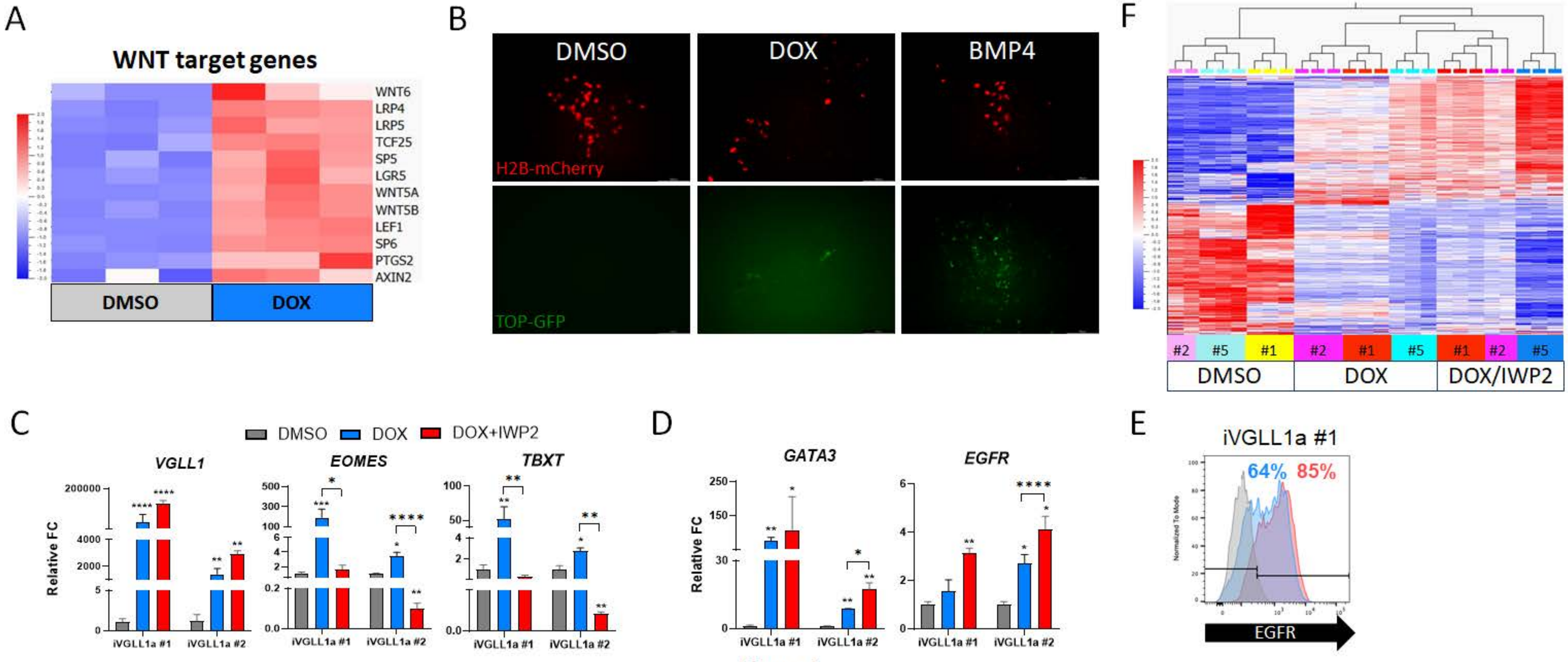

A
WNT target genes
WNT6
LRP4
LRP5
TCF25
SP5
LGR5
WNT5A
WNT5B
LEF1
SP6
PTGS2
AXIN2
DMSO
DOX
B
DMSO
DOX
BMP4
H2B-mCherry
TOP-GFP
F
#2
#5
#1
#2
#1
#5
#1
#2
#5
DMSO
DOX
DOX/IWP2
C
DMSO
DOX
DOX+IWP2
VGLL1
EOMES
TBXT
Relative FC
iVGLL1a #1
iVGLL1a #2
D
GATA3
EGFR
Relative FC
iVGLL1a #1
iVGLL1a #2
E
iVGLL1a #1
64%
85%
EGFR

Figure 4.

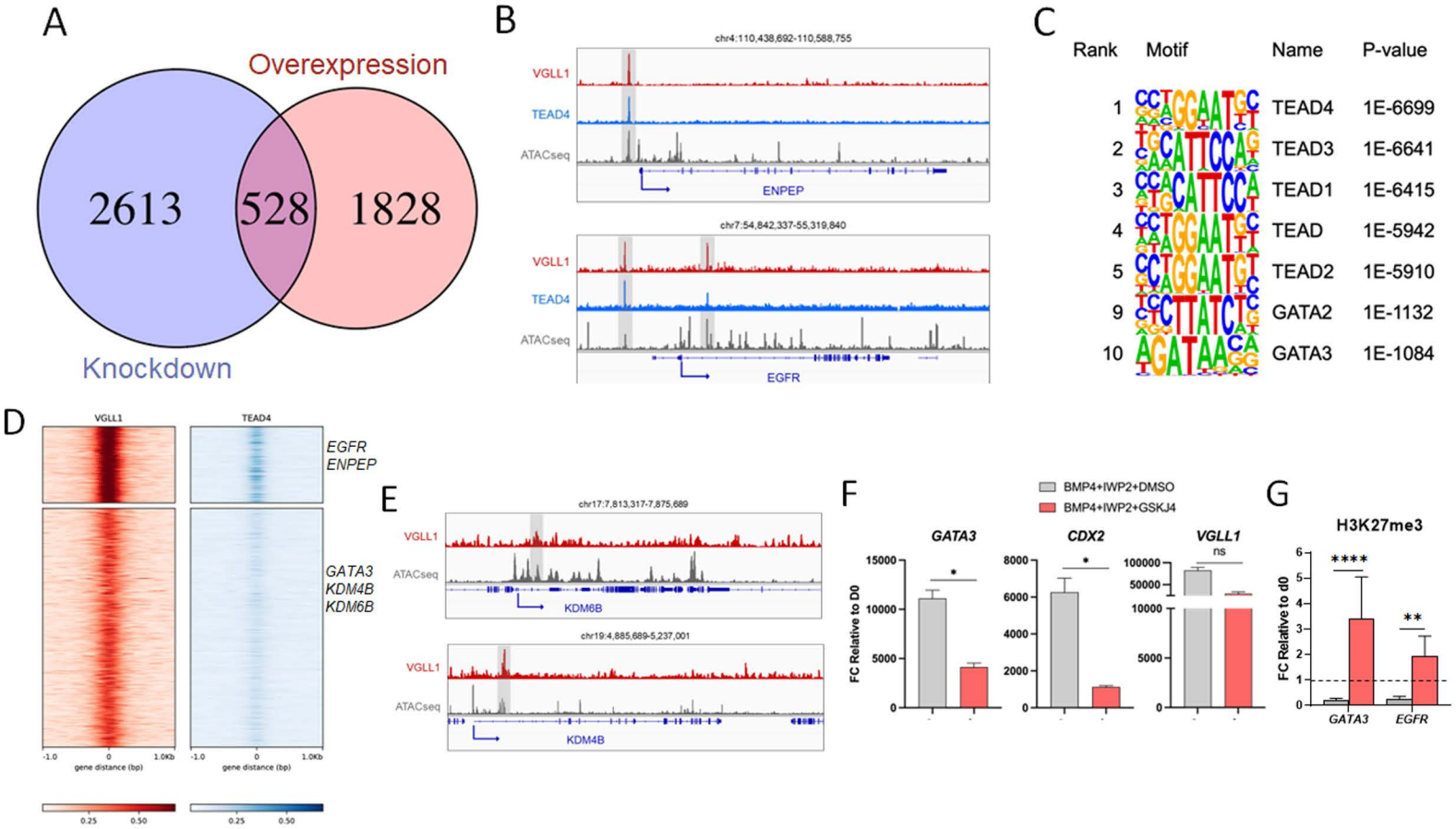

A
Overexpression
2613
528
1828
Knockdown
B
chr4:110,438,692-110,588,755
VGLL1
TEAD4
ATACseq
ENPEP
chr7:54,842,337-55,319,840
VGLL1
TEAD4
ATACseq
EGFR
C
Rank
Motif
Name
P-value
1
TEAD4
1E-6699
2
TEAD3
1E-6641
3
TEAD1
1E-6415
4
TEAD
1E-5942
5
TEAD2
1E-5910
9
GATA2
1E-1132
10
GATA3
1E-1084
D
VGLL1
TEAD4
EGFR
ENPEP
GATA3
KDM4B
KDM6B
-1.0
0
1.0Kb
gene distance (bp)
0.25
0.50
E
chr17:7,813,317-7,875,689
VGLL1
ATACseq
KDM6B
chr19:4,885,689-5,237,001
VGLL1
ATACseq
KDM4B
F
BMP4+IWP2+DMSO
BMP4+IWP2+GSKJ4
GATA3
CDX2
VGLL1
ns
FC Relative to D0
G
H3K27me3
FC Relative to d0
****
**
GATA3
EGFR

Figure 5

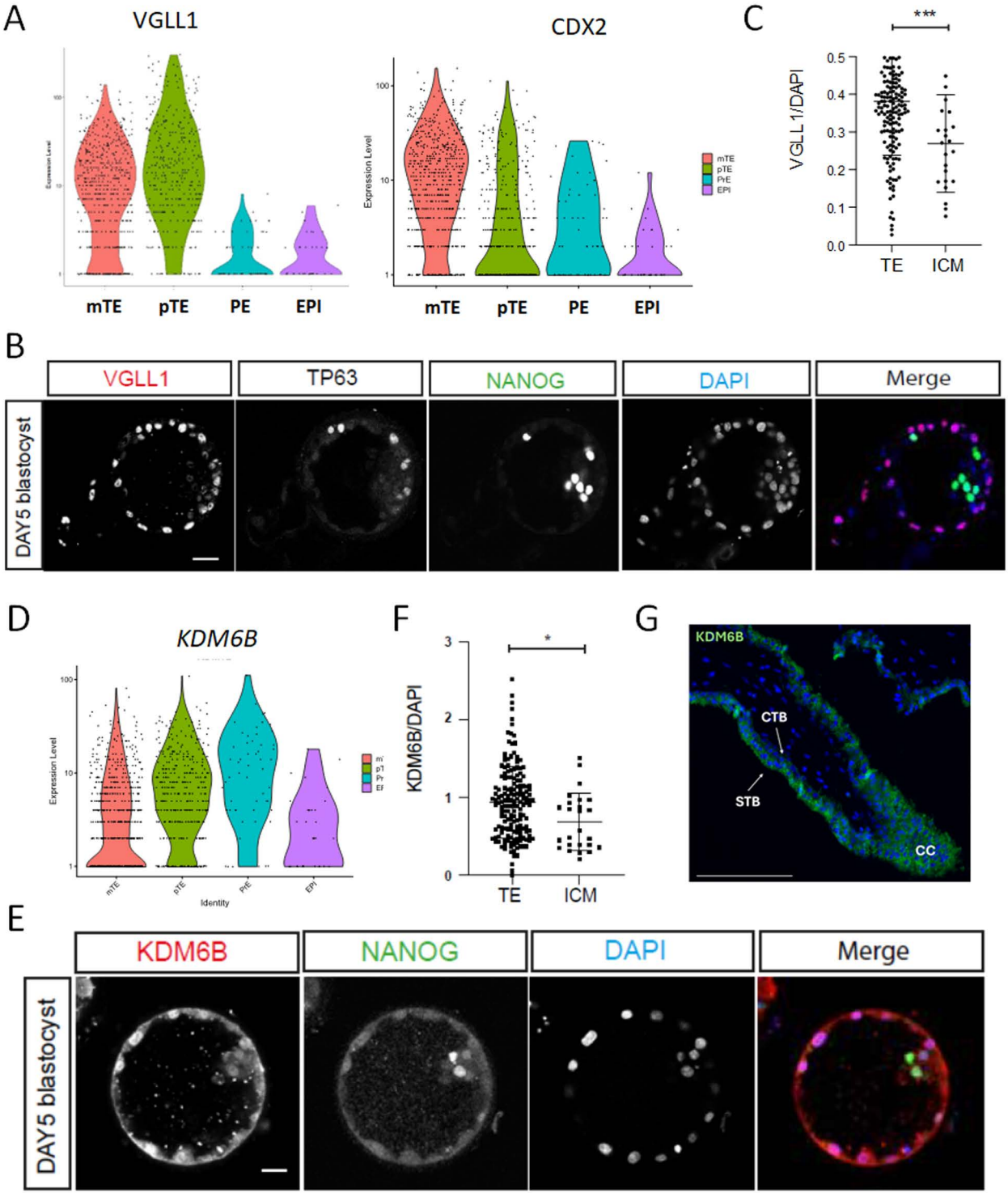

A
VGLL1
CDX2
Expression Level
mTE
pTE
PE
EPI
C
***
VGLL1/DAPI
TE
ICM
B
VGLL1
TP63
NANOG
DAPI
Merge
DAY5 blastocyst
D
KDM6B
Identity
F
*
KDM6B/DAPI
G
KDM6B
CTB
STB
CC
E
KDM6B
NANOG
DAPI
Merge
DAY5 blastocyst

Figure 6

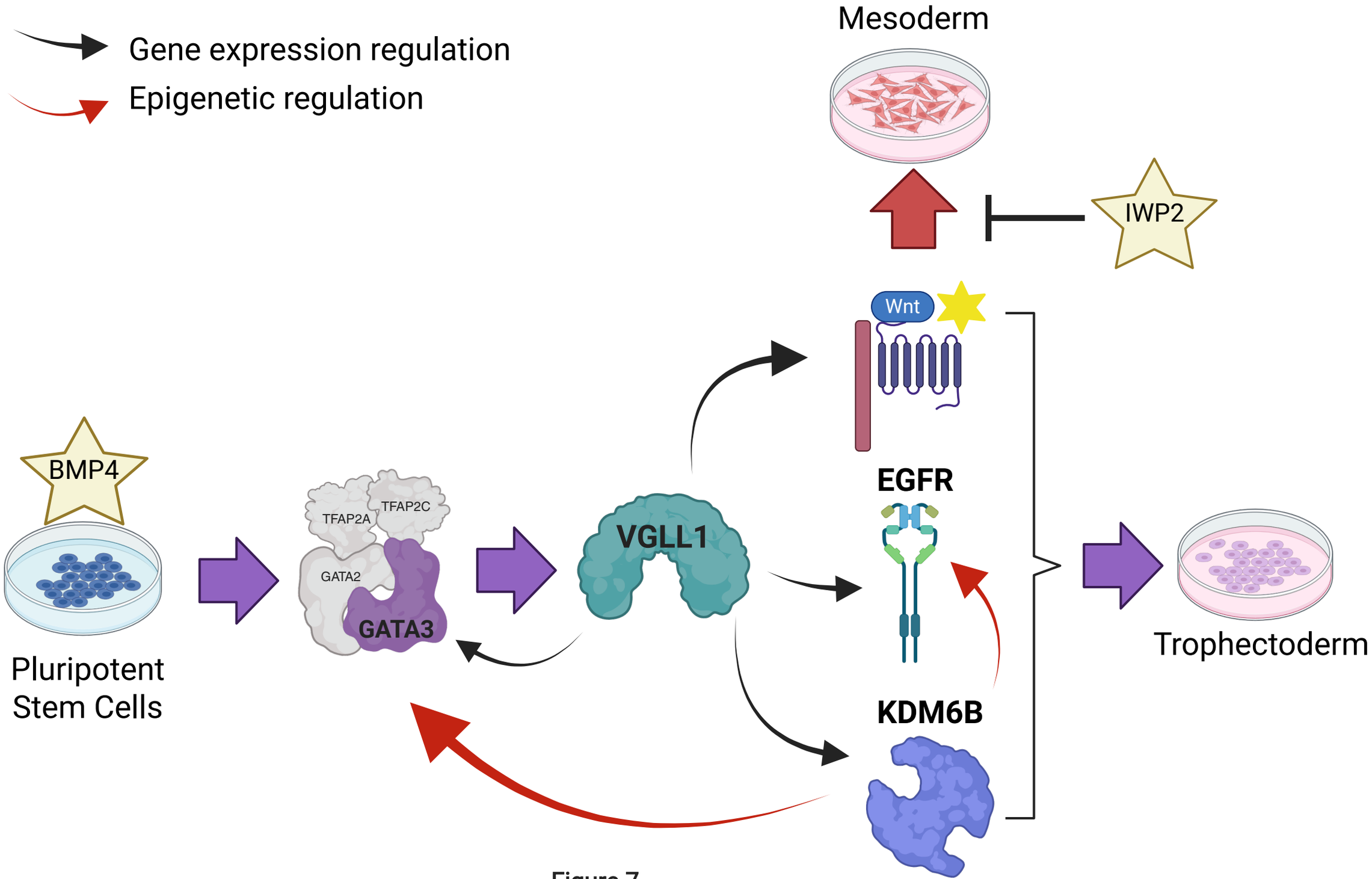

Gene expression regulation
Epigenetic regulation
Mesoderm
IWP2
Wnt
EGFR
BMP4
TFAP2A
TFAP2C
GATA2
GATA3
VGLL1
KDM6B
Pluripotent
Stem Cells
Trophectoderm

Figure 7