# The disruption of CtBP regulates DUX-dependent and -independent genetic program for the 2-cell-like state in murine embryonic stem cells

Kota Sugiyama[1], Kazuma Yoshioka[1], Nao Hayakawa[1], Ryo Masuda, Miyu Marutani, Selma Alamanda Abadi, Maiko Yamamoto and Yoshiyuki Seki

Department of Biomedical Sciences, School of Biological and Environmental Sciences, Kwansei Gakuin University, Hyogo, Japan

[1]Co-first author

[*]Correspondence: yseki@kwansei.ac.jp

**Abstract**

After fertilization, maternally deposited mRNA is cleared, and *de novo* mRNA is transcribed from the zygotic genome through zygotic genome activation (ZGA), which is called maternal-to-zygotic transition (MZT) occurring in the mouse at 2-cell (2C) stage. 2C-like cells (2CLCs) marked by MERVL expression and transcriptionally similar to 2C embryos spontaneously emerge from mouse embryonic stem cells (mESCs). Although the emergence of 2CLCs completely depends on DUX function, a recent knockout study clearly showed that DUX is dispensable for mouse embryos, suggesting that DUX-independent molecular pathways are not recapitulated in 2CLCs. We present here that the disruption of C-terminal binding protein 1/2 (*Ctbp1/2*) activates DUX-dependent and -independent molecular pathway associated with the development of early mouse embryos mediated by the upregulation of Preferentially expressed antigen of melanoma family-like 7 (PRAMEL7). Furthermore, the abnormality of the gene expression profile caused by *Dux* KO is partially rescued by the overexpression of PRAMEL7 in mESCs. Our study provides new insights into the DUX-independent molecular pathway for developing early mouse embryos.

32 **Introduction**

33 After fertilization, maternal factors direct development and trigger zygotic genome
34 activation (ZGA) at the around late one-cell stage in mice, called maternal-to-zygotic
35 transition (MZT) (Eckersley-Maslin et al., 2018a). The early-cleavage stage embryos
36 acquire cellular totipotency to generate both embryonic proper and extraembryonic
37 tissues during MZT. Although totipotency was thought to be exclusive to the early-
38 cleavage stage embryos, a part of the molecular network is recapitulated in the totipotent-
39 like cells (2-cell-like cells; 2CLC) marked by the expression of endogenous retrovirus
40 MERVL, which are an endogenous subpopulation of mouse embryonic stem cells (ESCs)
41 (Macfarlan et al., 2012a; Rodriguez-Terrones et al., 2018a).

42 The transition from pluripotent stem cells into 2CLCs is a useful *in vitro* system for
43 identifying the gene regulatory network for ZGA, and several genes that control the
44 emergence of 2CLCs from pluripotent cells have been identified (Hu et al., 2020; Ishiuchi
45 et al., 2015; Kreibich et al., 2019; Percharde et al., 2018; Rodriguez-Terrones et al.,
46 2018b). The transcription factor DUX (DUX4 in humans) is a critical factor for regulating
47 the transition from pluripotent stem cells into 2CLCs and the expression of ZGA genes
48 (De Iaco et al., 2017; Hendrickson et al., 2017; Whiddon et al., 2017a). Ectopic expression
49 of DUX enhances the population of 2CLCs, whereas the *Dux* disruption eliminates the
50 emergence of spontaneous 2CLCs during mESC culture. Several 2CLC regulatory genes
51 have been identified, but they are either direct activators (DPPA2/4, p53 and NELFA)
52 (Grow et al., 2021; Hu et al., 2020; Kreibich et al., 2019)or direct repressors (CAF-1,
53 LINE-1 and SETDB1) (Ishiuchi et al., 2015; Percharde et al., 2018)of *Dux* expression.

However, both *Dux* zygotic and maternal knockout embryos can survive full-term (Chen and Zhang, 2019; Guo et al., 2019), suggesting that DUX-independent molecular pathways not recapitulated in 2CLCs are critical for ZGA and the early development of mice.

Here, we report that the disruption of *Ctbp1/2* (C-terminal binding protein 1/2), transcriptional corepressor (Stankiewicz et al., 2014), activates subsets of genes that are transiently upregulated in a DUX-dependent at ZGA, a group of genes that are maternally expressed genes and upregulated after ZGA in DUX-independent manner. Preferentially Expressed Antigen in Melanoma (PRAME) 7 is directly upregulated by the loss of *Ctbp1/2* without DUX function, and upregulated PRAMEL7 activate subsets of genes that are transiently upregulated by ZGA in a DUX-dependent and a group of genes that are maternally expressed genes and upregulated after ZGA in DUX-independent manner. These data revealed that the disruption of *Ctbp1/2* recapitulates a part of DUX-independent ZGA cascades in ESCs, and *Ctbp1/2* DKO ESCs are useful *in vitro* systems to investigate DUX-independent molecular cascades ensuring ZGA in mice.

71 **Results**

72 ***Ctbp1/2 Deficiency enhances the transition from pluripotent stem cells to 2-cell-like***
73 ***Cells***

74 We have previously shown that the CtBP1/2 are required to transition from formative to
75 naïve pluripotency in mouse embryonic stem cells (mESCs) (Yamamoto et al., 2020). To
76 elucidate the molecular functions of CtBP1/2 in the transition from formative to naïve
77 pluripotency in mESCs, we compared the gene expression profile between wild-type
78 (WT) and *Ctbp1/2* double knockout (DKO) ESCs using RNA-sequencing (RNA-Seq).
79 MA-plot comparing WT ESCs with *Ctbp1/2* DKO ESCs showed that the 2-cell (2C)-
80 restricted genes, including *Dux*, *Zscan4* and *Zfp352*, as well as three repeat elements
81 MERVL, MT2 and GSAT_Mm are significantly upregulated by *Ctbp1/2* DKO (Figure
82 1A, B). Gene set enrichment analysis (GSEA) using a 2-cell embryo and 2-cell-like cell
83 (2CLC)-enriched genes (404) (Macfarlan et al., 2012a) showed that 2C-restricted genes
84 were significantly enriched in the upregulated genes by the loss of *Ctbp1/*2 (Figure 1C).
85 Next, to monitor the dynamics of the transition from pluripotent stem cells into 2C-like
86 cells, we established *Ctbp1/2* DKO ESCs carrying the turboGFP reporter under the
87 control of MERVL long terminal repeat (*Mervl*-EGFP) (Ishiuchi et al., 2015). The GFP-
88 positive fraction was about 0.1 % in WT ESCs, while remarkedly, *Ctbp1/2* DKO induced
89 a more than 50-fold increase of 2CLC (7.37%). To confirm the key markers of 2CLC, we
90 performed immunofluorescence analysis with anti-ZSCAN4, Mervl-GAG and OCT4
91 antibodies (Figure 1E). *Mervl*-EGFP-positive cells emerged by *Ctbp1/2* DKO expressed
92 ZSCAN4 and Mervl-GAG and lacked OCT4. Our data suggested that the downregulation

93 of *Ctbp1/2* might be involved in the zygotic genome activation *in vivo*. Consistent with
94 this idea, maternal mRNA of CtBP2 was reduced between the 2-cell to 4-cell stage based
95 on the single-cell RNA-Seq data of mouse early embryo (Gerlinger et al., 2014) (Figure
96 1F). However, the expression level of CtBP2 was maintained in spontaneous 2CLCs
97 (Figure 1G), suggesting that the molecular pathway triggered by the loss of *Ctbp1/2* in
98 the 2-cell embryo is stalled in spontaneous 2CLCs.

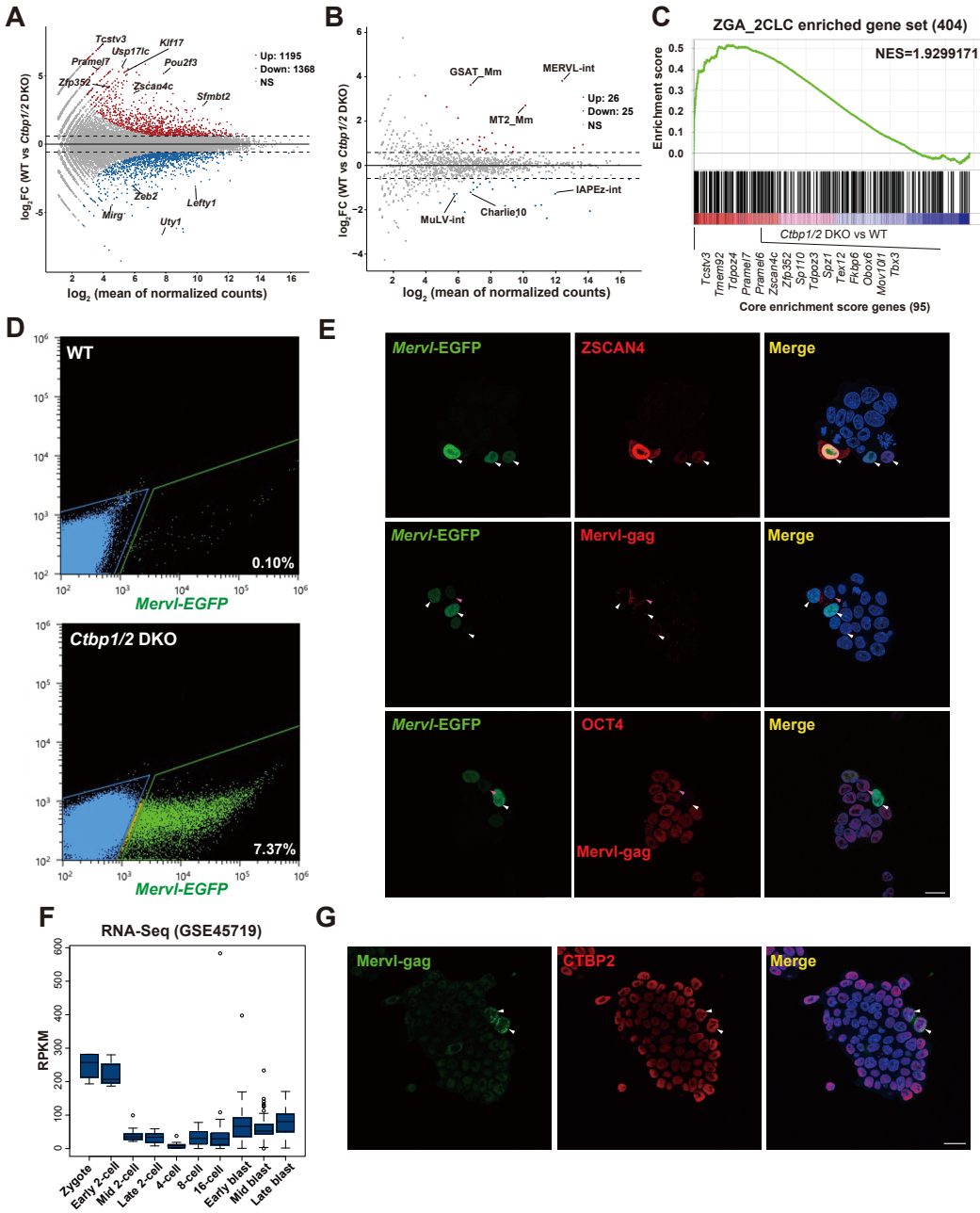



**Figure 1. The disruption of *Ctbp1/2* expands the 2-cell-like cells**

(A), (B) MA-plot showing mean expression and fold change per gene or repeat elements in *Ctbp1/2* DKO versus WT ESCs. Cut-off values were set to 0.05 ≥ *P* and over 2-fold changes. Red and blue dots indicate upregulated genes and downregulated genes, respectively. (C) Gene set enrichment analysis (GSEA) using ZGA, and 2-cell enriched genes defined in Macfalan et al. (2012) (GSE33923) (Macfarlan et al., 2012a) to compare *Ctbp1/2* DKO ESCs with WT ESCs. (D) FACS analysis of MERVL::TurboGFP in WT ESCs and *Ctbp1/2* DKO ESCs. (E) MERVL::TurboGFP-positive cells were marked by

107 ZSCAN4 and Mervl-gag proteins and mutually exclusive with OCT4-positive cells. Scale
108 bar: 10 μm. (F) mRNA levels of *Ctbp2* in early mouse development based on the
109 published single-cell RNA-Seq data (Gerlinger et al., 2014). (G) Immunofluorescence for
110 CtBP2 (red) in WT ESCs with MERVL::TurboGFP (green). Scale bar: 10 μm.
111

112 ***2C program induced by Ctbp1/2 Deficiency partially depends on DUX***

113 DUX, the homeobox transcription factor, is essential and sufficient for the emergence of
114 spontaneous 2CLCs from ESCs (De Iaco et al., 2017; Whiddon et al., 2017b). To
115 determine the dependency of the 2C-program driven by *Ctbp1/2* deficiency on DUX, we
116 deleted the *Dux* cluster locus in WT and *Ctbp1/2* DKO ESCs (Supplementary Figure 1C).
117 Firstly, we compared the population of MERVL-EGFP-positive cells between *Ctbp1/2*
118 DKO and *Ctbp1/2*/Dux TKO ESCs. Although *Dux* KO lost most of the MERVL-EGFP-
119 strong positive cells in *Ctbp1/2* DKO ESCs, MERVL-EGFP-weak positive cells were
120 found in *Ctbp1/2*/*Dux* TKO ESCs (Figure 2A). We have analyzed the expression levels
121 of 2C genes in WT, *Dux* KO, *Ctbp1/2* DKO and TKO ESCs using qRT-PCR. Upregulation
122 of Preferentially Expressed Antigen in Melanoma-like 7 (*Pramel7*) (Graf et al., 2017a),
123 which is upregulated at the 2C stage of the mouse embryo (Guo et al., 2019), by the loss
124 of *Ctbp1/2* was unaffected by *Dux* KO, while *Dux* KO downregulated the expression of
125 *Zscan4c* in *Ctbp1/2* DKO ESCs (Figure 2B). Next, to determine the effect of the loss of
126 *Ctbp1/2* and *Dux* in the transcriptional activity of MERVL-LTR, we performed a
127 luciferase assay using the reporter driven by MERVL-LTR. *Ctbp1/2* deficiency
128 significantly upregulated the reporter driven by MERVL-LTR, and the mutation of the
129 DUX-recognition sequence impaired this activation, but the activity was substantially

130 higher than those in WT ESCs (Supplementary Figure 2A). Furthermore, the activation
131 of MERVL-Luciferase (Luc) by the loss of *Ctbp1/2* was significantly suppressed by the
132 deletion of *Dux*. However, the reporter activity of MERVL-Luc and MERVL-Luc with
133 the mutation of the DUX-recognition sequence (DUXm) in *Ctbp1/2*/*Dux* TKO ESCs was
134 considerably higher than those in *Dux* KO ESCs. These findings suggest that the loss of
135 *Ctbp1/2* activates the reporter activity of MERVL-Luc without the DUX function. qRT-
136 PCR analysis showed that the activation of *Zscan4c* is more strongly dependent on DUX
137 than the activation of *Mervl* and *Naalad2* by the loss of *Ctbp1/2*. Published ChIP-Seq data
138 (Hendrickson et al., 2017) of DUX showed that DUX is localized at two distinct regions
139 around the *Zscan4c* locus (Supplementary Figure 2B). No MERVL sequence was found
140 at the major peak of DUX-binding, and the weak peak of DUX-binding was localized in
141 the MERVL sequence. In contrast, all DUX-binding peaks around the *Naalad2* locus are
142 co-localized with the MERVL sequence. These findings suggested that *Zscan4c* is
143 regulated by the DUX-binding elements with or without MERVL, while only DUX-
144 binding elements with MERVL regulate *Naalad2*. To confirm these possibilities, we
145 examined the reporter activity driven by DUX-binding elements without MERVL and
146 with MERVL. Reporter activity driven by the DUX-binding element without MERVL
147 upstream of *Zscan4c* was upregulated by the loss of *Ctbp1/2*, and this activation was
148 cancelled entirely by the deletion of *Dux* (Supplementary Figure 2C). The reporter
149 activity driven by the DUX-binding element with MERVL at *Zscan4c* and *Naalad2* was
150 also upregulated by the *Ctbp1/2* disruption, and this upregulation was significantly but
151 not entirely repressed by the loss of *Dux*. Consistent with the results of MERVL-Luc,

152 these reporter activities in *Ctbp1/2*/*Dux* TKO ESCs were higher than those in *Dux* KO
153 ESCs.
154 Next, to identify the 2C genes whose expression is upregulated by *Ctbp1/2* DKO without
155 DUX function, we performed RNA-Seq using WT, *Dux* KO, *Dppa2* KO, *Ctbp1/2* DKO,
156 *Ctbp1/2*/*Dux* KO and *Ctbp1/2*/*Dppa2* KO ESCs. DPPA2/DPPA4 complex is the upstream
157 element for the activation of *Dux* in ESCs (Eckersley-Maslin et al., 2018; Kreibich et al.,
158 2019). DUX and DPPA2-dependent upregulated genes by the loss of *Ctbp1/2* including
159 *Zscan4* and *Pramef25* were enriched in Cluster 2, which was separated by K-means
160 clustering analysis (K=5) using top 2000 differentially expressed genes (Figure 2C).
161 Cluster 3, which was enriched with a group of genes whose expression is reduced by *Dux*
162 KO and this reduction is cancelled by *Ctbp1/2* DKO, contained KRAB-ZFP cluster
163 (*Zfp900*/*981*/*600*/*534*/*458*) (Wolf et al., 2020). Cluster 4, in which the genes are
164 upregulated by *Ctbp1/2* DKO and TKO, contained the genes upregulated by ZGA
165 including *Pramel6/7*, *Pramef17*, *Cphx1/2/3* and *Crxos*. Interestingly, Cluster 5, enriched
166 with a group of genes whose expression was decreased by *Dppa2* KO and increased by
167 *Ctbp1/2* DKO, was enriched in GO regarding reproduction including *Tex11*, *Dazl* and
168 *Hormad1*. To verify whether the gene expression changes in ESCs induced by *Ctbp1/2*
169 DKO reflect the gene expression dynamics in early embryos, we analyzed the expression
170 changes of genes belonging to Cluster 2 (DUX-dependent genes), Cluster 4 (DUX-
171 independent genes) and Cluster 5 (DPPA2-dependent genes) in early mouse embryo using
172 public RNA-Seq data (Gerlinger et al., 2014). We analyzed the expression dynamics of
173 Cluster 2 genes in early embryos by removing genes with low expression in early embryos

174 (1≧FPKM). Most genes were transiently upregulated in 2-4 cell stages. In contrast, K-
175 means clustering (K=6) analysis of cluster 4 showed that they were classified into two
176 major groups: maternal factors whose expression decreased after fertilization (Cluster1,
177 2) and genes whose expression was activated after ZGA (Cluster3-6). Transcriptome data
178 showed that repressed genes by *Dux* KO tended to be upregulated by *Ctbp1/2* DKO
179 without DUX functions (Figure 2C). To validate the tendency of genes whose expression
180 is reduced by *Dux* KO to be activated by *Ctbp1/2* DKO in a DUX-independent manner,
181 we performed GSEA on *Ctbp1/2*/*Dux* TKO and *Dux* KO transcriptomes using the top 500
182 of downregulated genes by *Dux* KO as gene sets. We observed the significant enrichment
183 of the downregulated genes by the loss of *Dux* except for *Zscan4* in the upregulated genes
184 by *Ctbp1/2* DKO in the *Dux* KO ESCs (Figure 2D). Furthermore, the Venn diagram
185 overlapping the downregulated genes by *Dux* KO and the upregulated by *Ctbp1/2* DKO
186 in *Dux* KO showed that most of the down-regulated genes by *Dux* KO are upregulated by
187 *Ctbp1/2* DKO in *Dux* KO (Figure 2E). Comparison of the expression of repeat sequences
188 revealed that the Mervl elements (MERVL-int and MT2_MM) and major satellite repeats
189 (GSAT_MM) were upregulated by the loss of *Ctbp1/2*. The upregulation of Mervl
190 elements was in *Ctbp1/2* DKO significantly repressed by *Dux* KO, while the upregulation
191 of major satellite repeats by *Ctbp1/2* DKO was unaffected by *Dux* KO (Figure 2F).

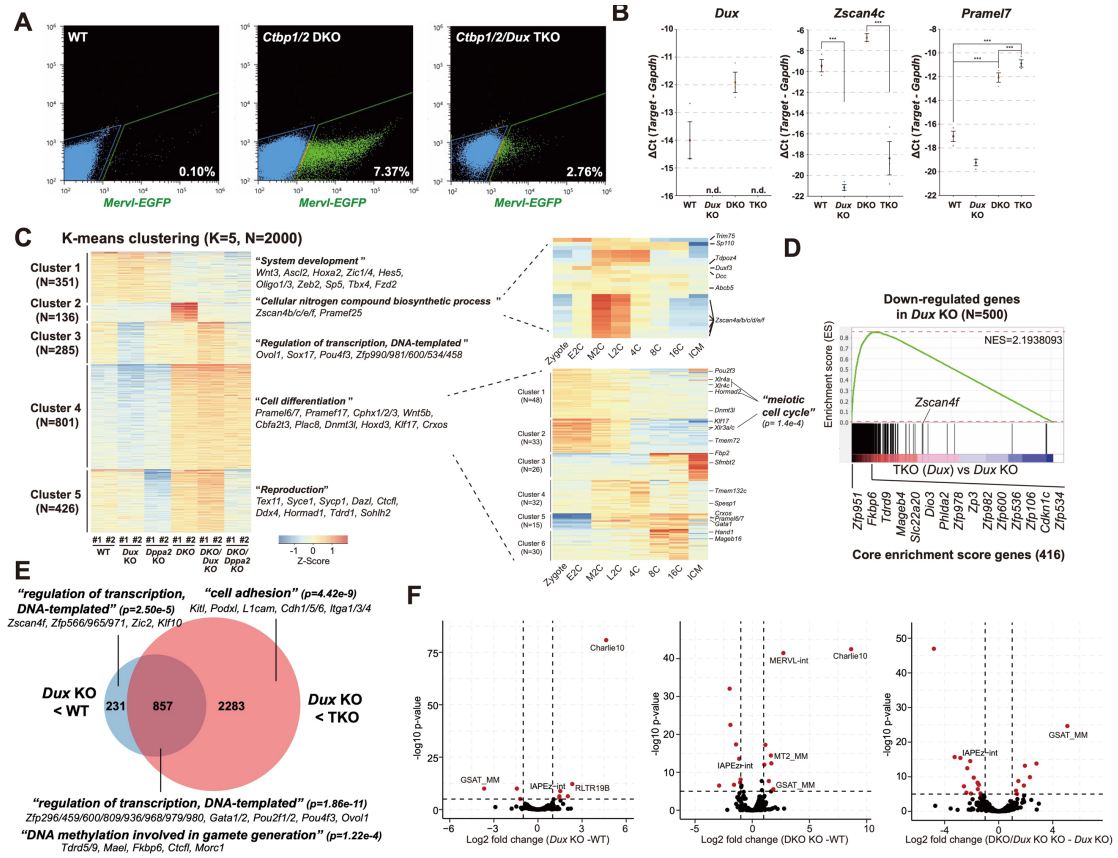



**Figure 2. *Ctbp1/2* disruption partially rescues the *Dux* KO phenotype in ESCs**

(A) FACS analysis of MERVL::TurboGFP in WT, *Ctbp1/2* DKO and *Ctbp1/2*/*Dux* TKO ESCs. (B) RT-qPCR analysis of *Dux*, *Zscan4c* and *Pramel7* in WT, *Dux* KO, *Ctbp1/2* DKO and Ctbp1/2/*Dux* TKO ESCs. The ΔCt of each gene was calculated by subtracting from the average Ct of Gapdh to that of each gene Ct. Means of ΔCt were calculated from three biological triplicates. ***$P < 0.001$, **$P < 0.005$, *$P < 0.05$ (Tukey-Kramer multiple comparison test). (C) K-means clustering analysis (K=5, N=2000) using RNA-Seq data of WT, *Dux* KO, *Dppa2* KO, *Ctbp1/2* DKO, *Ctbp1/2*/*Dux* TKO and *Ctbp1/2*/*Dppa2* TKO ESCs. K-means clustering analysis of the genes of Cluster2 and Cluster4 using published single-cell RNA-Seq data of early mouse embryos (Gerlinger et al., 2014). (D) Gene set enrichment analysis (GSEA) using down-regulated genes by Dux KO (N=500) to compare *Ctbp1/2*/*Dux* TKO ESCs with *Dux* KO ESCs. (E) Venn diagram illustrating the overlap between the down-regulated genes by *Dux* KO and the upregulated genes by *Ctbp1/2* DKO in *Dux* KO. (F) Volcano plot showing differentially accumulated repetitive elements between WT ESCs and *Dux* KO ESCs, WT ESCs and *Ctbp1/2* DKO ESCs, *Dux* KO ESCs and *Ctbp1/2*/*Dux* TKO ESCs.

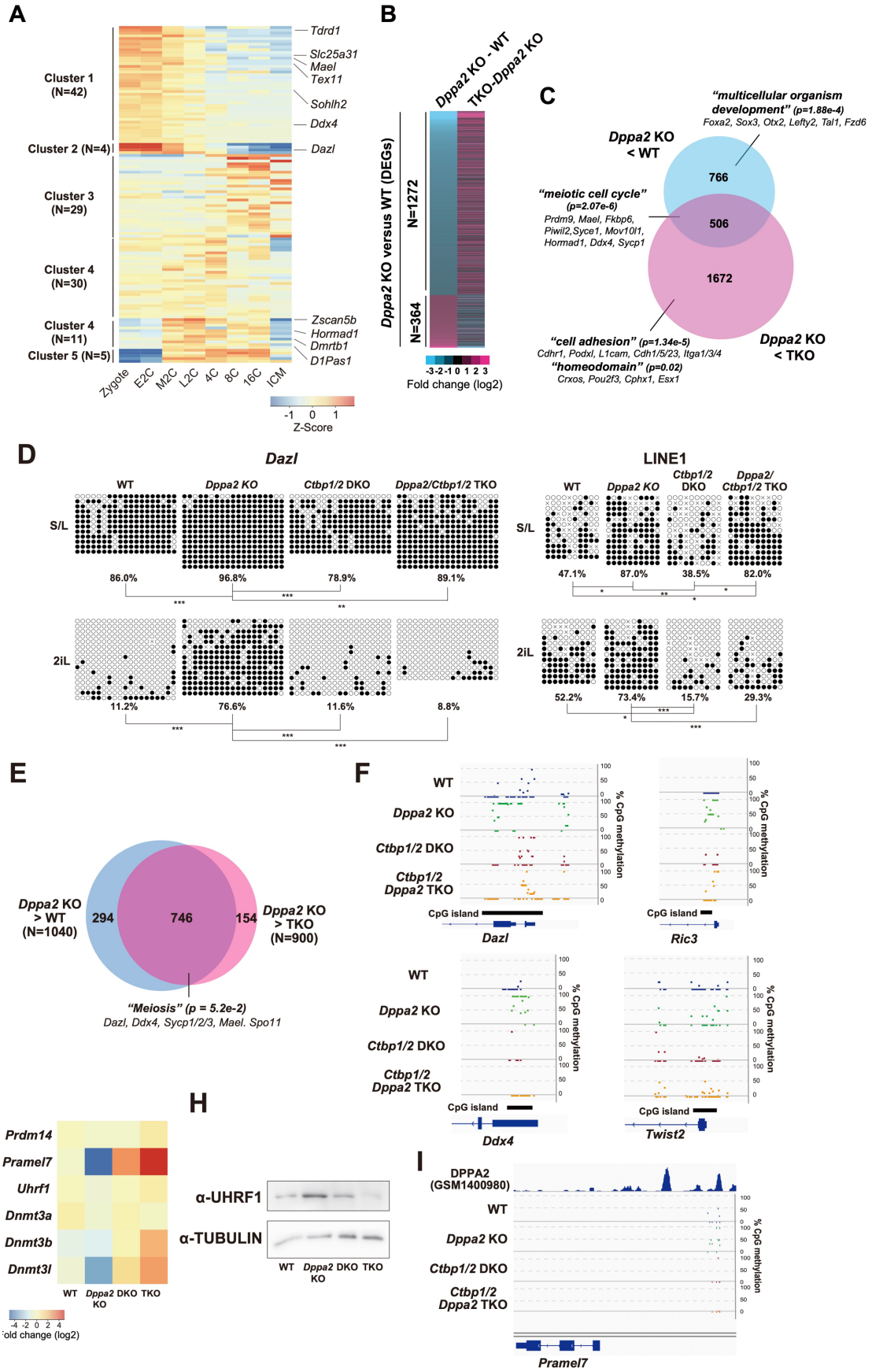
208

209 ***Dppa2 KO phenotype is partially rescued by the disruption of Ctbp1/2***

210 K-means clustering analysis of transcriptome data using WT, *Ctbp1/2* DKO, *Dppa2* KO
211 and *Ctbp1/2*/*Dppa2* TKO ESCs revealed that down-regulated genes by *Dppa2* KO tend
212 to be upregulated by *Ctbp1/2* DKO (Figure 2C) (Cluster 5). These genes were categorized
213 into two groups: maternally expressed genes and the genes activated after major ZGA
214 (Figure 3A). Comparing the expression changes caused by *Dppa2* KO in WT ESCs with
215 those caused by *Ctbp1/2* DKO in *Dppa2* KO ESCs showed a clear negative correlation
216 (Figure 3B, C). Germ-line-specific genes were enriched in the rescued gene category by
217 *Ctbp1/2* DKO in *Dppa2* KO ESCs (Figure 3C). It was previously reported that DPPA2
218 safeguard against *de novo* DNA methylation through the establishment of H3K4me3
219 enrichment in ESCs cultured in 2i plus LIF condition (Gretarsson and
220 Hackett)(Eckersley-maslin et al.). We compared the DNA methylation status at around
221 CpG island of *Dazl* and promoter regions of LINE1 in WT, *Dppa2* KO, *Ctbp1/2* DKO,
222 *Ctbp1/2*/*Dppa2* TKO ESCs with serum LIF and 2i LIF conditions using bisulfite
223 sequencing analysis. DNA methylation levels were consistently elevated by *Dppa2* KO
224 in WT ESCs but not in *Ctbp1/2* DKO, and this tendency was more pronounced in ESCs
225 with 2iL than SL (Figure 3D). We compared the global DNA methylation levels in WT,
226 *Dppa2* KO, *Ctbp1/2* DKO and *Ctbp1/2/Dppa2* TKO ESCs under 2i LIF condition using
227 reduced-representation bisulfite sequencing (RRBS). Because it has been reported that
228 DPPA2 preferentially binds to CpG island and protects aberrant *de novo* DNA
229 methylation (Gretarsson and Hackett), we extracted the promoter region with CpG island

230 in which the DNA methylation levels were elevated by *Dppa2* KO in ESCs (Figure 3E).
231 Interestingly, the DNA methylation levels at almost all these promoter regions, including
232 meiosis-associated genes, *Dazl*, *Ddx4* and *Sycp1/2/3*, were reduced by the disruption of
233 *Ctbp1/2* in *Dppa2* KO ESCs. (Figure 3E, F). To determine the molecular mechanism that
234 *Ctbp1/2* disruption impaired the abnormal elevation of DNA methylation caused by
235 *Dppa2* KO in ESCs, we compared the mRNA expression of the regulators of DNA
236 methylation and DNA methyltransferases by RT-qPCR analysis. *Pramel7,* which is
237 involved in degrading UHRF1 protein (Graf et al., 2017b), was significantly repressed in
238 *Dppa2* KO ESCs, and this repression was entirely cancelled by *Ctbp1/2* disruption
239 (Figure 3G). Consistent with the expression of *Pramel7*, the protein level of UHRF1 was
240 elevated by *Dppa2* KO, and this up-regulation was diminished by *Ctbp1/2* disruption
241 (Figure 3H). RRBS and published ChIP-Seq data of DPPA2 (Engelen et al., 2015) showed
242 that DNA methylation level was elevated upstream region of *Pramel7* associated with
243 DPPA2-binding (Figure 3I), suggesting that DPPA2 maintain DNA hypomethylation
244 through the activation of *Pramel7* expression.

245


A
Cluster 1 (N=42)
Cluster 2 (N=4)
Cluster 3 (N=29)
Cluster 4 (N=30)
Cluster 4 (N=11)
Cluster 5 (N=5)
Tdrd1
Slc25a31
Mael
Tex11
Sohlh2
Ddx4
Dazl
Zscan5b
Hormad1
Dmrtb1
D1Pas1
Zygote
E2C
M2C
L2C
4C
8C
16C
ICM
Z-Score
B
Dppa2 KO - WT
TKO-Dppa2 KO
Dppa2 KO versus WT (DEGs)
N=1272
N=364
Fold change (log2)
C
"multicellular organism development" (p=1.88e-4)
Foxa2, Sox3, Otx2, Lefty2, Tal1, Fzd6
Dppa2 KO < WT
766
"meiotic cell cycle" (p=2.07e-6)
Prdm9, Mael, Fkbp6, Piwil2,Syce1, Mov10l1, Hormad1, Ddx4, Sycp1
506
1672
"cell adhesion" (p=1.34e-5)
Cdhr1, Podxl, L1cam, Cdh1/5/23, Itga1/3/4
"homeodomain" (p=0.02)
Crxos, Pou2f3, Cphx1, Esx1
Dppa2 KO < TKO
D
Dazl
WT
Dppa2 KO
Ctbp1/2 DKO
Dppa2/Ctbp1/2 TKO
S/L
86.0%
96.8%
78.9%
89.1%
2iL
11.2%
76.6%
11.6%
8.8%
LINE1
47.1%
87.0%
38.5%
82.0%
52.2%
73.4%
15.7%
29.3%
E
Dppa2 KO > WT (N=1040)
294
746
154
Dppa2 KO > TKO (N=900)
"Meiosis" (p = 5.2e-2)
Dazl, Ddx4, Sycp1/2/3, Mael. Spo11
F
WT
Dppa2 KO
Ctbp1/2 DKO
Ctbp1/2 Dppa2 TKO
% CpG methylation
CpG island
Dazl
Ric3
Ddx4
Twist2
H
Prdm14
Pramel7
Uhrf1
Dnmt3a
Dnmt3b
Dnmt3l
WT
Dppa2 KO
DKO
TKO
α-UHRF1
α-TUBULIN
I
DPPA2 (GSM1400980)
Pramel7

246 **Figure 3. The loss of *Ctbp1/2* rescues abnormal DNA methylation by *Dppa2* KO.**
247 (A) K-means clustering analysis of the genes of Cluster 5 (Figure 2C) using published
248 single-cell RNA-Seq data of early mouse embryos (Gerlinger et al., 2014). (B) Heat map
249 showing relative expression of all differentially expressed genes between WT and *Dppa2*
250 KO or *Dppa2* KO and *Ctbp1/2*/*Dppa2* TKO ESCs. (C) The Venn diagram shows the
251 overlapping between the down-regulated genes by *Dppa2* KO in WT ESCs and the
252 upregulated genes by *Ctbp1/2* DKO in *Dppa2* KO ESCs. (D) Bisulfite sequencing of the
253 promoter region of *Dazl* and *LINE1* in WT, *Dppa2* KO, *Ctbp1/2* DKO and
254 *Ctbp1/2*/*Dppa2* TKO ESCs under Serum LIF and 2i LIF. ***$P < 0.001$, **$P < 0.005$, *$P$
255 $< 0.05$ (Mann-Whitney U-test). (E) Venn diagram showing the overlapping between the
256 region with CpG island in which DNA methylation is elevated by *Dppa2* KO in ESCs
257 and the region with CpG island in which Ctbp1/2 DKO reduces DNA methylation in
258 *Dppa2* KO ESCs. (F) Gene track representing the methylation levels derived from RRBS
259 at *Dazl*, *Ddx4*, *Ric3* and *Twist2* in WT, *Dppa2* KO, *Ctbp1/2* DKO and *Ctbp1/2*/*Dppa2*
260 TKO ESCs under 2iL. (G) Heat map showing the qRT-PCR data of the regulators of DNA
261 methylation and DNA methyltransferases in WT, *Dppa2* KO, *Ctbp1/2* DKO and
262 *Ctbp1/2*/*Dppa2* TKO ESCs. (H) Western blotting of UHRF1 and TUBULIN in WT,
263 *Dppa2* KO, *Ctbp1/2* DKO and *Ctbp1/2*/*Dppa2* TKO ESCs. (I) Gene track representing
264 the DPPA2-binding peaks reconstructed using published data (Engelen et al., 2015) and
265 the methylation levels derived from RRBS upstream of *Pramel7* in WT, *Dppa2* KO,
266 *Ctbp1/2* DKO and *Ctbp1/2*/*Dppa2* TKO ESCs.

267

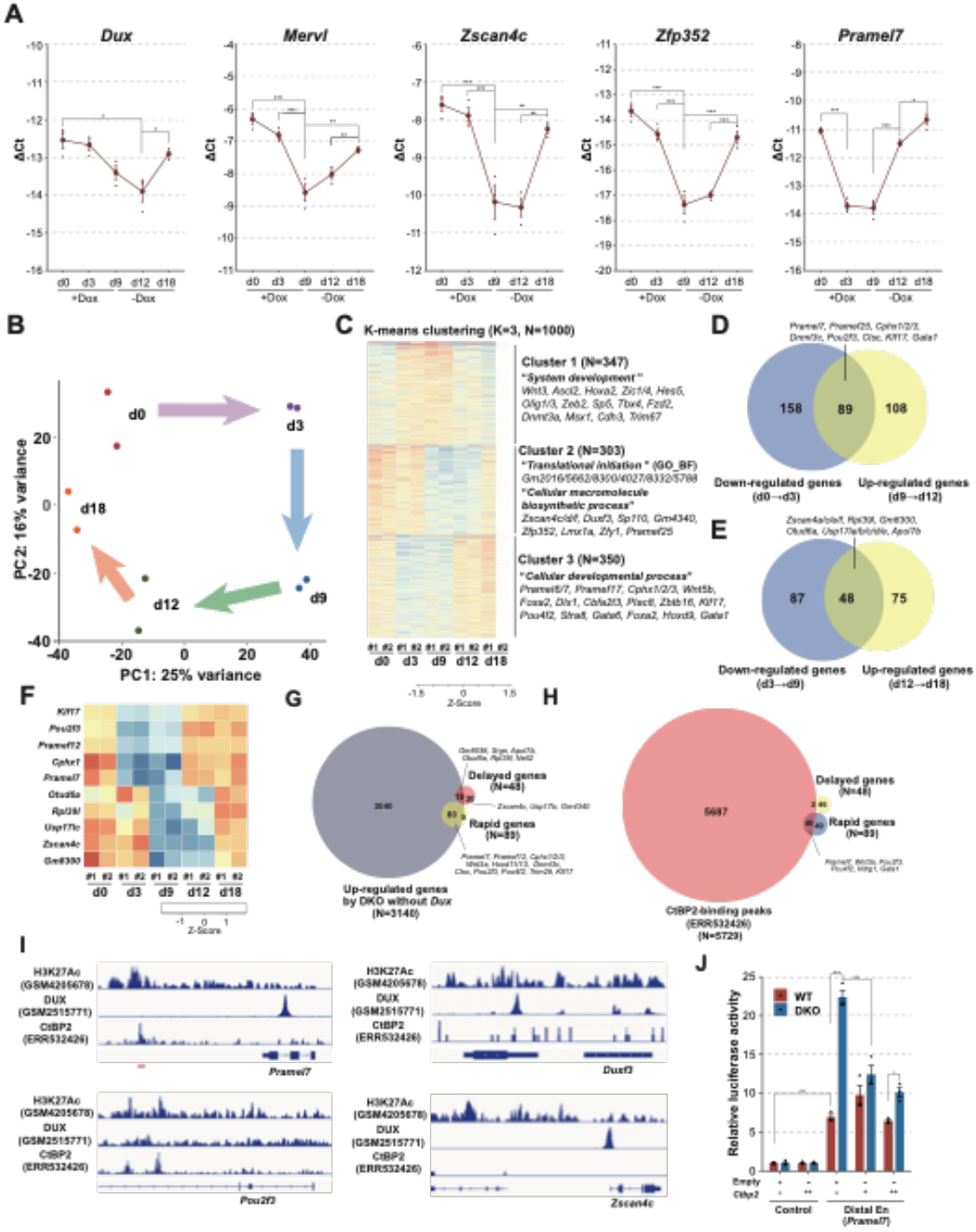
268 ***The loss of Ctbp1/2 controls two genetic cascades for ZGA: DUX-dependent and DUX-***
269 ***independent***

270 Next, to identify the direct target of CtBP1/2 controlling the emergence of MERVL-
271 EGFP-positive cells, we established *Ctbp1/2* DKO ESCs carrying the expression unit of
272 *Ctbp2* regulated by the doxycycline administration. We traced the expression dynamics
273 of some key genes associated with zygotic genome activation (ZGA) using qRT-PCR
274 analysis. *Dux*, *Mervl*, *Zscan4c* and *Zfp352* were significantly decreased from 3 to 9 days

275 after induction of *Ctbp2* in *Ctbp1/2* DKO ESCs, while these genes were upregulated
276 from 3 to 9 days (d12-d18) after the depletion of *Ctbp2* expression (Figure 4A). In
277 contrast, *Pramel7* was sharply downregulated (d0-d3) or upregulated (d9-d12) after
278 *Ctbp2* induction or depletion in *Ctbp1/2* DKO ESCs. These findings suggested that 2-
279 cell associated genes upregulated by the loss of *Ctbp1/2* are classified into two
280 categories: CtBP1/2-indirect target genes such as *Dux*, *Mervl*, *Zscan4c* and *Zfp352*, and
281 CtBP1/2-direct target genes such as *Pramel7*. To identify the direct and indirect targets
282 of CtBP2, we performed RNA-Seq using *Ctbp1/2* conditional KO (cKO) ESCs.
283 Principal component analysis (PCA) revealed that the global changes of gene
284 expression profiled by the short-term induction (d0-d3) or depletion (d9-d12) of *Ctbp2*
285 in *Ctbp1/2* DKO ESCs contributed PC1 axis in the opposite direction (Figure 4B). On
286 the other hand, the global gene expression changes profiled by the long-term induction
287 (d3-d9) or depletion of *Ctbp2* (d12-d18) in *Ctbp1/2* DKO ESCs contributed to the PC2
288 axis in the opposite direction. K-means clustering analysis revealed that CtBP2 rapidly
289 and positively regulated the genes associated with system development, including
290 *Ascl2*, *Zic1/4* and *Olig1/3* in ESCs (Figure 4C). Interestingly, the downregulated genes
291 by CtBP1/2 were separated into two groups. The genes in cluster 2 containing major
292 2C-associated genes, including *Dux*, *Zscan4*, *Sp110*, *Zfp352* and *Gm4340,* were
293 repressed 3 to 9 days after *Ctbp2* induction (Figure 4C). Other ZGA-associated genes,
294 *Pramel6/7* and *Cphx1/2/3,* were in cluster 3, in which the genes were rapidly
295 downregulated or upregulated by the loss or gain of *Ctbp2*. Next, to extract the direct or
296 indirect targets of CtBP2, we compared a group of genes rapidly or slowly repressed by

297 *Ctbp2* induction with a group of genes that are rapidly or delay activated by *Ctbp2*
298 depletion, respectively (Figure 4D, E). We defined overlapped genes as the candidates
299 for the direct target (rapid response genes) and indirect target of CtBP2 (delayed
300 response), and we displayed the RNA-Seq data of parts of these genes in the heatmap
301 (Figure 4F). Almost all rapid response genes and over half of the delayed response
302 genes were upregulated by the loss of CtBP2 without DUX function (Figure 4G). Next,
303 we overlapped the genes that CtBP2 is localized at around the transcription start site
304 (TSS) with rapid and delayed response genes using a Venn diagram using public ChIP-
305 seq data of CtBP2 in ESCs (Network et al., 2015) (Figure 4H). CtBP2-binding peaks
306 around the TSS of nearly half of the direct candidate genes for CTBP2, whereas
307 CTBP2-binding peaks have only been observed around the TSS of two indirect target
308 genes for CTBP2. Next, we displayed the ChIP-Seq track of DUX- and CtBP2-binding
309 peaks and H3K27Ac, a typical enhancer mark (Hendrickson et al., 2017) (Network et
310 al., 2015) (Zhang et al., 2020), at around the *Pramel7*, *Pou2f3* (rapid response genes)
311 and *Duxf3*, *Zscan4*c (delayed response genes) (Figure 3I). CtBP2 was localized within
312 the H3K27Ac-enriched distal enhancer of *Pramel7* and *Pou2f3*. In contrast, CtBP2-
313 binding peaks were absent, and DUX-binding peaks were localized around the TSS of
314 *Duxf3* and *Zsacn4c*. To determine the direct effect of CtBP2 at the distal enhancer of
315 *Pramel7*, we analyzed the effect of CtBP2 overexpression against the transcriptional
316 activity driven by the distal enhancer of *Pramel7* in WT and *Ctbp1/2* DKO ESCs
317 (Figure 4J). Luciferase activity was activated in *Ctbp1/2* DKO ESCs compared with
318 those in WT ESCs, and the CtBP2 overexpression completely diminished this

activation. In summary, our findings indicate that the loss of CtBP1/2 controls two pathways regulating 2C programs: one mediated by *Dux* activation (*Zscan4c* and *Usp17lc*) and the other directly activated by the loss of CtBP1/2 (*Prame7* and *Pou2f3)*.



323 **Figure 4. The loss of *Ctbp1/2* controls two genetic cascades for ZGA.**
324 (A) RT-qPCR analysis of *Dux*, *Mervl, Zscan4c Zfp352* and *Pramel7* in *Ctbp1/2* DKO
325 ESCs without Dox or with Dox. The ΔCt of each gene was calculated by subtracting the
326 average Ct of Gapdh from that of each gene Ct. Means of ΔCt were calculated from
327 three biological triplicates. ***$P < 0.001$, **$P < 0.005$, *$P < 0.05$ (Tukey-Kramer
328 multiple comparison test). (B) Principal component analysis (PCA) of *Ctbp1/2* DKO
329 ESCs without Dox or with Dox. (C) K-means clustering analysis of *Ctbp1/2* DKO ESCs
330 without Dox or with Dox using the top 1000 differentially expressed genes. (D) Venn
331 diagram shows the overlapping of down-regulated genes three days after Ctbp2
332 induction and upregulated genes three days after *Ctbp2* removal. (E) Venn diagram
333 showing the overlapping of down-regulated genes whose expression is down-regulated
334 from 3 days to 9 days after *Ctbp2* induction and upregulated genes whose expression is
335 upregulated from 3 days to 9 days after *Ctbp2* removal. (F) Heatmap showing the
336 selected genes from K-means clustering analysis (C). Venn diagram shows the
337 overlapping of upregulated genes by Ctbp1/2 DKO in Dux KO ESCs with rapid or
338 gradual responses against *Ctbp2* expression. (H) Venn diagram shows the overlapping
339 of the genes with CtBP2-binding with rapid or gradual responses against *Ctbp2*
340 expression. (I) Gene track of ChIP-Seq data of H3K27Ac (GSM4205678), DUX
341 (GSM2525771) and CtBP2 (ERR532426) around *Pramel7*, *Pou2f3*, *Duxf3* and *Zscan4c*.
342 The red bar indicates the region used for luciferase assay in (J). (J) The effect of CtBP2
343 overexpression for the enhancer activity of *Pramel7* distal enhancer (I) in *Ctbp1/2* DKO
344 ESCs. Relative luciferase activity was calculated with the values of the empty vector as
345 1. Means of relative luciferase activity were calculated from three biological triplicates.
346 ***$P < 0.001$, **$P < 0.005$, *$P < 0.05$ (Tukey-Kramer multiple comparison test).
347

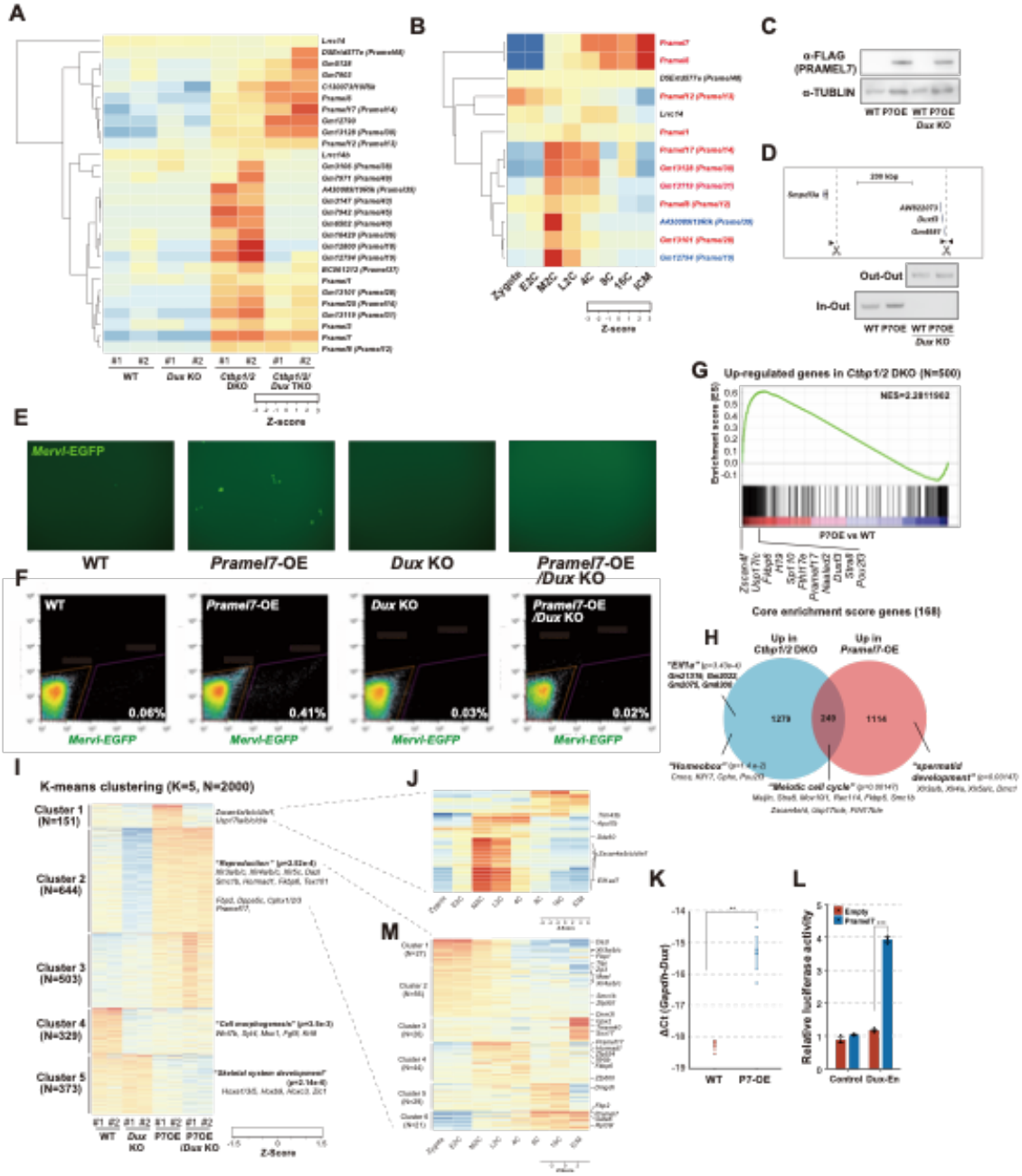
348 ***Prame family control genetic program of ZGA: DUX-dependent and -independent***

349 *Pramel7* belongs to the *Prame* gene family, which is restricted to the testis and various
350 cancers (Kern et al., 2021). Because *PRAME* is one of the most amplified gene families
351 in mammals and is considered the third largest gene family in the mouse genome
352 (Church et al., 2009), we focused on the expression of the *Prame* gene family in

353 *Ctbp1/2* DKO ESCs. We collected the amino acid sequences of the PRAME family
354 protein annotated on the mm10 genome sequence and created the phylogenetic tree
355 (Supplementary Figures 3A and B). The *Prame* family genes were clustery located at
356 six distinct chromosomes, and high similarity of PRAME family proteins tended to be
357 found on the same chromosome. Next, we selected the expression levels of *Prame*
358 family genes in WT, *Du*x KO, *Ctbp1/2* DKO and TKO from RNA-Seq data and
359 displayed the hierarchical clustering data in the heatmap (Figure 5A). Most *Prame*
360 family genes, except for *Lrrc14*, *Lrrc14b*, and *D5Ertd577e* (*Pramel48*), were
361 upregulated in *Ctbp1/2* DKO ESCs. According to gene expression patterns, the genes of
362 the PRAME family can be grouped into four categories. In the first category, including
363 *Lrrc14*, there are genes whose expression is unaffected by *Ctbp1/2* DKO. In the second
364 category, including *Pramel6* and *Pramel14*, the genes upregulated by *Ctbp1/2* DKO
365 were further elevated by *Dux* KO. In the third category, including *Pramel19* and
366 *Pramel39*, the genes upregulated by *Ctbp1/2* DKO were completely repressed by *Dux*
367 KO (DUX-dependent *Prame* family genes). In the fourth category, including *Pramel7*
368 and *Pramel12*, the genes upregulated by *Ctbp1/2* DKO were unaffected by *Dux* KO
369 (DUX-independent *Prame* family genes). Next, we analyzed the expression dynamics
370 of *Prame* family genes in the early mouse embryo using public scRNA-Seq data
371 (Gerlinger et al., 2014). *Pramel19* and *Pramel39*, DUX-dependent *Prame* family genes,
372 were transiently upregulated at the mid 2-cell stage after transient upregulation of *Dux*
373 at the early 2-cell stage. DUX-independent *Prame* family genes were subdivided into
374 three categories. The first category, including *Pramel6* and *Pramel7*, adjacent to

375 chromosome 2, were gradually upregulated from mid 2-cell to peak expression at ICM.
376 The genes, including *Pramel48* and *Pramel13* in the second category, were not
377 upregulated by ZGA. The genes belonging to the third category, including *Pramel14*
378 and *Pramel31,* were activated soon after ZGA and continued to be expressed until 4 cell
379 embryos. Because it has been reported that *Pramel7*, which is upregulated by the loss of
380 *Ctbp1/2* and ZGA not depending on DUX function, is critical for early mouse
381 development (Graf et al., 2017a), we focused on the PRAMEL7 functions on the
382 emergence of the 2CLCs from pluripotent stem cells. We generated ESCs with forced
383 expression of *Pramel7* with or without *Dux* KO (Figure 5C, D). *Pramel7*
384 overexpression expanded 2CLCs, and this expansion was completely cancelled by the
385 *Dux* KO (Figure 5E, F). Gene set enrichment analysis comparing WT and *Pramel7*-
386 overexpressing ESCs using the upregulated genes by the loss of *Ctbp1/2* showed that
387 the upregulated genes by the loss of *Ctbp1/2* tended to be upregulated by *Pramel7*
388 overexpression (Figure 5G). Germline-specific genes and ZGA-associated genes were
389 consistently upregulated by the loss of *Ctbp1/2* and *Pramel7* overexpression (Figure
390 5H). Some sets of ZGA-associated genes, including *Crxos*, *Cphx1* and *Klf17,* were only
391 upregulated by the loss of *Ctbp1/2*, suggesting that the loss of *Ctbp1/2* controls two
392 genetic cascades: PRAMEL7-dependent and PRAMEL7-independent pathway. Next,
393 we separated genes into five categories using K-means clustering analysis based on the
394 gene expression profiles in WT, *Dux* KO, *Pramel7*-OE and *Pramel7*-OE/*Dux* KO ESCs
395 (Figure 5I). The first category included the ZGA-associated genes whose expression
396 was upregulated by PRAMEL7 depending on DUX functions. Most genes belonging to

397 the first cluster were transiently upregulated at mid 2-cell embryos soon after the
398 transient upregulation of *Dux,* such as the *Zscan4* family; the rest genes were
399 upregulated after eight-cell embryos (Figure 5J). *Dux* expression was elevated in
400 PRAMEL7-overexpressing ESCs (Figure 5K), and enhancer activity was also
401 upregulated by PRAMEL7 overexpression (Figure 5L), suggesting that PRAMEL7
402 controls genetic program for 2CLCs through the upregulation of *Dux* expression in
403 ESCs. Next, we focused on cluster 2, in which the genes were downregulated by *Dux*
404 KO and this downregulation was diminished by PRAMEL7 overexpression (Figure 5I).
405 Cluster 2 was subdivided into two groups based on the gene expression dynamics in
406 early mouse development (Figure 5M). The first groups, clusters 1 and 2, contained the
407 maternal factor, including *Dazl*, *Zp3,* and *Mael*. The second group consisted of clusters
408 3, 4, 5 and 6, in which genes were upregulated after ZGA, including *Pramef17*, *Zp600*
409 and *Gata6*.




410

411 **Figure 5. PRMAEL7 controls two genetic cascades for ZGA.**

412 (A) The heatmap showing the hierarchical clustering of the expressions of Preme family
413 genes using RNA-Seq data in WT, *Dux* KO, *Ctbp1/2* DKO and *Ctbp1/2*/*Dux* KO ESCs.
414 (B) The heatmap shows the hierarchical clustering of the expressions of Preme family
415 genes using public single-cell RNA-Seq data in early mouse embryos (Gerlinger et al.,
416 2014). (C) Western blotting anti-FLAG (FLAG-PRAMEL) antibody using lysates of WT,

417 PRAMEL7-overexpressing, *Dux* KO and PRAMEL7-overexpressing *Dux* KO ESCs. (D)
418 Genotype PCR to confirm the large deletion of the *Dux* cluster. (E) Dark fields image
419 visualizing *Mervl*-EGFP signals in WT, PRAMEL7-overexpressing, *Dux* KO and
420 PRAMEL7-overexpressing *Dux* KO ESCs. (F) FACS analysis measuring the percentage
421 of *Mervl*-EGFP-positive cells in WT, PRAMEL7-overexpressing, *Dux* KO and
422 PRAMEL7-overexpressing *Dux* KO ESCs. (G) Gene set enrichment analysis (GSEA)
423 using the upregulated genes by the *Ctbp1/2* disruption to compare the gene expression of
424 PRMAEL7-overexpressing ESCs with WT ESCs. (H) Venn diagram showing the
425 overlapping of the upregulated genes by the *Ctbp1/2* disruption in ESCs with the
426 upregulated genes by PRAMEL7 overexpression in ESCs. (I) K-means clustering (K=5)
427 of the gene expression in WT, PRAMEL7-overexpressing, *Dux* KO, and PRAMEL7-
428 overexpressing *Dux* KO ESCs using top 2000 differentially expressed genes. (J and M)
429 K-means clustering analysis of the genes of Cluster 1 and Cluster 4 using published
430 single-cell RNA-Seq data of early mouse embryo(Gerlinger et al., 2014) (K) RT-qPCR
431 analysis of *Dux* in WT and PRAMEL7-overexpressing ESCs. ΔCt value was calculated
432 by subtracting the Ct value of *Dux* from that of *Gapdh*. The means of ΔCt was calculated
433 from three biological triplicates. ***$P < 0.001$, **$P < 0.005$, *$P < 0.05$ (Tukey-Kramer
434 multiple comparison test). (L) The effect of PRAMEL7 overexpression against luciferase
435 activity regulated by the *Dux* enhancer region. Relative luciferase activity was calculated
436 by setting the luciferase activity of the empty vector as 1. The means of relative luciferase
437 activity was calculated from three biological triplicates. ***$P < 0.001$, **$P < 0.005$, *$P$
438 $< 0.05$ (Tukey-Kramer multiple comparison test).

439

440 ***Dux controls a part of genetic cascades for the ZGA program through the PRAMEL7***
441 ***functions***

442 We generated *Pramel7* KO ESCs and compared the gene expression profiles of WT,
443 *Pramel7* KO, *Dux* KO, *Pramel7*-OE and *Pramel7*-OE/*Dux* KO ESCs (Supplementary
444 Figure 1D). Hierarchical clustering and principal component analyses showed similar
445 gene expression profiles of *Pramel7* KO ESCs and *Dux* ESCs (Figure 6A, B).

446 Furthermore, the upregulated or down-regulated genes by *Dux* KO tended to be
447 upregulated or down-regulated by *Pramel7* KO based on the GSEA analysis (Figure
448 6C, D). Venn diagram using the upregulated or downregulated genes by *Dux* KO and
449 *Pramel7* showed the shared genes and independent genes regulated by DUX and
450 PRAMEL7 (Figure 6E, F). The genes associated with ZGA were separated into DUX-
451 dependent (*Zscan4* family) and DUX/PRAMEL7-dependent (*Cphx* and *Zfp534/600*). K-
452 means clustering analysis of WT, *Dux* KO, *Pramel7* KO, *Pramel7*-OE and *Pramel7*-
453 OE/*Dux* KO showed that the commonly downregulated genes by *Dux* KO and *Pramel7*
454 KO (Cluster 3 and 5), including (*Zfp600*, *Pou2f2* and *Fkbp6*, *Xlr3c*) were upregulated
455 by PRAMEL7 overexpression in *Dux* KO ESCs (Figure 6G and H). *Pramel7* expression
456 was significantly reduced in *Dux* KO ESCs (Figure 6I), and the down-regulated genes
457 in *Dux* KO ESCs were substantially upregulated by PRAMEL7 overexpression (Figure
458 6J and K), which indicates that the parts of *Dux* KO phenotypes depend on the
459 reduction of *Pramel7* expression.

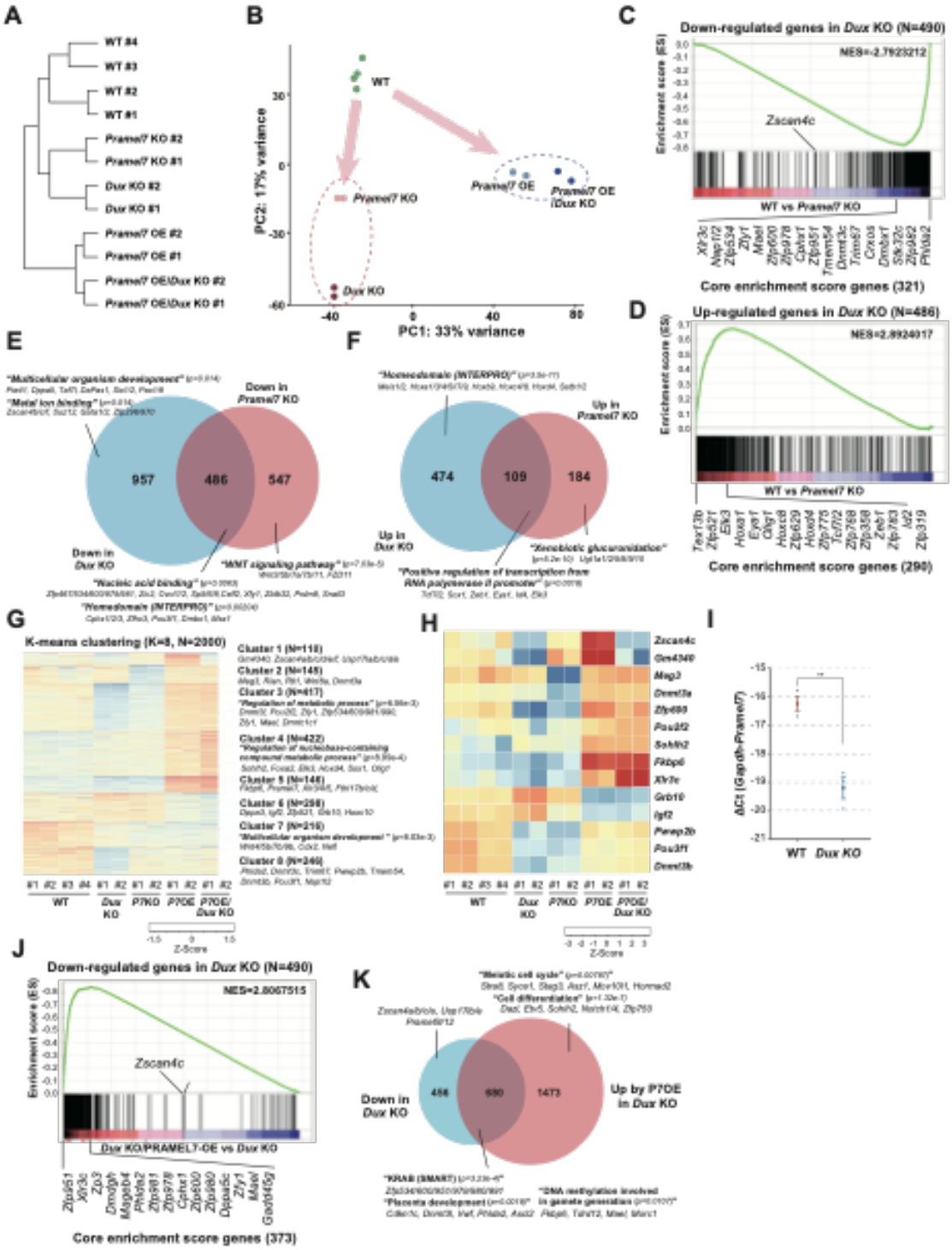



**Figure 6. DUX regulates parts of the genetic program for ZGA by activating *Pramel7*.** (A) Hierarchical clustering analysis of WT ESCs, *Pramel7* KO ESCs, *Dux* KO ESCs, PRMAEL7-overexpressing ESCs and PRMAEL7-overexpressing *Dux* KO ESCs based on the gene expression. (B) Principal component analysis of WT ESCs, *Pramel7* KO ESCs, *Dux* KO ESCs, PRMAEL7-overexpressing ESCs and PRMAEL7-overexpressing

*Dux* KO ESCs based on the gene expression. (C) Gene set enrichment analysis (GSEA) using the down-regulated genes by the *Dux* KO to compare the gene expression of *Pramel7* KO ESCs with WT ESCs. (D) Gene set enrichment analysis (GSEA) using the upregulated genes by the *Dux* KO to compare the gene expression of *Pramel7* KO ESCs with WT ESCs. (E) Venn diagram showing the overlapping of the down-regulated genes by *Dux* KO in ESCs with the down-regulated genes by *Pramel7* KO in ESCs. (F) Venn diagram showing the overlapping of the upregulated genes by *Dux* KO in ESCs with the upregulated genes by *Pramel7* KO in ESCs. (G) K-means clustering (K=8) of the gene expression in WT ESCs, *Dux* KO ESCs, *Pramel7* KO ESCs, *Prame7*-overexpressing ESCs and *Prame7*-overexpressing *Dux* KO ESCs using the top 2000 differentially expressed genes. (H) Heatmap showing the selected genes from K-means clustering analysis (G). (I) qRT-PCR analysis of *Pramel7* in wild type and *Dux* KO ESCs. ΔCt value was calculated by subtracting the Ct value of *Pramel7* from that of *Gapdh*. The means of ΔCt was calculated from three biological triplicates. ***$P < 0.001$, **$P < 0.005$, *$P < 0.05$ (Tukey-Kramer multiple comparison test). (J) Venn diagram showing the overlapping of the down-regulated genes by *Dux* KO in ESCs with the upregulated genes by PRAMEL7 overexpression in *Dux* KO ESCs. (K) Gene set enrichment analysis (GSEA) using the down-regulated genes by the *Dux* KO to compare the gene expression of PRMAEL7-overexpressing in *Dux* KO ESCs with *Dux* KO ESCs.

487 **Discussion**

488 In this study, we revealed that the disruption of *Ctbp1/2* in ESCs activates DUX-
489 dependent and independent molecular cascades to establish totipotent cells. Most DUX-
490 dependent genes activated by *Ctbp1/2* DKO are transiently expressed at around 2C stage
491 embryo, while DUX-independent genes activated by *Ctbp1/2* DKO are categorized into
492 two groups: Maternal expressed genes and the genes upregulated after major ZGA, some
493 of which are regulated by DPPA2. PRAMEL7, which is directly activated by the
494 disruption of *Ctbp1/2* in a DUX-independent manner, controls DUX-dependent genetic
495 cascades through DUX activation and DUX-independent genetic cascades. We present
496 new insights into the DUX-independent molecular pathways ensuring the early mouse
497 embryo.

498 Our data clearly showed that the reduction of CtBP protein levels in ESCs activates a part
499 of the genetic program regulating the development of early mouse embryos. scRNA-seq
500 data of early mouse embryos shows that a large amount of *Ctbp2* mRNA is maternally
501 deposited in unfertilized eggs and significantly reduced after fertilization (Gerlinger et al.,
502 2014). In contrast, the protein level of CtBP2 was consistent between the 2CLCs that
503 emerged from pluripotent stem cells and pluripotent stem cells, indicating that the
504 expression levels of CtBP proteins are quite different between 2C embryos and 2CLCs.
505 This difference may explain the dependence of the emergence of 2CLC and early mouse
506 development on DUX functions. However, the molecular mechanism of the clearance of
507 *Ctbp2* mRNA after fertilization is currently unclear.

508 Transcriptome analysis of several knockout lines of ESCs provided evidence that the

509 activated genes by the disruption of *Ctbp1/2* are subdivided into three groups: both DUX
510 and DPPA2-dependent genes, DPPA2-dependent genes and both DUX and DPPA2-
511 independent genes. DUX-and DPPA2-dependent genes were enriched in the transient
512 upregulated genes at 2C embryo, including the *Zscan4* family. In contrast, DUX and
513 DPPA2-independent genes contained maternal factors and the upregulated genes after
514 ZGA in the early mouse embryo. A previous study has shown that totipotent-like stem
515 cells (TLSCs) induced by the chemical inhibitors of DOT1L, KDM5B and G9a/GLP
516 expressed a large number of minor and major ZGA genes and maternal genes (Yang et
517 al., 2022). The re-establishment of broad H3K4me3 domains induced by the inhibition of
518 KDM5B is thought to be involved in the upregulation of maternal genes. However,
519 because TLSCs were not established using *Dux* KO ESCs, the role of DUX in the
520 transition from ESCs to TLSCs and the upregulation of maternal genes are not largely
521 unknown. Our data demonstrate that the upregulation of maternal genes by *Ctbp1/2* DKO
522 does not depend on DUX function and is partially regulated by *Pramel7* induction.
523 *Prame* is one of the most amplified gene families in mammals, and the mouse *Prame*
524 gene family is the third largest family in the genome (Chang et al., 2011). We presented
525 that the disruption of *Ctbp1/2* activates many *Prame* family genes in ESCs, some of which
526 are upregulated by DUX function and the rest directly by CtBP1/2. Interestingly, *Prame*
527 genes that depend on DUX are transiently upregulated at 2-cell embryos, while *Prame*
528 genes independent of DUX are upregulated after major ZGA. Many *Prame* genes are
529 repressed soon after induction of CtBP2 in *Ctbp1/2* DKO ESCs, indicating that CtBP1/2
530 directly repress *Prame* family genes in ESCs. However, the positive regulators of *Prame*

531 family genes are largely unknown. Expression of *Prame* family genes has been
532 discovered in various cancer and germline cells (Kern et al., 2021), and the expression
533 pattern of *Prame* family genes is similar to that of *Pou5f1* (Oct4) (Bortvin et al., 2003),
534 suggesting that transcriptional network centred upon OCT4 is a positive regulator of
535 *Prame* family genes.
536 Although DUX is a master regulator for the transition from pluripotent stem cells into
537 2CLC, *Dux* KO embryos are normally born and fertile. The data presented here show that
538 most of the *Dux* KO phenotype is rescued by PRAMEL7 overexpression in ESCs. It has
539 been reported that *Prame* family genes are upregulated from the early phase of 2CLC
540 emergence traced by the *Zscan4* and *Mervl* reporter system based on the scRNA-seq
541 (Rodriguez-Terrones et al., 2018b)(Eckersley-Maslin et al., 2016) and MERVL-positive
542 2CLCs are eliminated by *Dux* KO (De Iaco et al., 2017). Based on these findings, we
543 speculate that DUX controls two major genetic cascades for the development of early
544 mouse embryos: direct transcriptional regulation and PRAME-mediated transcriptional
545 regulation. *Dux* KO ESCs showed a significant reduction of *Pramel7* transcription, while
546 the expression of *Pramel7* is normally activated in *Dux* KO embryos at the mid 2-cell
547 stage (Guo et al., 2019). These findings suggest that the DUX-independent mechanism
548 activates the *Prame* family *in vivo*, and this mechanism might be induced in ESCs by the
549 disruption of *Ctbp1/2*.
550
551 Our results provide new insights into understanding DUX-independent genetic cascades,
552 ensuring the development of early mouse embryos. Although the visualization of 2CLC

553 marked by the transcription of MERVL contributed to identifying the regulators for the
554 transition between pluripotent stem cells and 2CLCs, the knockout phenotypes of the
555 identified factors were functionally dependent on DUX. Developing the visualization
556 system for new 2CLC regulating by DUX-independent genetic cascades in ESCs will be
557 needed.
558
559

560 **Material and Methods**

561 **Cell Culture**

562 E14tg2a ESCs (RIKEN Cell Bank) were cultured in Glasgow minimum essential medium
563 (GMEM; Wako) containing 10% fetal calf serum (FCS; Invitrogen), 1 mM glutamine
564 (Wako), nonessential amino acids (Wako) and 500 μM Monothioglycerol Solution
565 (Wako). The culture was supplemented with 1000 U/ml of leukaemia inhibitory factor
566 (LIF) (Wako) in the absence of feeder cells. ESCs were cultured in N2B27 medium with
567 2i (PD0325901, 0.4 μM: WAKO, CHIR99021, 3 μM: Sigma-Aldrich).

568

569 **Generation of knockout ESCs using a CRISPR-Cas9 system**

570 Guide RNAs for *Ctbp1/2*, *Dux* and *Pramel7* were generated in pX330-U6- Chimeric_BB-
571 CBh-hSpCas9 (Addgene plasmid #42230; deposited by Feng Zhang; Cong et al., 2013).
572 The guide RNA sequences are listed in Table S1. The pX330 was co-transfected with
573 pCAG-EGFP into E14tg2a with MERVL::TurboGFP (Addgene plasmid #69072;
574 deposited by Maria-Elena Torres-Padilla; (Ishiuchi et al., 2015). The GFP-positive
575 fraction was sorted by Cell Sorter (No. SH800; Sony Biotechnology, Inc., Tokyo, Japan).
576 Genomic DNA deletions were sequenced (Supplementary Figure 1).

577

578 **qRT-PCR and RNA-Seq**

579 Total mRNA was purified with a PureLink$^{TM}$ RNA Mini Kit (Ambion, Inc.). For
580 quantitative RT-PCR, cDNA was synthesized by ReverTra Ace® qPCR RT master Mix
581 (Toyobo), and then analyzed by qPCR with FastStart SYBR Green Master Mix (Roche)

582 using a specific primer set for each gene (Supplementary Table X). RNA-seq libraries
583 were generated using Collibri™ Stranded RNA library prep for illumina™ Systems
584 (ThermoFisher Scientific) and sequenced using HiSeq X™ Ten (Illumina) (150-bp
585 paired-ends reads).
586
587 **Data analysis of RNA-Seq**
588 Total mRNA was purified with a PureLink™ RNA Mini Kit (Ambion, Inc.). RNA-seq
589 libraries were generated using Collibri™ Stranded RNA library prep for illumina™
590 Systems (ThermoFisher Scientific) and sequenced using HiSeq X™ Ten (Illumina).
591 Reads were aligned by Hisat2 to the Mus musculus genome (mm10). Raw read counts of
592 genes were calculated by R package FeatureCounts using mm10 KnownGenes. GTF from
593 UCSC for annotation. Normalized count data, principal component and K-means
594 clustering analysis were exported by the interactive Expression and Pathway analysis
595 (iDEP) webserver using the count data or normalized expression data of early mouse
596 embryos (Ge et al., 2018). Gene set enrichment analysis (GSEA) was used to determine
597 whether upregulated genes by *Ctbp1/2* knockouts were enriched for the upregulated genes
598 in 2-cell embryos or 2CLCs as described previously (Macfarlan et al., 2012b). Venn
599 diagrams showing overlapping genes were constructed by BioVenn, a web application
600 (Hulsen et al., 2008).
601
602 **Bisulfite sequencing of specific genes and RRBS**

603 Genomic DNA was purified using Wizard® SV Genomic DNA Purification System
604 (Promega). For bisulfite sequencing of specific genes, the unmethylated cytosine of the
605 genomic DNA was converted to uracil using EpiTect Bisulfite Kit (QIAGEN) according
606 to the manufacturer's instructions. Bisulfited DNA was amplified by the region-specific
607 primer, followed by the sequencing. According to the manufacturer's instructions, DNA
608 libraries for RRBS analysis were prepared using the Premium RRBS kit V2 and
609 sequenced using HiSeq X™ Ten (Illumina) (150-bp paired-ends reads). RRBS reads
610 were trimmed using trim galore (www.bioinformatics.babraham.ac.uk/projects/
611 trim_galore/) with the RRBS option invoked. Bismark was used to align reads to UCSC
612 version mm10 of the mouse genome. CpG methylation was extracted using
613 bismark_methylation_extractor and displayed by Broad Institute's Integrative Genomic
614 Viewer (IGV) (Robinson et al., 2011).
615
616 **Immunofluorescence analysis**
617 Cells were fixed by 4% paraformaldehyde at 4°C for 20 min, followed by incubation with
618 PBS containing 0.5% Triton-X at 4°C for 5 min. Permeabilized cells were incubated with
619 a primary antibody solution containing anti-GFP (Nacarai Tesque, 04404-84, 1:500) with
620 anti-ZSCAN4 (Merk Millipore, AB4340, 1:1000) or anti-MuERVL-Gag (Epigentek, A-
621 2801-050, 1:250) or anti-OCT4 (Santa Cruz, Sc-8628, 1:500) or anti-CtBP2 (BE
622 Biosciences, 612044, 1:500) followed by secondary antibody solution containing anti-
623 Rat IgG Dylight™ 488 antibody (Rockland, 612-141-120, 1:500) or anti-Rabbit IgG
624 Dylight™ 549 antibody (KPL, 042-04-15-06, 1:500) or anti-Mouse IgG Dylight® 550

625 antibody (Abcam, ab96876, 1:500) or anti-Goat IgG Dylight™ 549 antibody (Rockland,
626 605-742-002, 1:500) with 1 μg/ml of DAPI solution (DOJINDO, 340-07971).
627
628 **Luciferase assay**
629 The enhancer candidate region of *Pramel7* or *Dux* was cloned into pGL4.23, driving the
630 *luciferase* gene. Wild-type or *Ctbp1/2* DKO ESCs were transfected with the reported
631 constructs with *Ctbp2* or *Pramel7* expression vector using HilyMax (DOJINDO). Cells
632 were harvested 36 hours after transfection, and the luciferase activities were measured
633 using the Dual-luciferase Reporter Assay System (Promega).
634
635 **Western blotting analysis**
636 Cells were lysed by boiling SDS sample buffer (WAKO). Before lysed proteins were
637 applied to polyacrylamide-SDS gels, 2-mercaptoethanol was added to denature the
638 proteins. The lysed proteins were separated on polyacrylamide-SDS gels, blotted on a
639 polyvinylidene fluoride membrane, and probed using the following primary antibodies:
640 anti-UHRF1 (MBL, D289-3, 1:1000) or anti-TUBLIN (SIGMA, T5168, 1:5000).
641 Following the primary antibody reaction, the membrane were incubated with horseradish
642 peroxidase (HRP)-coupled secondary antibodies. All antibodies were detected with
643 Immunostar LD (WAKO).
644
645

## References


**Bortvin, A., Eggan, K., Skaletsky, H., Akutsu, H., Berry, D. L., Yanagimachi, R., Page, D. C. and Jaenisch, R.** (2003). Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. *Development* **130**, 1673–1680.

**Chang, T. C., Yang, Y., Yasue, H., Bharti, A. K., Retzel, E. F. and Liu, W. S.** (2011). The expansion of the PRAME gene family in Eutheria. *PLoS One* **6**,.

**Chen, Z. and Zhang, Y.** (2019). Loss of DUX causes minor defects in zygotic genome activation and is compatible with mouse development. *Nat Genet*.

**De Iaco, A., Planet, E., Coluccio, A., Verp, S., Duc, J. and Trono, D.** (2017). DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat Genet* **49**, 941–945.

**Eckersley-maslin, M. A., Parry, A., Blotenburg, M., Krueger, C., Ito, Y., Nila, V., Franklin, R., Narita, M., Santos, C. S. D. and Reik, W.** facilitates multi-lineage commitment.

**Eckersley-Maslin, M. A., Svensson, V., Krueger, C., Stubbs, T. M., Giehr, P., Krueger, F., Miragaia, R. J., Kyriakopoulos, C., Berrens, R. V., Milagre, I., et al.** (2016). MERVL/Zscan4 Network Activation Results in Transient Genome-wide DNA Demethylation of mESCs. *Cell Rep* **17**, 179–192.

**Eckersley-maslin, M., Alda-Catalinas, C., Blotenburg, M. and Kreibich, E.** (2018). Dppa2 and Dppa4 directly regulate the Dux driven zygotic transcriptional programme. 1–26.

**Engelen, E., Brandsma, J. H., Moen, M. J., Signorile, L., Dekkers, D. H. W., Demmers, J., Kockx, C. E. M., Ozgür, Z., Van Ijcken, W. F. J., Van Den Berg, D. L. C., et al.** (2015). Proteins that bind regulatory regions identified by histone modification chromatin immunoprecipitations and mass spectrometry. *Nat Commun* **6**,.

**Ge, S. X., Son, E. W. and Yao, R.** (2018). iDEP: An integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics* **19**, 1–24.

**Gerlinger, M., Swanton, C., Greaves, M. and Maley, C. C.** (2014). Single-Cell RNA-Seq Reveals. **343**,.

**Graf, U., Casanova, E. A., Wyck, S., Dalcher, D., Gatti, M., Vollenweider, E.,**

**Okoniewski, M. J., Weber, F. A., Patel, S. S., Schmid, M. W., et al.** (2017a). Pramel7 mediates ground-state pluripotency through proteasomal-epigenetic combined pathways. *Nat Cell Biol* **19**, 763–773.

**Graf, U., Casanova, E. A., Wyck, S., Dalcher, D., Gatti, M., Vollenweider, E., Okoniewski, M. J., Weber, F. A., Patel, S. S., Schmid, M. W., et al.** (2017b). Pramel7 mediates ground-state pluripotency through proteasomal-epigenetic combined pathways. *Nat Cell Biol* **19**, 763–773.

**Gretarsson, K. H. and Hackett, J. A.** epigenome for development. *Nat Struct Mol Biol*.

**Grow, E. J., Weaver, B. D., Smith, C. M., Guo, J., Stein, P., Shadle, S. C., Hendrickson, P. G., Johnson, N. E., Butterfield, R. J., Menafra, R., et al.** (2021). p53 convergently activates Dux/DUX4 in embryonic stem cells and in facioscapulohumeral muscular dystrophy cell models. *Nat Genet* **53**, 1207–1220.

**Guo, M., Zhang, Y., Zhou, J., Bi, Y., Xu, J., Xu, C., Kou, X., Zhao, Y., Li, Y., Tu, Z., et al.** (2019). Precise temporal regulation of Dux is important for embryo development. *Cell Res* **7**,.

**Hendrickson, P. G., Doráis, J. A., Grow, E. J., Whiddon, J. L., Lim, J.-W., Wike, C. L., Weaver, B. D., Pflueger, C., Emery, B. R., Wilcox, A. L., et al.** (2017). Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat Genet*.

**Hu, Z., Tan, D. E. K., Chia, G., Tan, H., Leong, H. F., Chen, B. J., Lau, M. S., Tan, K. Y. S., Bi, X., Yang, D., et al.** (2020). Maternal factor NELFA drives a 2C-like state in mouse embryonic stem cells. *Nat Cell Biol* **22**, 175–186.

**Hulsen, T., de Vlieg, J. and Alkema, W.** (2008). BioVenn - A web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **9**, 1–6.

**Ishiuchi, T., Enriquez-Gasca, R., Mizutani, E., Bošković, A., Ziegler-Birling, C., Rodriguez-Terrones, D., Wakayama, T., Vaquerizas, J. M. and Torres-Padilla, M.-E.** (2015). Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat Struct Mol Biol*.

**Kern, C. H., Yang, M. and Liu, W. S.** (2021). The PRAME family of cancer testis antigens is essential for germline development and gametogenesis. *Biol Reprod* **105**, 290–304.

**Kreibich, E., Blotenburg, M., Alda-Catalinas, C., Reik, W., Krueger, C. and**

**Eckersley-Maslin, M.** (2019). Dppa2 and Dppa4 directly regulate the Dux-driven zygotic transcriptional program. *Genes Dev* **33**, 194–208.

**Macfarlan, T. S., Gifford, W. D., Driscoll, S., Lettieri, K., Rowe, H. M., Bonanomi, D., Firth, A., Singer, O., Trono, D. and Pfaff, S. L.** (2012a). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63.

**Macfarlan, T. S., Gifford, W. D., Driscoll, S., Lettieri, K., Rowe, H. M., Bonanomi, D., Firth, A., Singer, O., Trono, D. and Pfaff, S. L.** (2012b). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* 1–10.

**Network, R., Informatics, B., National, S., Resource, A., Sciences, B. and Science, C.** (2015). P LURIPOTENT S TEM C ELLS Ctbp2 Modulates NuRD-Mediated Deacetylation of in Active Embryonic Stem Cell Genes During Exit from Pluripotency. 2442–2455.

**Percharde, M., Lin, C., Yin, Y., Huang, B., Shen, X., Ramalho-Santos, M., Percharde, M., Lin, C., Yin, Y., Guan, J., et al.** (2018). Article A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity Article A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity. *Cell* 1–15.

**Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. and Mesirov, J. P.** (2011). Integrative genomics viewer. *Nat Biotechnol* **29**, 24–26.

**Rodriguez-Terrones, D., Gaume, X., Ishiuchi, T., Weiss, A., Kopp, A., Kruse, K., Penning, A., Vaquerizas, J. M., Brino, L. and Torres-Padilla, M. E.** (2018a). A molecular roadmap for the emergence of early-embryonic-like cells in culture. *Nat Genet* **50**, 106–119.

**Rodriguez-Terrones, D., Gaume, X., Ishiuchi, T., Weiss, A., Kopp, A., Kruse, K., Penning, A., Vaquerizas, J. M., Brino, L. and Torres-Padilla, M. E.** (2018b). A molecular roadmap for the emergence of early-embryonic-like cells in culture. *Nat Genet* **50**, 106–119.

**Stankiewicz, T. R., Gray, J. J., Winter, A. N. and Linseman, D. A.** (2014). C-terminal binding proteins: Central players in development and disease. *Biomol Concepts* **5**, 489–511.

**Whiddon, J. L., Langford, A. T., Wong, C. J., Zhong, J. W. and Tapscott, S. J.** (2017a). Conservation and innovation in the DUX4-family gene network. *Nat Genet* **49**, 935–940.

**Whiddon, J. L., Langford, A. T., Wong, C. J., Zhong, J. W. and Tapscott, S. J.** (2017b). Conservation and innovation in the DUX4-family gene network. *Nat Genet* **49**, 935–940.

**Wolf, G., de Iaco, A., Sun, M. A., Bruno, M., Tinkham, M., Hoang, D., Mitra, A., Ralls, S., Trono, D. and Macfarlan, T. S.** (2020). Krab-zinc finger protein gene expansion in response to active retrotransposons in the murine lineage. *Elife* **9**, 1–22.

**Yamamoto, M., Suwa, Y., Sugiyama, K., Okashita, N., Kawaguchi, M., Tani, N., Matsubara, K., Nakamura, A. and Seki, Y.** (2020). PRDM14-CtBP1/2-PRC2 complex regulates transcriptional repression during transition from primed to naïve pluripotency. *J Cell Sci* jcs.240176.

**Yang, M., Yu, H., Yu, X., Liang, S., Hu, Y., Luo, Y., Izsvák, Z., Sun, C. and Wang, J.** (2022). Chemical-induced chromatin remodeling reprograms mouse ESCs to totipotent-like stem cells. *Cell Stem Cell* **29**, 400-418.e13.

**Zhang, T., Zhang, Z., Dong, Q., Xiong, J. and Zhu, B.** (2020). Histone H3K27 acetylation is dispensable for enhancer activity in mouse embryonic stem cells. *Genome Biol* **21**, 1–7.

765 **Author Contributions**

766 Y.S. conceived the study and experimental design, performed experiments and wrote the
767 manuscript. K.S., K.Y. and N.H. performed most experiments. R.M. M.M., A.S. and Y.M
768 performed some experiments. K.S., K.Y. and N.H. contributed equally to this work.

769

770 **Acknowledgements**

771 We thank all members of Seki laboratory for stimulating discussions. This study was
772 supported by JSPS KAKENHI Grant Number JP18H02422, JP20H05375 and
773 JP22H04682.

774

775 **Conflict of interest**

776 The authors declare no competing interests.

777